Copyright © ECMS2017

Printed: ISBN: 978-0-9932440-4-9
CD: ISBN: 978-0-9932440-5-6

European Council for Modelling
and Simulation

Cover pictures
front side: © Kravka |
Dreamstime.com
back side: © Stefano Ferrario |
pixabay.com

Printed by
Digitaldruck Pirrot GmbH
66125 Sbr.-Dudweiler,
Germany

PROCEEDINGS

31st European Conference on Modelling and Simulation ECMS 2017

May 23rd – May 26th, 2017
Budapest, Hungary

Edited by:

Zita Zoltay Paprika

Péter Horák

Kata Váradi

Péter Tamás Zwierczyk

Ágnes Vidovics-Dancs

János Péter Rádics

Organized by:

ECMS - European Council for Modelling and Simulation

Hosted by:

Corvinus University of Budapest, Corvinus Business School

Budapest University of Technology and Economics

Sponsored by:

Budapest University of Technology and Economics

Central-European Training Center for Brokers / Közép-európai Brókerképző Alapítvány

Corvinus University of Budapest, Corvinus Business School

KELER CCP Ltd.

MVM Group
CLAAS Hungária Kft.
FÉMALK Zrt.

International Co-Societies:
IEEE - Institute of Electrical and Electronics Engineers
ASIM - German Speaking Simulation Society
EUROSIM - Federation of European Simulation Societies
PTSK - Polish Society of Computer Simulation
LSS - Latvian Simulation Society

ECMS 2017 ORGANIZATION

Conference Chair

Zita Zoltay Paprika

Corvinus University of Budapest
Corvinus Business School
Hungary

Conference Co-Chairs

Kata Váradi

Corvinus University of Budapest
Corvinus Business School
Hungary

Péter Tamás Zwierczyk

Budapest University of Technology
and Economics
Hungary

Programme Chair

Péter Horák

Budapest University of Technology
and Economics
Hungary

Programme Co-Chairs

Ágnes Vidovics-Dancs

Corvinus University of Budapest
Corvinus Business School
Hungary

János Péter Rádics

Budapest University of Technology
and Economics
Hungary

President of European Council for Modelling and Simulation

Khalid Al-Begain

University of South Wales, United Kingdom

Vice-President of European Council for Modelling and Simulation

Lars Nolle

Jade University of Applied Science, Wilhelmshaven, Germany

Managing Editor

Martina-Maria Seidel

St. Ingbert, Germany

EUROPEAN COUNCIL BOARD 2017

Khalid Al-Begain University of South Wales, United Kingdom	President of European Council for Modelling and Simulation and Past President of European Council for Modelling and Simulation (2006-2010)
Lars Nolle Jade University of Applied Science, Wilhelmshaven, Germany	Vice-President of European Council for Modelling and Simulation elected in 2016
Evtim Peytchev Nottingham Trent University United Kingdom	Past President of European Council for Modelling and Simulation (2012-2014) Treasurer (2015-today)
Andrzej Bargiela United Kingdom	Past President of European Council for Modelling and Simulation (2002-2004, 2004-2006, 2010-2012)
Eugène Kerckhoffs The Netherlands	Founder, Past President and Historian of ECMS
Kata Váradi Corvinus University of Budapest, Corvinus Business School Hungary	Conference Co-Chair of ECMS 2017
Peter T. Zwierczyk Budapest University of Technology and Economics Hungary	Programme Co-Chair of ECMS 2017
Frank Herrmann Ostbayerische Technische Hochschule Germany	Conference Chair of ECMS 2016
Michael Manitz Universität Duisburg-Essen Germnay	Programme Chair of ECMS 2016

INTERNATIONAL PROGRAMME COMMITTEE

Agent-Based Simulation

Track Chair: **Michael Möhring**
University of Koblenz-Landau, Germany

Co-Chair: **Ulf Lotzmann**
University of Koblenz-Landau, Germany

Simulation in Industry, Business, Transport and Services

Track Chair: **Alessandra Orsoni**
University of Kingston, United Kingdom

Co-Chair: **Edward J. Williams**
University of Michigan-Dearborn, USA

Simulation of Intelligent Systems

Track Chair: **Zuzana Kominková Oplatková**
Tomas Bata University of Zlín, Czech Republic

Co-Chair: **Roman Senkerik**
Tomas Bata University of Zlín, Czech Republic

Finance and Economics and Social Science

Track Chair: **Kata Váradi**
Corvinus University of Budapest, Corvinus Business School, Hungary

Co-Chairs:
Barbara Dömötör
Corvinus University of Budapest, Corvinus Business School, Hungary

Ágnes Vidovics-Dancs
Corvinus University of Budapest, Corvinus Business School , Hungary

Modelling, Simulation and Control of Technological Processes

Track Chair: **Jiři Vojtěšek**
Tomas Bata University in Zlín, Czech Republic

Co-Chairs:
Petr Dostál
Tomas Bata University in Zlín, Czech Republic

František Gazdoš
Tomas Bata University in Zlín, Czech Republic

Simulation and Optimization

Track Chair: **Frank Herrmann**
OTH Regensburg, Germany

Co-Chairs:
Thorsten Claus
Technical University Dresden, Germany

Michael Manitz
University of Duisburg-Essen, Germany

High Performance Modelling and Simulation

Track Chair: **Mauro Iacono**
Seconda Università degli Studi di Napoli, Italy

Co-Chairs:
Daniel Grzonka
Cracow University of Technology, Poland

Agnieszka Jakobik
Cracow University of Technology, Poland

Rostislav V. Razumchik
Institute of Informatics Problems FRC CSC RAS,
Russia

Honorary Track Chair:
Joanna Kolodziej
Cracow University of Technology, Poland

IPC Members in Alphabetical Order

Pavel Abaev, Peoples' Friendship University of Russia, Russia

Frederic Amblard, Université Toulouse 1 Capitole, France

Piotr Arabas, Warsaw University of Technology and NASK, Poland

Monika Bakosova, Slovak University of Technology in Bratislava, Slovakia

Hans-Peter Barbey, University of Applied Sciences in Bielefeld, Germany

Enrico Barbierato, Politecnico di Milano, Italy

Jan Bielanski, AGH University of Science and Technology, Poland

Vladimir Bobal, Tomas Bata University in Zlin, Czech Republic

Riccardo Boero, Los Alamos National Laboratory, USA

Aleksander Byrski, AGH University of Science and Technology, Poland

Petr Chalupa, Tomas Bata University in Zlin, Czech Republic

Emile Chappin, Delft University of Technology, The Netherlands

Adriana Chis, National College of Ireland Dublin, Ireland

Marina Chukalina, Russian Academy of Science, Russia

Franco Cicirelli, University of Calabria, Italy

Catherine Cleophas, RWTH Aachen, Germany

Gregoire Danoy, University of Luxembourg, Luxembourg

Ciprian Dobre, University Politehnica of Bucharest, Romania

František Dušek, University of Pardubice, Czech Republic

Nóra Ágota Felföldi-Szűcs, Corvinus University of Budapest, Hungary

Massimo Ficco, Seconda Università degli Studi di Napoli, Italy

Robert Forstner, SimPlan AG in Regensburg, Germany

Christopher Frantz, Otago Polytechnic, New Zealand

Amineh Ghorbani, Delft University of Technology, The Netherlands

Horacio Gonzalez-Velez, National College of Ireland, Ireland

Claudius Gräbner, (ICAE) Johannes Kepler University Linz, Austria

Marco Gribaudo, Politecnico di Milano, Italy

Alexander Grusho, Institute of Informatics Problems FRC CSC RAS, Russia

Magdalena Handerek, AGH University of Science and Technology, Poland

Stefan Hannig, ZF Friedrichshafen AG Passau, Germany

Benjamin Hildebrandt, University Duisburg-Essen, Germany

Daniel Honc, University of Pardubice, Czech Republic

Mark Hoogendorn, VU University of Amsterdam, The Netherlands

Thomas Hußlein, OptWare GmbH in Regensburg, Germany

Zsuzsa Huszár, National University of Singapore, Singapore

Teruaki Ito, The University of Tokushima, Japan

Michal Janosek, University of Ostrava, Czech Republic

Jácint Juhász, Babes-Bolyai University, Romania

Bogumil Kamiński, Warsaw School of Economics, Poland

Bart Kamphorst, VU University of Amsterdam, The Netherlands

Michał Konopczak, Warsaw School of Economics, Poland

Petia Koprinkova, Bulgarian Academy of Sciences, Bulgaria

Victor Korolev, Moscow State University, Russia

Ina Kortemeier, University Duisburg-Essen, Germany

Igor Kotenko, SPIIRAS, Russia

Martin Kotyrba, University of Ostrava, Czech Republic

Achyutha Krishnamoorthy, Cochin University of Science and Technology, India

Julia Kruk, Belarusian National Technical University, Belarus

Mateusz Krzysztoń, Warsaw University of Technology and NASK, Poland

Marek Kubalcik, Tomas Bata University in Zlin, Czech Republic

Frederick Lange, Maschinenfabrik Reinhausen Regensburg, Germany

Andrea Marin, Università Ca' Foscari di Venezia, Italy

Michal Marks, Research and Academic Computer Network (NASK), Poland

Stefano Marrone, Seconda Università degli Studi di Napoli, Italy

Agnieszka Mars, Jagiellonian University Cracow, Poland

Radek Matusu, Tomas Bata University in Zlin, Czech Republic

Nicolas Meseth, University of Applied Sciences in Osnabrueck, Germany

Christian Müller, University of Applied Sciences in Wildau, Germany

Maximilian Munninger, University of Applied Sciences in Regensburg, Germany

Tamas Nagy, Corvinus University of Budapest, Hungary

Valeriy Naumov, Service Innovation Research Institute, Finland

Catalin Negru, University Politehnica of Bucharest, Romania

Libero Nigro, University of Calabria, Italy

Dmitry P. Nikolaev, Russian Academy of Science, Russia

Igor Nikolic, Delft University of Technology, The Netherlands

Lars Nolle, Jade University of Applied Science Wilhelmshafen, Germany

Jakub Novak, Tomas Bata University in Zlin, Czech Republic

Beatrix Oravecz, Corvinus University of Budapest, Hungary

Francesco Palmieri, Università degli Studi di Salerno, Italy

Falk Stefan Pappert, Universität der Bundeswehr München, Germany

Libor Pekar, Tomas Bata University in Zlin, Czech Republic

Marc-Philip Piehl, University Duisburg-Essen, Germany

Michal Pluhacek, Tomas Bata University in Zlin, Czech Republic

Zoltan Pollak, Corvinus University of Budapest, Hungary

Florin Pop, University Politehnica of Bucharest, Romania

Simone Righi, Hungarian Academy of Sciences, Hungary

Boris Rohal-Ilkiv, Slovak University of Technology in Bratislava, Slovakia

Michael Römer, Martin-Luther-Universität Halle-Wittenberg, Germany

Juliette Rouchier, GREQAM-CNRS, France

Konstantin Samouylov, Peoples' Friendship University, Russia

Leonid Sevastyanov, Peoples' Friendship University, Russia

Oleg Shestakov, Moscow State University, Russia

Sergey Ya. Shorgin, Institute of Informatics Problems FRC CSC RAS, Russia

Markus Siegle, Universität der Bundeswehr München, Germany

Stelios Sotiriadis, University of Toronto, Canada

Grażyna Suchacka, Opole University, Poland

Katarzyna Sum, Warsaw School of Economics, Poland

Grzegorz Szafrański, Institute of Finance, Lodz University, Poland

János Száz, Corvinus University of Budapest, Hungary

Magdalena Szmajduch, Cracow University of Technology, Poland

Piotr Szuster, Cracow University of Technology, Poland

Pawel Szynkiewicz, Polish Academy of Science Warsaw, Poland

Armando Tacchella, Università degli Studi di Genova, Italy

Kristóf Tamás, Corvinus University of Budapest, Hungary

Enrico Teich, Technical University of Dresden, Germany

Marco Trost, Technical University of Dresden, Germany

Christopher Tubb, University of South Wales, United Kingdom

Tobias Uhlig, Universität der Bundeswehr München, Germany

Nikolai Ushakov, Norwegian University of Science and Technology, Norway

Ward van Breda, VU University of Amsterdam, The Netherlands

Enrico Vicario, Università degli Studi di Firenze, Italy

Andrea Vinci, CNR – National Research Council of Italy, Italy

Narayanan Viswanath, Government Engineering College Thrissur, India

Jaroslav Vitku, GoodAI, Czech Republic

Thorsten Vitzthum, University of Applied Sciences in Regensburg, Germany

Eva Volna, University of Ostrava, Czech Republic

Andrzej Wilczyński, Cracow University of Technology, Poland

Victor Zakharov, Institute of Informatics Problems FRC CSC RAS, Russia

Alexander Zeifman, Vologda State University, Russia

Bihary Zsolt, Corvinus University of Budapest, Hungary

PREFACE

It is a pleasure to welcome the more than 100 researchers at the 31st European Conference on Modelling and Simulation (ECMS) conference from May 23rd till May 26th 2017, and have their research collected in this proceedings.

The 31st ECMS conference will be hosted by Corvinus University of Budapest and Budapest University of Technology and Economics. Organizing ECMS 2017 is a great opportunity for these two long-established Hungarian universities to give rise to a fruitful cooperation – bringing the two geographically very close institutions even closer. Modelling and simulation are exciting and useful methodologies to study various problems in several research fields.

Besides the disciplines closely related to the two organizing universities (business, economics, finance, engineering, kinematics, mechanisms, mechanical simulations), we are eager to get acquainted with all the fields related to the conference. The main goal of the conference is to share your research questions and expertise with the ECMS community as well. It is always a nice experience to bring inspired researchers together. This conference is a forum to communicate and share new ideas and methods that will foster the development across all aspects of computational methods and their applications in modelling and simulation of different fields of applied science.

We hope you will find your field of interest in our conference tracks which cover exciting topics about intelligent systems, virtual prototyping, operating simulation, applied modelling, simulation and optimization, controlling of technological processes and the modelling and simulation of cases in finance, economics and social sciences.

The high prestige of the ECMS 2017 program is enhanced by our two keynote speakers, one from the field of economics, and the other from the field of engineering. István P. Székely – director of the European Commission, Economic and Financial Affairs, and honorary professor of Corvinus University of Budapest –will talk about the economic modelling and economic policy surveillance in Europe. Jin Ooi – professor of particulate solid mechanics at the University of Edinburgh – will give a lecture on discrete element modelling of granular materials.

The conference will also give you the opportunity to visit Budapest, one of the most exciting cities in Europe. In addition to a decent scientific program, we sincerely hope that you will take a chance to enjoy the culture of Budapest.

We are looking forward to welcoming you in Budapest!

Zita Zoltay Paprika, Péter Horák

Budapest, May 2017

TABLE OF CONTENTS

Plenary Talks - Abstracts

Agent-Based Simulation

Finance and Economics and Social Science

Simulation in Industry, Business, Transport and Services

Simulation of Intelligent Systems

Modelling, Simulation and Control of Technological Processes

Simulation and Optimization

High Performance Modelling and Simulation

Modelling and Simulation of Data Intensive Systems - Special Session -

Probability and Statistical Methods for Modelling and Simulation of High Performance Information Systems - Special Session -

ECMS 2017

SCIENTIFIC PROGRAM

Plenary Talks

Challenges to policy-oriented modelling and model-based policy formulation during the crises: A user perspective

István P. Székely

European Commission
and
Corvinus University of Budapest

ABSTRACT

The events of the past decade have posed unprecedented challenges to economic policy making and more broadly to the design of economic systems at the national and supranational levels. First, the global financial crisis hit, which although originating in the US financial system,, spread quickly to Europe through linkages in the global financial system. This was followed by the sovereign credit crisis in Europe, which alongside revealing a serious policy coordination problem, catalysed fundamentally changed views on the growth potential of some of the European economies. Finally, just when the European economy was settling down onto a moderate recovery path, suddenly the reality of Brexit emerged. Overall, the crisis in Europe has lasted much longer than expected and longer than in past episodes. It has taken many unexpected turns and increased already heightened uncertainty. Not surprisingly, existing models that were used to support policy formulation and assessment failed to describe most of these developments. Forecasts failed to foresee many of the events and turned out to be most imprecise when they were most needed. Policy assessment became a more difficult exercise. Moreover, predictions regarding the short-term economic impact of Brexit turned out to be rather poor, at least prior to the activation of article 50.

In many areas, policy making relies on a range of models that originate from before the 2008/2009 economic and financial crisis, both at the national and the supranational levels. These models could not predict the crisis, nor could they capture the dynamics

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

of the ensuing readjustment process. In fact, the crisis served more effectively to reveal many of the fundamental problems with these models. Problems that had existed before but had not been surfaced during the period of great moderation. The workhorse model of policy making, the New Keynesian DSGE model, could not capture the interaction between the real economy and the financial system, because its original version was built on the assumption that this interaction was broadly irrelevant. In fact, there were no banks or money in this model, and there was no credit default either. Moreover, this model was built on the assumption that the behaviour of a representative agent could describe the behaviour of an economy at the aggregate level. In order words, the heterogeneity of consumers or firms was also viewed as largely irrelevant. As some put it, DGSE models crashed when the crisis hit, just when they were most needed to understand the reaction of the economy and to formulate a policy response. Many of the policy coordination mechanisms in Europe, most importantly the Stability and Growth Pact also base themselves on some form of partial model, since fiscal policy effort is measured using the concept of potential output. While the model to determine potential output does incorporate hysteresis, and it did capture the increase in the NAWRU, it was still challenged in many ways by the crisis.

The keynote speech will review the ways in which the crisis has had important implications for these models and the forms in which model builders and users have reacted to such challenges. Models have been improved in several ways to address the problems that had been revealed. Nevertheless, further work is needed to ensure that the models that national and supranational policy makers rely on provide useful input into the process of policy formulation.

Discrete element modelling of cohesionless, cohesive and bonded granular materials - from model conceptualisations to industrial scale applications

Jin Y. Ooi

Institute for Infrastructure & Environment,
School of Engineering,
University of Edinburgh, U.K.
j.ooi@ed.ac.uk

ABSTRACT

Handling and processing of bulk granular materials is of major importance in many industries including agriculture, chemical, food, pharmaceutical, construction and mining. In recent decades, the Discrete Element Method (DEM) is increasingly adopted for studying the complex behaviour of granular materials and simulating the industrial processes in many applications. DEM computes at individual particle interaction level and by incorporating relevant inter-particle forces and coupling with hydrodynamic forces of any surrounding fluid in particle-fluid systems, it provides important insights into the particle level phenomena – this in turn inform the bulk and industrial level processes.

In this keynote lecture, the developments of three DEM contact models to model: a) free flowing cohesionless solids; b) cohesive fine powders and c) cementitious materials are described. In particular, the meso-scale approach of modelling cohesive material is proposed using a visco-elasto-plastic frictional adhesive contact model developed recently at the University of Edinburgh [1,2]. The cohesive model reflects the physical phenomena of adhesive contact forces in fine cohesive particles and accounts for both elastic and plastic contact deformation with the adhesion being dependent on contact plasticity. Also a bonded contact model based on the

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Timoshenko beam theory which can be used to model concrete, rock and other deformable-breakable materials will also be presented [3].

The suitability of these contact models to predict real physical behaviours will be discussed. The scaling laws of the contact model parameters to produce the same load-deformation behaviour invariant of the particle size used in the simulations are presented [4,5]. Averaging (coarse-graining) technique based on statistical mechanics [6] is applied to the DEM particle data to provide important insights into the mobilized stress field during an industrial process. The studies demonstrate successful applications of DEM simulations of large scale applications using DEM models with appropriate scaling laws and material characterization experiments.

References:

1. S. Thakur, J. Morrissey, J. Sun, J. Chen, J.Y. Ooi, (2014) “Micromechanical analysis of cohesive granular materials using the discrete element method with an adhesive elasto-plastic contact model”, *Granular Matter* 16, 383 – 400
2. S. Thakur, H. Ahmadian, J. Sun, J.Y. Ooi, (2014) “An experimental and numerical study of packing, compression, and caking behaviour of detergent powders”, *Particuology* 12, 2–12.
3. Brown, N., Ooi, J.Y., Chen, J.F. (2014) “A bond model for DEM simulation of cementitious materials and deformable structures” Granular Matter, 16:299–311.
4. A. Janda and J.Y. Ooi (2016) “DEM modelling of cone penetration and unconfined compression in cohesive solids”, *Powder Technol.* 293, 60–68.
5. S. Thakur, J.Y. Ooi and H. Ahmadian (2016) “Scaling of discrete element model parameters for cohesionless and cohesive solid”, *Powder Technol.* 293, 130–137.
6. T. Weinhart, C. Labra, S. Luding, J.Y. Ooi (2016) “Influence of coarsegraining parameters on the analysis of DEM simulations of silo flow”, *Powder Technol.* 293, 138–148.

Agent-Based Simulation

STATISTICAL MODEL CHECKING OF MULTI-AGENT SYSTEMS

Libero Nigro, Paolo F. Sciammarella

Software Engineering Laboratory
University of Calabria, DIMES - 87036 Rende (CS) – Italy
Email: l.nigro@unical.it, p.sciammarella@dimes.unical.it

KEYWORDS
Statistical Model Checking, Multi-agent systems, Actors, UPPAAL SMC, Iterated Prisoner's Dilemma.

ABSTRACT

This paper proposes an original approach to modelling and simulation of multi-agent systems which is based on statistical model checking (SMC). The approach is prototyped in the context of the popular UPPAAL SMC toolbox. Usefulness and validation of the approach are checked by applying it to a known complex and adaptive model of the Iterated Prisoner's Dilemma (IPD) game, by studying the emergence of cooperation in the presence of different social interaction structures.

INTRODUCTION

In the last years multi-agent systems (MAS) have proved to be a fundamental paradigm for modelling and simulation (M&S) of complex and adaptive systems (North & Macal, 2007)(Cicirelli & Nigro, 2016b). Power and flexibility of MAS stem from the ability of modeling individual behaviour of agents and their social interactions, and then to observe the emergence of properties at the population level.

In this work a minimal actor computation model (Cicirelli & Nigro, 2006a) is adopted for supporting MAS. A unique feature of this model is a light-weight notion of actors which (i) are threadless agents, (ii) hide an internal data status, and (iii) communicate to one another by asynchronous message passing. Message exchanges are ultimately regulated by a customizable *control structure* which can reason on time (simulated or real-time). The model can be effectively hosted by widespread languages like Java.

Novel in this paper is a mapping of the actor model onto UPPAAL SMC (David *et al*., 2015) so as to exploit statistical model checking techniques (Younes *et al*, 2006)(Larsen & Legay, 2016). SMC automatizes multiple executions, estimates a required number of simulation runs, uses statistical properties (e.g., Monte Carlo-like simulations and sequential hypothesis testing) to infer system properties from the observables of the various runs.

To the best of authors knowledge, no similar support in UPPAAL SMC of MAS with asynchronous messages was previously published. Originality is mainly tied to an exploitation of dynamic message template processes which are not supported by other tools.

The resultant approach is practically demonstrated through a case study concerned with a complex, adaptive and scalable model based on the Iterated Prisoner's Dilemma (IPD) game (Axelrod *et al*., 2002). The model is very challenging and it aims to study the emergence of cooperation among competitive agents when varying, e.g., the social interaction network. SMC results confirm previous indications of Axelrod et al. work, although with a smaller cooperation degree.

The paper is structured as follows. First, the actor computational model is summarized. Then major features of UPPAAL SMC are recapitulated. After that, a structural mapping of actors onto UPPAAL SMC is proposed. The paper goes on by detailing the developed IPD case study and the achieved experimental results. Conclusions are finally presented by stating on-going and future work.

ACTOR COMPUTATIONAL MODEL

Actors (Agha, 1986) with asynchronous message passing are a reference model for concurrent and distributed systems. Many variants of this model, though, motivated by specific uses and application contexts, are developed in the literature for both theory and/or practice reasons.

In this work a light-weight (threadless) and control-based version of the actor model is adopted (Cicirelli & Nigro, 2013)(Cicirelli & Nigro, 2016a), addressing specifically high-performance time-dependent applications. A system is a federation of theatres. A theatre (Logical Process or LP), allocated for execution on a computing node, hosts a collection of actors plus a control machine (CM) and interface to a transport layer. The goal of CM is to transparently manage the cloud of exchanged messages and to deliver them, one at time and in an interleaved way, according to a control structure. Message processing is atomic (*macro-step* semantics). As a consequence, a cooperative concurrency schema among the agents of a theatre is ensured. The CMs of a system synchronize one to another in order to ensure a coherent global time notion. Details about control structure design and global synchronization algorithms based on a time server are described in (Cicirelli *et al*., 2016a). In the following the focus will be on a single theatre and on the assumed actor programming style, e.g., in Java.

Actors are objects of classes which inherit from the abstract Actor base class which exposes, among others, the basic non blocking message *send* operation. An actor owns some local data variables and acquaintances (known actors) to whom messages can be sent (including itself for proactivity).

To simplify modelling/programming, message handling is split among separate methods which can have

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

parameters. Each method represents a “message server” (*msgsrv*) which receives and handles the specific associated message. When a msgsrv gets actually invoked it depends ultimately on a decision of the control machine. No mailbox is provided per actor, rather all the sent messages get buffered in a cloud managed by the control machine.

Figure 1 clarifies the actor programming style through a simple ping-pong application. A main program (not shown for brevity) creates the two actors, sends them an *init* message carrying the identity of the partner (acquaintance) and then sends, e.g., a first *ping* message to the ping actor.

```
public class Ping extends Actor{       public class Pong extends Actor{
  protected Pong pong;                   protected Ping ping;
  public void init(Pong pong){           public void init(Ping ping){
     this.pong=pong;                        this.ping=ping;
  }//init                                }//init
  public void ping(){                    public void pong(){
     System.out.println("ping");            System.out.println("pong");
     pong.send( "pong" );                   ping.send( "ping" );
  }//ping                                }//pong
}//Ping                                }//Pong
```

Figure 1 – Ping Pong actors

The send operation can attach a relative timestamp to a message: $target.send(delay, msg - name, args)$, where delay can be established by sampling a probability distribution function or it can be deterministic. For the timed send, the msgsrv $msg - name$ with $args$ (an array of objects) will be invoked on the target actor after delay time units are elapsed from current time (returned by the $now()$ service). When the timestamp is missing, it defaults to 0.

A library of control machines ranging from pure concurrency (time insensitive), to simulated time and real time was developed. Actors do not know the identity of the regulating control machine.

CONCEPTS OF UPPAAL SMC

UPPAAL (Behrmann *et al.*, 2004) is a popular toolbox for modelling and analysis of real-time systems. A model is a network of timed automata (TA).

TA are modelled as *template processes*, which can have parameters, can be instantiated, and consist of *locations*, *edges* and *atomic actions*. TA synchronize to one another by CSP-like channels (*rendezvous*) which carry no data values. Asynchronous communication is provided by broadcast channels. The sender of a broadcast channel in no case blocks. The synchronization can be heard by 0, 1 or multiple receivers. Clock variables allow measuring the time elapsed from a given instant (clock reset). Locations of an automaton are linked by edges. Every edge can be annotated by a command with four (optional) elements: (i) a *non deterministic selection*, (ii) a *guard*, (iii) a *synchronization* (? for input and ! for output) on a channel, and (iv) an *update* consisting of a set of clock resets and a list of variable assignments. A clock *invariant* can be attached to a location as a *progress* condition. The automaton can stay in the location provided the invariant is not violated. *Committed* and *urgent* locations which must be exited without passage of time, are also supported. Committed locations have priority w.r.t. urgent locations. TA models can greatly benefits, in latest versions of the toolbox, by the use of C-like functions.

The *symbolic model checker* of UPPAAL handles the parallel composition of the TA of a model, i.e., all the possible action interleavings are considered. For exhaustive property assessment the verifier tries to build the model state graph.

The problem with symbolic model checking is that it could not be practically applied to realistic complex systems which can imply an enormous (possible infinite) or a not decidable (e.g., continuous time and stochastic behavior are combined) state graph. Property checking in these cases can only be approximated or estimated.

In recent years the UPPAAL toolbox was extended to support statistical model checking (Younes, 2006)(David *et al.*, 2015). UPPAAL SMC avoids the construction of the state graph and checks properties by performing a certain number of simulation runs. After that some statistics techniques are used to infer results from the simulation runs. SMC refines and extends basic UPPAAL. Only broadcast synchronizations are allowed among stochastic TA (STA). Stopwatches, floating point (double) variables which can be assigned the value of a clock, and *dynamic templates* which can be instantiated and terminated at run time are supported too. On a stochastic TA model the following query types can be issued (meta-symbols **(** and **|** and **)** are in bold).

1. simulate N [**(**clock**|**#**|**void**)**<=bound] {Expr1,…, Exprk}
2. Pr[**(**clock**|**#**|**void**)**<=bound] (**(**<>**|**[]**)** Expression)
3. Pr[**(**clock**|**#**|**void**)**<=bound] (**(**<>**|**[]**)** Expression) **(**<=**|**>=**)** PROB
4. Pr[**(**clock**|**#**|**void**)**<=bound] (**(**<>**|**[]**)** Expression) **(**<=**|**>=**)**
   Pr[**(**clock**|**#**|**void**)**<=bound] (**(**<>**|**[]**)** Expression)
5. E[**(**clock**|**#**|**void**)**<=bound; N] (**(**min:**|**max:**)** Expression)

Expressions can specify an automaton is in a certain location, or some constraints on data variables or clocks etc. All the queries are evaluated according to a bound which can be related to the (implicit) global time or to a clock or to a number of simulation steps (#). Query 1 asks N simulations and collects information about the listed expressions. Query 2 evaluates the probability the given expression holds (<>) within the assigned bound or always holds within the bound ([]) with a confidence interval. Query 3 checks if the estimated probability is lesser/greater than a given probability value. Query 4 compares two probabilities. Query 5, finally, estimates the minimum/maximum value of an expression.

Quantitative estimation of a query of type 2 rests on Monte Carlo-like simulations. Qualitative queries of the type 3 and 4 use sequential hypothesis testing, and precompute a required number of runs. An important feature provided by UPPAAL SMC is *visualization* of simulation output. Following a satisfied query, the modeler can right click on the executed query and choice an available diagram (histogram, probability distribution etc.) to be plotted. At the time of this writing, UPPAAL SMC is supported by the development version 4.1.19.

MAPPING ACTORS ONTO UPPAAL SMC

The translation of actors onto UPPAAL SMC relies on dynamic template processes. A dynamic automaton must be announced in the global declarations thus:

```
dynamic tName( params ); //only int params
```

and its behavior specified as for normal TA. The dynamic automaton can then be instantiated in the update of a command with a *spawn* expression:

```
spawn tName( args );
```

Similarly, it can be terminated by an *exit()* expression in the update of a command in the tName template process.

In the following, dynamic templates will be used only for messages, although they could also be exploited for creating/destroying actors dynamically. Let aid be an int type range for the agent unique identifiers, and msg_id a type range for the messages unique identifiers, in the MAS model. An array of broadcast channels corresponding to all the possible message servers in the model is then introduced:

```
broadcast chan msgsrv[aid][msg_id];
```

Two typical examples of message templates, respectively instantaneous and timed, are shown in the Figures 2 and 3 (see also the case study later in this paper):

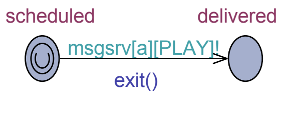


Fig. 2 – Immediate message

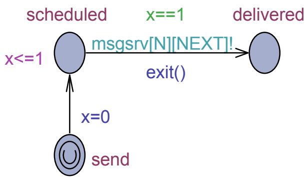


Fig. 3- Timed message

Message automata starts in a urgent location and admits two conceptual locations: *scheduled* and *delivered*. The *scheduled* can be time-sensitive. In Fig. 3 the message cannot be delivered before 1 time unit is elapsed from the send time. In Fig. 2 the scheduled message must be immediately delivered. Delivering is achieved by sending a synchronization on the msgsrv channel corresponding to the destination actor and the involved message id.

An actor automaton (see, e.g., Fig. 5) receives a message server invocation from a normal location, and then processes it through (in general) a cascade of committed locations which ends in a normal location too. Since UPPAAL SMC requires input determinism, that is, only one message an actor can receive at a time, Fig. 4 sketches a *message reception schema* which can be adopted. The actor a in the *receive* location can get one message from $m1$, $m2$, …, mn. First a non deterministic selection is performed on the message id, immediately followed by a committed location which identifies the particular received message ID and starts its corresponding processing actions.

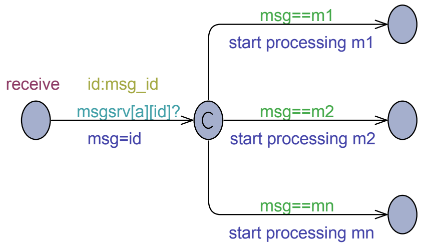


Fig. 4 – Input determinism during message reception

As a consequence of the above design: (a) the control structure of the cloud of sent messages is automatically realized by UPPAAL SMC through the activation/deactivation of dynamic message templates; (b) message processing in an actor automaton is guaranteed to terminate *before* any new message server can be delivered thus ensuring the macro step semantics.

A CASE STUDY USING THE ITERATED PRISONER'S DILEMMA

The Prisoner's Dilemma (PD) (Axelrod, 1984) is a binary game in which two players have to decide independently and without any form of communication, between two alternative choices: to defect (D, e.g., 0) or to cooperate (C, e.g., 1). The decision implies that each player gets a payoff as follows: $(D,D) \to (P,P)$, $(D,C) \to (T,S)$, $(C,D) \to (S,T)$, $(C,C) \to (R,R)$, where P means punishment for mutual defect, T temptation to defect, S sucker's payoff, and R reward for mutual cooperation. Classically $T > R > P >$S and $R > (S+T)/2$. Common adopted values are $S=0$, $P=1$, $R=3$, $T=5$.

Under the uncertainty of partner decision, players acting rationally direct themselves to defect in order to optimize their payoff, with (D,D) being the Nash equilibrium of the game. Indeed players spontaneously are driven by selfish behavior due to suspect about the opponent decision. In this situation it would be extremely risky to decide C, in fact if the partner choses D, the first player would achieve a 0 payoff and the partner the maximum reward of 5.

But if the one shot game admits the only outcome of (D,D), things are not determinate in the case the game is long iterated, with the number of iterations being unknown to players. The Axelrod book (Axelrod, 1984) triggered much interest in the social science toward studying conditions under which cooperation can emerge.

In the basic Iterated Prisoner's Dilemma (IPD) (see Axelrod tournaments (Axelrod, 1980a-b)), a certain number of players N, each equipped with a suitable strategy, repeatedly plays in turn with each of the other $N-1$ partners and the payoff is accumulated so as to detect some dominant strategy. Each player has memory of what the opponent did in the previous move. The winner of the first tournament was the strategy Tit-for-Tat (TFT) proposed by Anatol Rapaport. TFT cooperates on the first move (i.e., it is a nice strategy) and then mimics the opponent decision taken in the previous move. Also in the second Axelrod tournament, with more competing strategies, TFT emerged as the winner strategy, but in addition the experiment revealed that "altruistic" strategies instead of "selfishness" and "greedy" behavior, in the long time can do better toward cooperation, particularly if strategies can evolve and adapt, thus learning from the experience, during the iterated game.

In (Cohen *et al.*, 1999)(Axelrod *et al.*, 2002) the IPD was studied from a different perspective, to investigate the role of a social interaction network upon player behavior. In particular, the goal was to check the influence of link persistence (also said context preservation) on the emergence of cooperation, in the presence of learning and evolution of the strategies. The study confirms cooperation is possible under link persistence.

The (Axelrod *et al.*, 2002) complex and adaptive system was chosen as a challenging test bed for the approach described in this paper.

Case study description

The case study consists of a time step simulation of a large MAS of $N = 256$ players, where each agent plays PD with four neighbors whose identity varies with the adopted interaction network. Three cases are investigated:

- (PTG) a persistent toroidal $16x16$ grid with neighbors established according to the Manhattan neighborhood (NEWS – North, East, South and West);
- (PRN) a persistent random network, where neighbors are established randomly once at the start of the each run;
- (TRN) a temporary random network, where neighbors are defined at each step (also said period).

At the beginning of each simulation run, each player is assigned a strategy (y, p, q) of three probability values in [0,1], where y is the probability of choosing C at the first period, p is the probability of choosing C when the partner's last move was C, and q is the probability of choosing C when the partner's last move was D. The space of strategies includes the binary strategies $ALL-C$ $(y = p = q = 1)$, TFT $(y = p = 1, q = 0)$, anti TFT ($aTFT$: $y = p = 0, q = 1$) and $ALL - D$ $(y = p = q = 0)$. However, the model initially configures the population of shuffled agents by an even distribution of strategies where $y = p$ and p and q can assume the sixteen probability values in the vector $[1/32, 3/32, \ldots, 31/32]$.

At each period, each player plays 4 times the PD game separately with each of its neighbors, and the payoff is accumulated (and finally normalized) and the last move recorded, move by move, for both the player and its neighbors.

At the end of each period, following the PD moves, each player A adapts its behavior by copying (imitation) the strategy of the best performing neighbor (say it B), would the payoff of B be strictly greater than the period payoff of A. In addition, since the adaptation process can realistically be affected by errors (a comparison error can occur during the selection of the best performing neighbor, and a copying error can introduce a noise during the copying process) the following hypothesis are made. At each adaptation time, there exists a 10% chance the comparison between A and B payoffs is wrong performed and the best payoff misunderstood. Moreover, even in the case the strategy of A was not replaced with that of B, there is 10% chance that each "gene" of A strategy, i.e., the parameters y, p, q, be affected by a Gaussian noise with mean 0 and standard deviation 0.4.

The main goal of the case study is to monitor the fitness of the model vs. time, using a number T of 2500 steps or periods. First the average payoff per period is determined by adding all the period payoff of players and dividing the total period payoff by the population size N. Then the fitness is extracted by accumulating, at each time t, all the population average payoffs up to t, and dividing this sum by t. Other observables are the average values at each time of the probabilities p and q, averaged over all the population, so as to monitor the trend of strategy adaptation. Of course, a fitness value definitely moving toward 1 mirrors the emergence of defection, whereas its tendences to 3 (actually to a value greater than 2) testifies cooperation. The above described observables must be checked in all the possible model configurations.

Multi-agent model in UPPAAL SMC

Two basic agents were introduced: the $Manager$ (one instance) who is in charge of enforcing the time stepped simulation mode, and the $Player$ (N instances) who contains the details of the IPD game, along with global declarations which include the following:

```
//model global declarations
const int N=256;
const int dim=16; //sqrt(N)
typedef int[0,N-1] pid; //player id
typedef int[N,N] mid; //manager id
typedef int[0,N] aid; //agent id
const int MSG=5; //number of distinct message ids
typedef int[0,MSG-1] msg_id;
broadcast chan msgsrv[aid][msg_id];
clock now; //simulation time
```

Manager admits INIT, NEXT and DONE message ids, whereas Player can receive an INIT, PLAY or ADAPT message. NEWS links are randomly interpreted when a TRN or PRN topology is used instead of PTG. The following is the set of used dynamic messages. Each automaton (see Fig. 2 and Fig. 3) delivers to an actor a specific message id through the msgsrv array of channels. Except for the Next automaton (Fig. 3), all the others are instantaneous messages whose model follows that of Fig. 2.

```
//dynamic message declarations
dynamic InitManager();
dynamic Next();
dynamic Done();
dynamic InitPlayer( const pid a );
dynamic Play( const pid a );
dynamic Adapt( const pid a );
```

Fig. 5 depicts the Player automaton (whose only parameter is $const\ pid\ a$). A player first gets an INIT message then it waits for a PLAY message to which it responds with a Done to Manager. After that the player expects an ADAPT message to which it answers with another Done message to Manager. Functions $init\_player()$, $do\_play()$ and $do\_adapt()$ respectively prepare and implement the game. For instance, in the PTG case, init_player() identifies the four neighbors NEWS which persist along all the simulation. The function also establishes the strategy assigned to the player. do_play() realizes the 4 moves with each of its neighbors, accumulates the payoff, and stores the last move for player a and for each of its neighbors. Two versions of do_adapt() were built: the first one ignores any error during the adaptation process; the second one considers probabilistic comparison and copy errors as stated in the case study description.

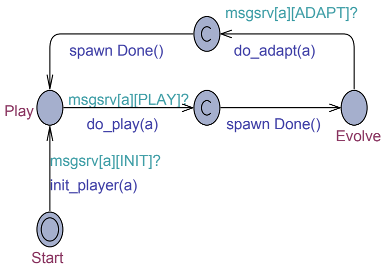


Figure 5 – $Player$ automaton

Fig. 6 shows the Manager automaton (which admits the parameter $const\ mid\ m$). The Manager first receives an INIT from the Main automaton (see Fig. 7), which causes a INIT message to be sent to all the players, then it sends proactively to itself a NEXT message with constant delay of 1 time unit. On receiving the NEXT message, the Manager enters its basic cycle: first data structures are reset for the next period, then a PLAY message is sent to all players, followed by N DONE replies to be received. When all the DONE messages arrived, the Manager goes on by sending an ADAPT message to all the players, after that N DONE replies are awaited. Finally, a NEXT message is sent again to itself for triggering the next cycle and the story recommences.

It is important to note that scheduled groups of instantaneous (concurrent) messages, such as INIT, PLAY or ADAPT to players, are actually delivered in a non deterministic way. This is a direct consequence of the asynchronous spawning mechanism and the use of urgent locations in dynamic message templates, which are exited in an interleaved arbitrary way. This same non determinism is a key for achieving actor shuffling and even distribution of all the pairs of p,q probabilities.

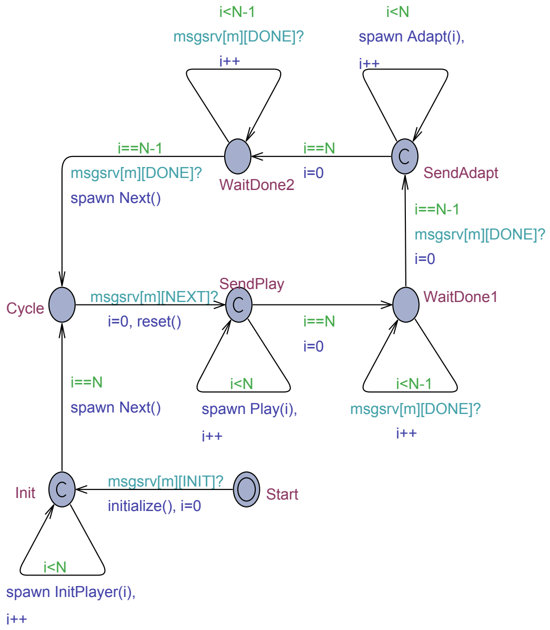


Figure 6 – $Manager$ automaton

Model bootstrap is ensured through the $Main$ automaton (see Fig. 7) which only sends the INIT message to the Manager by spawing the InitManager message template.

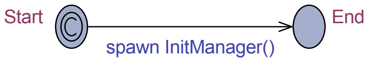


Figure 7 – $Main$ automaton

System configuration relies on implicit template instantiation and is composed as follows:

$system\ Main, Manager, Player;$

One instance of Main (anonymous), one instance of Manager (with id N) and N instances of Player (with ids from 0 to $N-1$) are created.

Implicit system (simulated) time increases with timed NEXT message delivery. The decoration clock now, initialized to 0 (default) and automatically advanced, was introduced so as to make explicit the simulation time to statistic functions such as $avg\_payoff()$, $fitness()$, $avg\_p()$, $avg\_q()$ etc., which depend also on some further decoration variables like $totpayoff$ (double).

EXPERIMENTAL RESULTS

Some preliminary experiments were carried out for observing the shape of the average period payoff in the first few time steps. The query:

$$simulate\ 1\ [<=50]\ \{\ avg\_payoff()\ \}$$

was used. Results for the PTG model without adaptation errors are shown in Fig. 8. Basic behavior holds with or without adaptation errors and also for PRN and TRN topologies. Fig. 8 confirms the indications in (Cohen *et al.*, 1999) at page 24 and page 42. The average payoff starts at 2.25 then sharply decreases, after which it will tend to a final possible regime. The initial value is due to an equivalent average strategy (y, p, q) of $(.5, .5, .5)$ being randomly initially distributed. The sharp decline is due to the presence of akin $ALL-C$ strategies which play with akin $ALL-D$ strategies. As a consequence, $ALL-D$ tends to dominate, but as $ALL-D$ plays with other $ALL-D$ it causes a sudden decrease in the payoff.

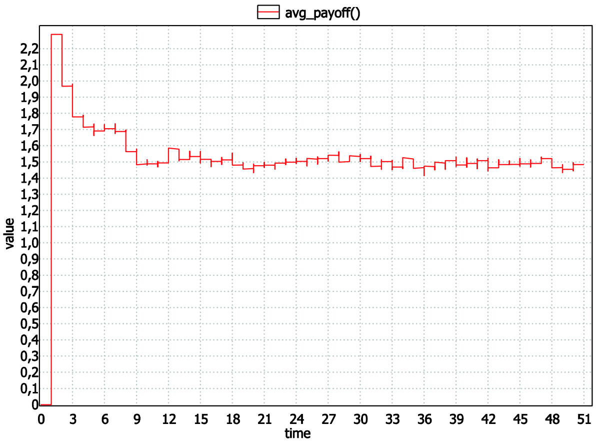


Figure 8 – Average payoff in the first 50 steps – no errors

The three models PTG, PRN and TRN were repeatedly studied by using the following query:

$$simulate\ 1\ [\leq 2500]$$
$$\{\, avg\_payoff(), fitness(), avg\_p(), avg\_q()\, \}$$

Figures from 9 to 11 depict the observed average period payoff, the temporal average $fitness$, and the average p value and q value for the three models, in the most general case when adaptation errors are involved. In Fig. 12 it is reported the watched TRN behavior without adaptation errors.

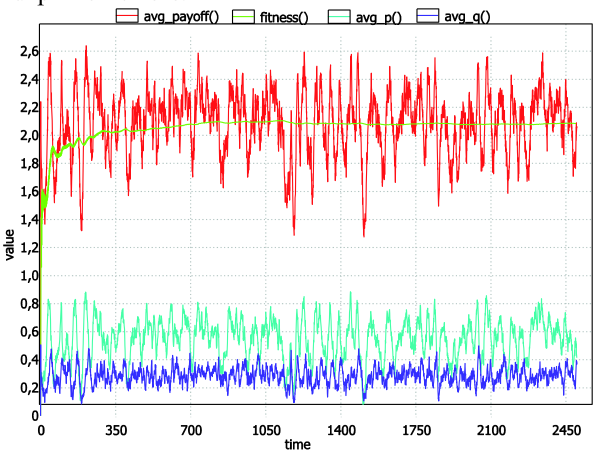


Figure 9 - PTG sampled behavior with adaptation errors

In the absence of adaptation errors, no new strategies can dynamically be introduced. Rather, some strategies can dominate whereas others can reduce its number or even disappear from the game. In these scenarios, fluctuations of the payoff tend soon to stabilize (see Fig. 12) and cooperation can or cannot possibly occur. All depends from the initial random assignment of strategies to shuffled players, i.e., who plays with who initially. However, in the TRN model, where neighbors are redefined at each time step, in no case cooperation was observed.

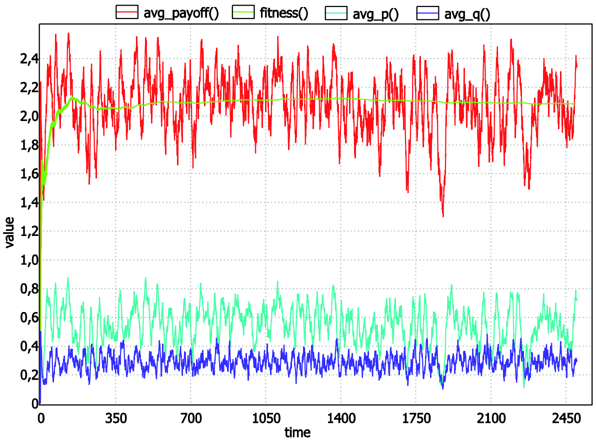


Figure 10 - PRN sampled behavior with adaptation errors

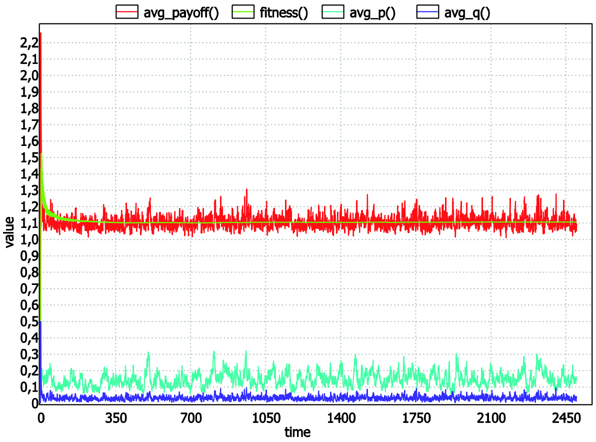


Figure 11 - TRN sampled behavior with adaptation errors

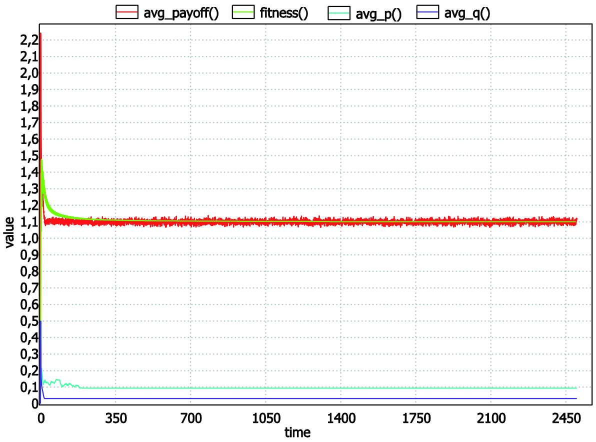


Figure 12 - TRN sampled behavior without adaptation errors

Under the presence of adaptation errors, new strategies can be created at run time and persistent topologies (PTG and PRN) manifest a more evident character to host a cooperative regime (Fig. 9 and Fig. 10). In the TRN model, instead, defection prevails (Fig. 12). The next step was to check the emergence and maintaining of a cooperative regime in the three models with errors, by a query like the following:

$$Pr[<= 800]\ (\ []\ (now < 500\ ||\ fitness() > 2.0)\ )$$

which asks to estimate through a number of runs the probability that after 500 steps a cooperation regime is eventually reached. Given the low number of time steps, it was inconclusive for PTG and PRN. For TRN, after 36 runs, Uppaal SMC suggested Pr is in the interval $[0, 0.0973938]$ with a confidence degree of 95%, which testifies the attainement of a defection regime.

The achieved experimental results agree with the results documented in (Axelrod *et al.*, 2002), although with a lesser value of the cooperation level. Results were also validated by porting the IPD models in Java. Transportation was facilitated by the Uppaal SMC formal approach. Fig. 13 depicts the observed payoff of the three topologies (with adaptation errors) averaged after 30 runs of the Java models.

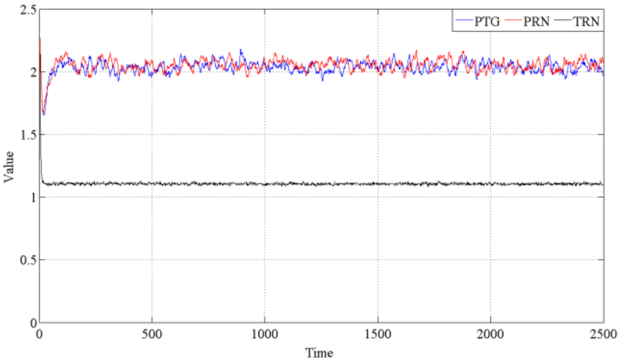


Figure 13 – Average payoff after 30 runs

From the Java models it emerged, after 100 runs, that beyond a threshold of 2000 time steps, both PTG and PRN reach and maintain the cooperative regime with a probability value of almost 1.

The experimental work confirms the intuition that link persistence (i.e., playing with the same partners during all the game) is a key for cooperation because it favours players trustiness. All of this has an obvious social interpretation nowadays when one considers people interactions through a social network.

Experiments were carried out on a Linux machine, Intel Xeon CPU E5-1603@2.80GHz, 32GB, using Uppaal 4.1.19 64bit.

CONCLUSIONS

This work develops an original approach in Uppaal SMC (David *et al.*, 2015), which enables modelling and analysis of complex adaptive asynchronous multi-agent systems (MAS). Benefits of the approach are formal modelling and the exploitation of statistical model checking techniques (Larsen & Legay, 2016).

For demonstration purposes, a scalable version of the Iterated Prisoner's Dilemma (IPD) game (Axelrod *et al.*, 2002) was modelled and thoroughly experimented. The goal was not to add to the theory of IPD, but only using a particular version of it as a benchmark. Practical limitations of the approach relate to the MAS model size, which can imply very long execution times. Prosecution of the research is directed to:

- Porting the approach in the Plasma Lab tool (Plasma Lab, on-line) so as to integrate scalable Java-based actor models for efficient SMC analysis;
- Adapting the IPD models toward an investigation of new player strategies;
- Extending the Uppaal SMC approach to modelling and analysis of distributed probabilistic real-time actor systems.

REFERENCES

Agha, G. (1986). *Actors: a model of concurrent computation in distributed systems*. MIT Press, Cambridge, MA.

Axelrod, R. (1980a) Effective choice in the prisoner's dilemma, *J. of Conflict Resolution*, **24**, pp. 3-25.

Axelrod, R. (1980b) More effective choice in the prisoner's dilemma, *J. of Conflict Resolution*, **24**, pp. 379-403.

Axelrod, R. (1984). *The evolution of cooperation*. Basic Books, New York.

Axelrod, R., R.L. Riolo, M.D. Cohen (2002). Beyond geography: cooperation with persistent links and in the absence of clustered neighborhoods. *Personality and Social Psycology Review*, **6**(2):341.346.

Behrmann, G., A. David, K.G. Larsen (2004). A tutorial on Uppaal. In: *Formal Methods for the Design of Real-Time Systems*, Lecture Notes in Computer Science, Vol. 3185, Springer-Verlag, pp. 200-236.

Cicirelli, F., L. Nigro (2013). An agent framework for high performance simulations over multi-core clusters. *Communications in Computer and Information Science (CCIS)*, Volume 402, pp. 49-60, Springer.

Cicirelli, F., L. Nigro (2016a). Control centric framework for model continuity in time-dependent multi-agent systems. *Concurrency and Computation: Practice and Experience*, **28**(12):3333-3356, Wiley.

Cicirelli, F., L. Nigro (2016b). Exploiting social capabilities in the minority game. *ACM Trans. on Modeling and Computer Simulation*, **27**(1), article 6, DOI: http://dx.doi.org/10.1145/2996456.

Cohen, M.D., R.L. Riolo, R. Axelrod (1999). The emergence of social organization in the Prisoner's Dilemma: how context-preservation and other factors promote cooperation. *Santa Fe Institute Working Paper*: 1999-01-002.

David, A., K.G. Larsen, A. Legay, M. Mikucionis, D.B. Poulsen (2015). Uppaal SMS Tutorial. *Int. J. on Software Tools for Technology Transfer*, Springer, **17**:1-19, 06.01.2015, DOI 10.1007/s10009-014-0361-y.

Larsen, K.G., A. Legay (2016). On the power of statistical model checking. In *7th Int. Symposium, ISoLA 2016*, pp. 843-862.

North, M.J., C.M. Macal (2007). *Managing business complexity-Discovering strategic solutions with Agent-based Modeling and Simulation*. Oxford University Press.

Plasma Lab, on-line, https://project.inria.fr/plasma-lab/.

Reynisson, A.H., M. Sirjani, L. Aceto, M. Cimini, A. Jafari, A. Ingolfsdottir, S.H. Sigurdarson (2014). Modelling and simulation of asynchronous real-time systems using timed Rebeca. *Science of Computer Progr.*, vol. 89, pp.41-68.

Varshosaz, M., R. Khosravi (2012). Modelling and verification of probabilistic actor systems using pRebeca. *LNCS 7635*, pp. 135-150, Springer.

Younes, H.L.S., M. Kwiatkowska, G. Normaln, D. Parker (2006). Numerical vs. statistical probabilistic model checking. *Int. J. on Software Tools for Technology Transfer*, vol. 8, no. 3, pp. 216-228.

DRIVING BEHAVIOUR CLUSTERING FOR REALISTIC TRAFFIC MICRO-SIMULATORS

Alessandro Petraro, Federico Caselli, Michela Milano
Department of Computer Science and Engineering
University of Bologna
viale Risorgimento 2, Bologna, Italy
alessandro.petraro2@studio.unibo.it, f.caselli@unibo.it,
michela.milano@unibo.it

Marco Lippi
Department of Sciences and Methods for Engineering
University of Modena and Reggio Emilia
via Amendola 2, Reggio Emilia, Italy
marco.lippi@unimore.it

KEYWORDS

Agent-based modelling; Traffic Micro Simulators; Clustering Algorithms

ABSTRACT

Traffic simulators are effective tools to support decisions in urban planning systems, to identify criticalities, to observe emerging behaviours in road networks and to configure road infrastructures, such as road side units and traffic lights. Clearly the more realistic the simulator the more precise the insight provided to decision makers. This paper provides a first step toward the design and calibration of traffic micro-simulator to produce realistic behaviour. The long term idea is to collect and analyse real traffic traces collecting vehicular information, to cluster them in groups representing similar driving behaviours and then to extract from these clusters relevant parameters to tune the micro-simulator. In this paper we have run controlled experiments where traffic traces have been synthetized to obtain different driving styles, so that the effectiveness of the clustering algorithm could be checked on known labels. We describe the overall methodology and the results already achieved on the controlled experiment, showing the clusters obtained and reporting guidelines for future experiments.

INTRODUCTION

Vehicular mobility is a complex man-made socio-technical system involving the road and communication infrastructures, road side units and drivers. Vehicular mobility affects many aspects of our everyday life, shaping the environment around us, underpinning economic growth and affecting our health and quality of life. If we restrict ourselves to urban mobility, it accounts for more than 40% of CO_2 emissions and more than 70% of other pollutants from transport[1]. Understanding mobility patterns, identifying criticalities in such a complex system, assessing the performance of the road network in specific areas is essential for decision makers that have to manage the system and plan interventions to improve the infrastructure.

Important tools that support such decision making process are traffic simulators, namely mathematical or agent-based models mimicking the traffic dynamics on a road network. They are classified in macro- and micro-simulators (Helbing et al., 2002). Macro-simulators are often based on traffic flows and describe the collective behaviour of vehicle dynamics in terms of vehicle density and average speed in time. Micro-simulators, instead, model the single vehicle and account for the driver behaviour that influences the speed and position of the vehicle in time. Thus, in micro-simulators, traffic patterns emerge by the interaction of each vehicle dynamics. It is extremely difficult to model single drivers and obtain a realistic emerging behaviour.

To provide useful support to decision makers, traffic simulators should be realistic and model real traffic patterns. It is clear that if the drivers are modelled in an unrealistic way, many emerging and realistic patterns are lost, while unrealistic patterns arise.

In this paper we propose a method for configuring realistic micro-simulators and we introduce the first results achieved in this direction. The idea is to use real traffic traces, to cluster them in groups sharing similar driving behaviour and then use the cluster features to configure the agents in the traffic simulator. To assess the feasibility of this methodology, we have started with a controlled experiment and generated synthetic traffic traces to understand which parameters are discriminant for the clustering algorithm and how to ex-

[1] https://ec.europa.eu/transport/themes/urban/urban_mobility_en

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

tract and manipulate them. We have compared traces that are ordered and not ordered, long and short and we have come out with some guidelines for conducting experiments on real traces.

The paper is organized as follows. In the next section, we first propose the concept underlying our method, then we introduce SUMO (Krajzewicz et al., 2012), the micro-simulator used, and we show how we applied the clustering algorithm. Finally, we present an experimental evaluation on a set of synthetic traces providing results and guidelines for future tests.

CONCEPT

We have designed a process aimed at configuring a micro-simulator to obtain realistic behaviour. The system pipeline is depicted in Figure 1. The starting point concerns the collection of real traffic traces from vehicles. The main parameters we have to collect are speed, position and acceleration at any point in time. These time series should be processed in real-time, stored in a data base and then manipulated to extract a meaningful training set for the clustering algorithm.

One important aspect of this processing phase is that, often, time series have different lengths, or contain many values that refer to stops (either at traffic lights or in congestions) that do not provide any meaningful insight on the driving styles. Therefore we have basically three possibilities: (1) either cut the time series to obtain feature data of the same length, (2) order the time series and cut them after the sorting or (3) use aggregate values (average speed/acceleration, standard deviation, maximum speed/acceleration). In the experimental result section, we will consider these aspects and test each alternative to understand its effectiveness in extracting driving styles.

The clustering algorithm then creates clusters on the feature space and obtains groups of traces that hopefully share some driving style characteristics. The cluster dimension (namely the number of vehicle traces in the cluster), and other aggregate parameters (max speed, max acceleration, standard deviations of speed and acceleration) are extracted from clusters and fed into the simulators to generate vehicles with the same characteristics. In the future, self-driving cars could also take advantage of these clustered driving styles, so as to learn typical human behaviors.

This paper is a first but significant step toward the process described above. Here we focus on a controlled experiment, where traffic traces are synthetic and are generated to have specific features. In this way we artificially create traffic traces exposing a controlled number of driving styles, to double check if the clustering algorithm finds homogeneous clusters with respect to the driving style.

Clearly, the experimental setup should cover scenarios where driving styles are very different one another and scenarios where driving styles are somehow more similar. We will show in the experimental result section the generated scenarios and the corresponding results.

THE SUMO SIMULATOR

SUMO (Simulation of Urban MObility) (Krajzewicz et al., 2012) is a microscopic, time-discrete traffic flow simulation platform used in this paper to collect the traffic information for the controlled experiments.

In SUMO each vehicle is uniquely simulated: it has an unique identifier, a departure time and a route (defined as a list of streets) that it will follow through the road network. Furthermore, each vehicle can be characterized by a set of features (called type) describing how it will behave in the simulation, including physical information (ie its maximum speed and acceleration) and other more specific parameters that regulate advanced aspects of the simulation, such as the driver's willingness to respect the speed limits.

Since our goal is to assess the feasibility of our concept with SUMO, to create the different driver behaviours we focused on values that can be easily collected from the observation of real traffic. In particular, to characterize the different types we used the maximum acceleration and deceleration and the willingness of the driver to follow the speed limit. This last parameter can be controlled in SUMO via a Gaussian distribution where $speedFactor$ is the mean and $speedDeviation$ is the standard deviation. When a vehicle enters the simulation SUMO computes its speed factor, using the Gaussian defined by its type. This means that the real maximum speed of a vehicle on a lane is $vehicleSpeedFactor \times laneSpeedLimit$.

SUMO allows many kinds of data outputs for each simulation, and also offers an API that can be leveraged to control the simulation online through another program. When used in this mode, SUMO allows the client to access many aspects of the simulation undergoing at every time step and also to change the values. This makes collecting custom information very efficient, without the need to parse large output files.

We used the software platform to generate custom traffic demands and interacted directly with the simulator online. We used these features to create the synthetic scenarios that were used to validate our approach, and to extract the traffic traces to perform the analysis of the simulations.

CLUSTERING

Given a collection of traffic traces, our goal is to look for drivers that share similar characteristics. From a machine learning point of view, this is an unsupervised learning task, since we do not know in advance the categories into which the examples should be partitioned. Differently from a supervised learning setting, when one is given a collection of object-target pairs with the aim of learning a function that associates the objects to their targets, in an unsupervised setting we are only given unlabeled observations, with the goal of automatically detecting relevant, common patterns among the examples (Hastie et al., 2001). This setting is particularly appropriate for our scenario, since it is unlikely to pretend to know in advance a precise set of driver categories, but it is much more reasonable to search for

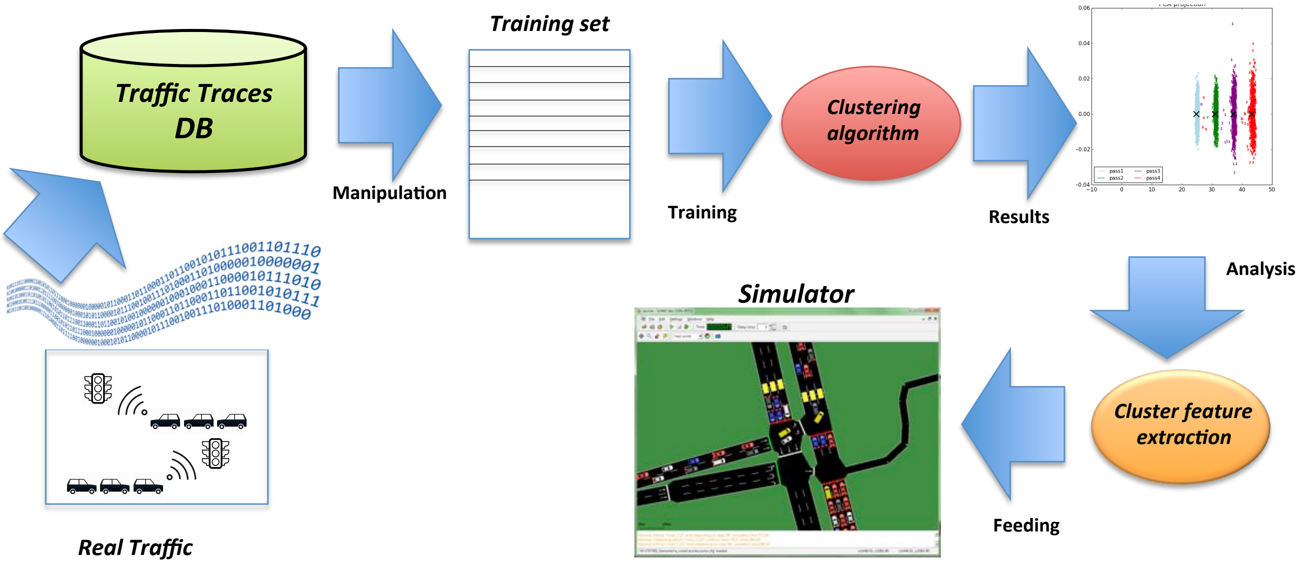


Fig. 1: Process flow for configuring realistic micro-simulators

emerging behaviors directly from data observations.

To this aim, we employ one of the most used and studied clustering algorithms, namely k-means (Hartigan and Wong, 1979). We hereby remark that other more sophisticated algorithms could indeed be applied to the same problem: yet, our goal in this paper is to provide a proof-of-concept of the whole system, thus we selected such a simple yet effective algorithm, leaving for future work the analysis of alternatives. Given n data points and an integer k, k-means partitions the data into a set of k clusters. This is done by finding the k cluster centers with an iterative procedure: starting with k initial centers (e.g., randomly chosen), the remaining points are associated each to the closest center. Then, the centers of mass of the so-obtained clusters are computed, which produces a new set of k cluster centers. Such steps are repeated until convergence. The algorithm is guaranteed to converge to a local optmimum, but it greatly depends on the initialization of the cluster centers. In this paper we employ two standard techniques that typically improve clustering results: (i) a smart initialization of the centers, named k-means++ (Arthur and Vassilvitskii, 2007), that tries to have an initial set of spatially distant points; (ii) multiple re-starts of the algorithm, finally choosing the best solution according to a certain criterion. More details on the metrics used to assess the goodness of clustering will be given in the experimental section. Here we just anticipate that, since we operate in a controlled (simulated) environment, clearly we also know the true labels of the vehicles (since we generate them). This allows us to use clustering evaluation metrics that exploit knowledge of the labels.

With respect to our specific problem, the main issue that has to be solved when feeding data to the clustering algorithm is how to represent each traffic trace. In fact, clustering algorithms (including k-means) typically assume all data samples to have the same dimensionality, that is they are represented by the same number of features. In the case of traffic traces, instead, this is clearly not true, since the length of the trace of each vehicle depends on how long the vehicle has been observed (in the controlled experiment, how much time it spends within the simulation). Moreover, the trace may possibly contain information about both the speed and the acceleration of the vehicle. Several solutions can be designed to address these issues. In this work we considered four different possibilities: (i) compute aggregated statistics (e.g., mean, variance, etc.) of each trace, to be used as feature vectors; (ii) choose the m largest values of the trace (in decreasing order); (iii) choose the m smallest values of the trace (in increasing order); (iv) choose the first m timestamps of the trace.

EXPERIMENTS

We now present the experimental evaluation that we conducted in a controlled setting implemented within SUMO, version 0.27.1. We considered three different scenarios of increasing difficulty for the clustering algorithm, to study the performance of our approach with respect to the heterogeneity in the generated traffic traces. Each simulation has run for a period of 4 hours, with about 14,000 vehicles. Then, we extracted the 25% of vehicles with the longest traces in the simulation, so that several lengths for the feature vectors could be tested in the experiments. Therefore, each scenario contains about 3,600 vehicles. As for the road network, we used a portion of the Bologna metropolitan area, depicted in Figure 2. In all the simulations, we kept SUMO parameter $speedDev = 0.1$ (representing the speed standard deviation).

For each scenario, we have applied the clustering algorithm on (i) aggregated data (ii) time series containing only speed values in time, (iii) time series containing acceleration values in time, and (iv) combined time

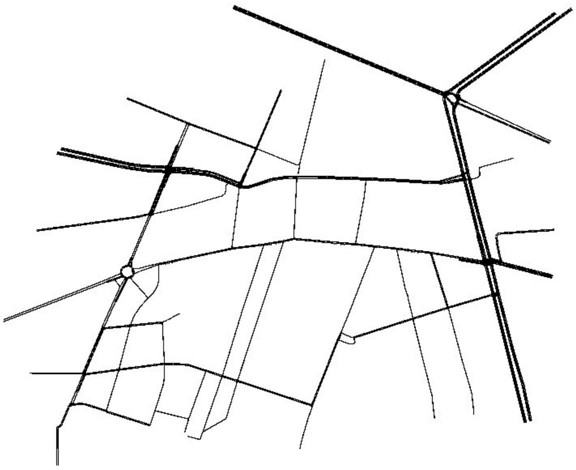


Fig. 2: Simulation scenario: a portion of the road network in the metropolitan area of Bologna.

series of speed and acceleration. The feature vector of aggregated data consists of three values: maximum speed, maximum acceleration and maximum deceleration. Since these will be exactly the parameters tuned in SUMO to generate different class vehicles in each scenario, we expect clustering with aggregated data to be very effective, and to act as a sort of upper bound for all the other approaches. Since perfectly knowing which are the discriminating parameters of the vehicle categories is clearly not realistic, it is interesting to compare the results of aggregated statistics with those obtained by directly employing the time-series (or a portion thereof).

When we did not compute aggregated statistics of the time series, we had to cut feature vectors to a fixed length of m timesteps. We considered values for $m = 5$ up to $m = 60$ with step 5 (from preliminary experiments, we observed no improvements with larger values of m). The trimmed traffic traces were considered both unsorted and in descending sorted order: experiments show that ascending ordering is useless, since low values of the time series are very similar across all the vehicle classes, typically corresponding to vehicle stops. For descending sorting, we ordered the traces for decreasing speed, and then considered the corresponding acceleration (thus, the acceleration values are not in descending order). In all the experiments, we set $k = 4$ for k-means. A validation of the clustering performance as a function of this parameter is left for future work. To evaluate clustering performance, we employ four metrics: Homogeneity (h), Completeness (c), v-measure (v) and Adjusted Rand Index (ARI). Homogeneity and completeness are strictly intertwined metrics: the first measures the degree of homogeneous labels within each cluster, whereas the second measures at what extent members of a certain class are assigned to the same cluster. Formally, they are defined as:

$$h = 1 - \frac{H(C|K)}{H(C)} \tag{1}$$

Class	a	d	sf
1	1.6	4.0	0.8
2	2.2	4.5	1.0
3	2.8	5.0	1.2
4	3.4	5.5	1.4

TABLE I: Parameters employed for the generation of vehicles in the easy scenario: a is acceleration, d is deceleration, sf is the speed factor.

$$c = 1 - \frac{H(K|C)}{H(K)} \tag{2}$$

where $H(C)$ is the entropy of the (true) classes, $H(K)$ is the entropy of the clusters, and $H(C|K)$, $H(K|C)$ are the two conditional entropies. The v-measure is the harmonic mean between h and c. The Adjusted Rand Index (ARI) is instead defined starting from Rand Index (RI):

$$RI = \frac{A + B}{Z} \tag{3}$$

where A is the number of pairs of examples that belong to the same class and to the same cluster, B is the number of pairs of examples that belong to different classes and also to different clusters, Z is a normalization factor (the sum of all possible pairs of examples). ARI is defined as an adjustment of RI taking into account chance normalization:

$$ARI = \frac{RI - E[RI]}{\max(RI) - E[RI]} \tag{4}$$

where $E[RI]$ is the expected RI of a random cluster assignment, and $\max(RI)$ is the maximum possible value of RI.

A. Easy scenario

In the first scenario, we generated four vehicle categories having different physical properties (acceleration and deceleration) and also different aggressiveness (willingness to ignore the speed limit). This is clearly the easiest scenario for the clustering algorithm, as vehicle categories should be well separated and distinguishable. The chosen parameters are shown in Table I. The simulation covers 3,595 vehicles, 36.2% of class 1, 25.1% of class 2, 21.2% of class 3, 17.5% of class 4. Table IIa reports the clustering metrics employed with the different settings. As expected, different driving styles are easily recognizable as they greatly vary for what concerns the speed and acceleration of the vehicles. Unsorted traffic traces lead to reasonable clustering results, with a v-measure in the best case equal to 0.754 (using both speed and acceleration). Descending, sorted time series produce instead very good results: the shortest time series (e.g., $m = 5$) work best, as large speed and acceleration values are the most informative data. Results with aggregated data confirm an almost perfect clustering, as expected. In Figure 3a we show a representation of the best clustering results obtained

Trace	Sorting	m	h	c	v	ARI
S	None	10	0.745	0.772	0.758	0.788
A	None	15	0.635	0.671	0.652	0.693
SA	None	20	0.741	0.768	0.754	0.786
S	Dec.	5	0.977	0.979	0.978	0.989
A	Dec.	5	0.979	0.980	0.979	0.989
SA	Dec.	5	0.872	0.882	0.877	0.915
Agg	–	–	0.985	0.985	0.985	0.993

(a)

Trace	Sorting	m	h	c	v	ARI
S	None	15	0.656	0.685	0.670	0.697
A	None	15	0.572	0.570	0.571	0.602
SA	None	25	0.716	0.739	0.727	0.755
S	Dec.	5	0.987	0.988	0.988	0.994
A	Dec.	50	0.318	0.326	0.322	0.233
SA	Dec.	5	0.884	0.889	0.887	0.914
Agg	–	–	0.993	0.993	0.993	0.997

(b)

Trace	Sorting	m	h	c	v	ARI
S	None	10	0.723	0.737	0.730	0.741
A	None	15	0.712	0.723	0.718	0.748
SA	None	20	0.727	0.740	0.733	0.747
S	Dec.	30	0.485	0.843	0.616	0.476
A	Dec.	25	0.590	0.661	0.623	0.509
SA	Dec.	15	0.474	0.740	0.578	0.459
Agg	–	–	0.989	0.988	0.989	0.994

(c)

TABLE II: Best clustering results obtained in each scenario: (a) easy, (b) intermediate, (c) hard. h is cluster homogeneity, c is completeness, v is the v-measure, and ARI is Adjusted Rand Index. For traces, S stays for speed, A for acceleration and SA for a combination of the two. Agg indicates aggregated features.

for this scenario ($m = 5$, descending ordering, negligible differences if using acceleration or speed). Examples have been projected to a two-dimensional feature space via Principal Component Analysis (PCA) (Jolliffe, 2002), where colors represent the cluster assignment by k-means, and digits indicate the true vehicle label.

B. Intermediate scenario

The second scenario we consider has an intermediate level of complexity, since only the aggressiveness of the vehicles changes. Again, we generated four traffic categories corresponding to four different values for SUMO parameter *speedFactor*: {0.8, 1.0, 1.2, 1.4}, which are the same employed in the easy scenario (but this time without differentiating also the physical properties). The simulation covers 3,643 vehicles, 32.6% of class 1, 26.1% of class 2, 21.5% of class 3, 19.8% of class 4. Table IIb reports the clustering metrics employed with the different settings. Similarly to the previous scenario, the clustering of unordered traffic still achieves metrics above 0.7, with slightly larger values of m with respect to the easy scenario. As in the easy scenario, switching to a descending ordering of the traces has lead to the best overall performance, and again $m = 5$ provided the best results. Differently from the previous setting, the acceleration in this case is ineffective in identifying the driving styles (since it is the speed factor that actually differentiates vehicles), whereas using speed or pairs speed/acceleration results in effective clustering. Interestingly, unsorted time series suffer somehow less of this problem, probably indicating that the *trend* in the variation of the acceleration contains information that can be exploited to characterize the driver's behaviour. Results with aggregate data again indicate an almost perfect clustering. The best results obtained for this scenario ($m = 5$, descending ordering, speed traces) are shown in Figure 3b.

k	m	h	c	v	ARI
4	20	0.727	0.740	0.733	0.747
6	25	0.793	0.683	0.734	0.742
8	25	0.806	0.585	0.678	0.592
10	25	0.821	0.521	0.638	0.489

TABLE III: Performance measurements with different numbers of clusters k in the hard scenario. In this setting we employ unsorted traces with both speed and acceleration. We report h as cluster homogeneity, c as completeness, v as the v-measure, and ARI as Adjusted Rand Index.

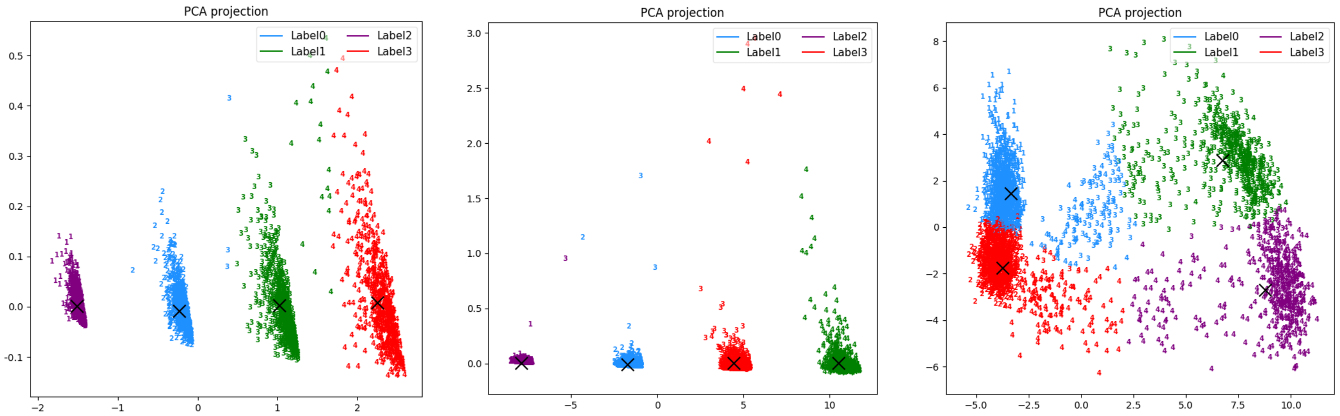
Class	a	d	sf
1	2.2	4.5	0.8
2	2.8	4.5	0.8
3	2.2	4.5	1.2
4	2.8	4.5	1.2

TABLE IV: Parameters employed for the generation of vehicles in the hard scenario: a is acceleration, d is deceleration, sf is the speed factor.

C. Hard scenario

The third scenario is the hardest for the clustering algorithm. Here, some vehicles have the same physical properties but different aggressiveness, whilst other have the same aggressiveness but different physical properties. Therefore, driving styles will be similar even among different categories, which makes the clustering more challenging. The chosen parameters for SUMO are shown in Table IV. The simulation covers 3,604 vehicles, 29.4% of class 1, 31.0% of class 2, 19.8% of class 3, 19.8% of class 4.

As shown in Table IV, this scenario is much more challenging than the previous ones, and results are very different. In this case, in fact, the best overall results were obtained when the data is left unsorted. The best v-measure achieved is 0.733 for $m = 20$ and time-series that combine speed and acceleration. These results can be explained with the observation that, in a complex



(a) (b) (c)

Fig. 3: Results on the three scenarios: easy (a), intermediate (b) and hard (c). Colors indicate the cluster assignment by k-means, whereas digits represent the true vehicle label. Black crosses indicate cluster centroids.

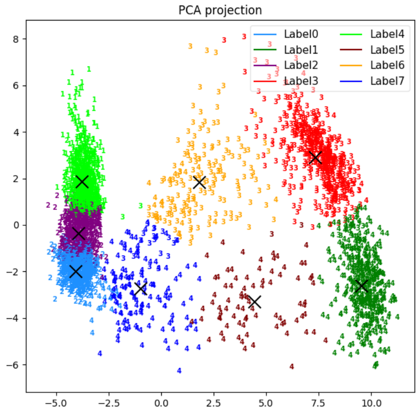


Fig. 4: Results on the hard scenarios obtained with 8 clusters. Colors indicate the cluster assignment by k-means, whereas digits represent the true vehicle label. Black crosses indicate cluster centroids.

scenario, the largest speed and the corresponding acceleration values are not sufficient to identify the driving style. It is instead much more informative to observe the *trend* in the time-series, which allows to consider, for example, how long it takes, for a driver, to accelerate and decelerate up to a certain speed. This information is clearly encoded in the unsorted time-series, but not in the ordered case. Table IIc reports the clustering metrics employed with the different settings, and Figure 3c shows the clustering after PCA projection.

For this scenario, in Table III we also report results for different numbers of clusters, obtained with unsorted traces with both speed and acceleration information. As the number of clusters grows, homogeneity increases, which means that the algorithms finds more clusters, but with stronger intra-cluster similarities: as a trade-off, not surprisingly, completeness decreases, as vehicles of the same class are sometimes split across different clusters. For example, Figure 4 shows that the bottom blue, brown and green clusters cover almost completely the fourth class of vehicles.

D. Discussion

The controlled experiments conducted with the SUMO simulator confirm that the proposed clustering approach could be effectively used to extract common driving behaviors from traffic traces. Clearly, aggregated statistics work extremely well, but this can happen only when there is a strong correlation between maximum values of speed and acceleration and driving style. Although such features are certainly crucial to detect common relevant patterns, they are not necessarily the only significant information in all real scenarios. Using decreasing ordering for time-series, and exploiting short feature vectors, typically leads to very good performance when the driving styles to be recognized are well distinguishable. On the other hand, it is very interesting to notice that unsorted traffic traces lead to very similar performance, in terms of clustering metrics, in all the three scenarios, thus showing to be robust across different, heterogeneous settings.

RELATED WORK

The problem of identifying common driving styles has been the subject of several studies, although without considering its integration within the context of micro-simulation tuning and optimization. Data mining techniques, including clustering, are employed in (Constantinescu et al., 2010) to identify common driving behaviors, by constructing aggregated statistics from a collection of GPS data. Supervised learning techniques for the classification of data coming from inertial sensors are presented in (Van Ly et al., 2013). In (Wang and Lukic, 2011) a review of the most widely employed features for driver characterization is presented, with the aim to build support systems for hybrid electric vehicle control strategy. Typical and aggressive driving style behaviors are classified in (Johnson and Trivedi, 2011), by exploiting smartphones as sensor platforms, and by employing a simple nearest neighbor classifier.

CONCLUSIONS

In this paper we presented a methodology for driving style characterization, that could be exploited for a more realistic design and calibration of micro-traffic simulators. The experimental results that we conducted in a controlled environment suggest that driving styles are best identified from the extreme behaviours of a driver, namely from the top-m largest values in the speed and acceleration time-series, but only when the vehicle categories are characterized by marked differences. Within this setting, we observed that speed values are typically more informative for the characterization of a driver. On the other hand, in more complex scenarios, it results to be more useful to consider a portion of the unsorted traffic trace, which allows to observe also the trend in the speed and acceleration time-series. Among future research directions, we are currently exploring the use of unsupervised deep networks, namely Stacked Denoising Autoencoders (Vincent et al., 2010), to perform automatic feature extraction from traffic time-series. More clustering algorithms will also be tested, and their performance will be evaluated as a function of the parameter representing the expected number of clusters, as we already started investigating in the reported experiments. Dynamic time warping (Berndt and Clifford, 1994) could also be an alternative for the comparison of time series with different lengths. It would be interesting also to assess the performance of the approach when vehicle categories are extremely imbalanced (i.e., few vehicles for some of the categories). Finally, our ultimate goal would be to use our system with real traffic traces.

REFERENCES

Arthur, D. and S. Vassilvitskii (2007). “k-means++: The advantages of careful seeding”. In: *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, pp. 1027–1035.

Berndt, D. J. and J. Clifford (1994). “Using dynamic time warping to find patterns in time series.” In: *KDD workshop*. Vol. 10. 16. Seattle, WA, pp. 359–370.

Constantinescu, Z., C. Marinoiu, and M. Vladoiu (2010). “Driving style analysis using data mining techniques”. In: *International Journal of Computers Communications & Control* 5.5, pp. 654–663.

Hartigan, J. A. and M. A. Wong (1979). “Algorithm AS 136: A k-means clustering algorithm”. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28.1, pp. 100–108.

Hastie, T., R. Tibshirani, and J. Friedman (2001). *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc.

Helbing, D. et al. (2002). “Micro-and macro-simulation of freeway traffic”. In: *Mathematical and computer modelling* 35.5-6, pp. 517–547.

Johnson, D. A. and M. M. Trivedi (2011). “Driving style recognition using a smartphone as a sensor platform”. In: *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on*. IEEE, pp. 1609–1615.

Jolliffe, I. (2002). *Principal component analysis*. Wiley Online Library.

Krajzewicz, D. et al. (2012). “Recent Development and Applications of SUMO - Simulation of Urban MObility”. In: *International Journal On Advances in Systems and Measurements* 5.3&4, pp. 128–138.

Van Ly, M., S. Martin, and M. M. Trivedi (2013). “Driver classification and driving style recognition using inertial sensors”. In: *Intelligent Vehicles Symposium (IV), 2013 IEEE*. IEEE, pp. 1040–1045.

Vincent, P. et al. (2010). “Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion”. In: *Journal of Machine Learning Research* 11.Dec, pp. 3371–3408.

Wang, R. and S. M. Lukic (2011). “Review of driving conditions prediction and driving style recognition based control algorithms for hybrid electric vehicles”. In: *Vehicle Power and Propulsion Conference (VPPC), 2011 IEEE*. IEEE, pp. 1–7.

ALESSANDRO PETRARO is a software engineer at Cubbit, a digital startup accelerated by the investment fund Barcamper Ventures (PrimoMiglio SGR). His main interests are in the fields of artificial intelligence, machine learning and distributed networks. His e-mail address is: alessandro.petraro2@studio.unibo.it and his complete profile can be found at https://www.linkedin.com/in/alessandro- petraro-b9aa308b

FEDERICO CASELLI is research associate at Department of Computer Science and Engineering at the University of Bologna. His main research interest are in the fields of transportation systems, simulations, distributed networks and decision support systems.

MARCO LIPPI is assistant professor in Computer Engineering at the University of Modena and Reggio Emilia. His main research interests are in the fields of artificial intelligence and machine learning, with applications to bioinformatics, argumentation mining, transportation systems and computer vision. His e-mail address is: marco.lippi@unimore.it and his webpage can be found at http://www.agentgroup.unimore.it/Lippi.

MICHELA MILANO is full professor of Artificial Intelligence at the Department of Computer Science and Engineering at the University of Bologna. Her main research interest cover decision support systems and their integration with machine learning algorithms. She is Editor in Chief of the International Journal Constraints, member of the EurAI board, Executive Counsilor of AAAI, author of more than 130 papers on peer reviewed journals and conferences and coordinator of EU projects and industrial collaborations.

Finance and Economics and Social Science

SIMULATION MODELS OF TWO DUOPOLY GAMES

Ingolf Ståhl
Department of Economic Statistics
Stockholm School of Economics
Box 6153
SE-11383, Stockholm, Sweden
Email: ingolf.stahl@hhs.se

KEYWORDS
Duopoly theory, Cournot, Bertrand, Discrete Event Simulation, GPSS.

ABSTRACT

The paper presents simulation of two classic duopoly games, those of Cournot and Bertrand. The simulation is done with a very simple Discrete Event Simulation system, aGPSS, developed for economics students. With the Cournot model one can study under which conditions the Cournot solution will be obtained even if the parties lack knowledge about demand and costs. The Bertrand model, with stochastic demand, allows the study of cases when a firm can charge a higher price than the other firm and yet sell because the firm charging a lower price sells more than it has expected.

INTRODUCTION

In recent years there has been a strong trend towards modifying some original game theory ideas, completely based on mathematical deduction, by introducing behavioral aspects, often learnt from experiments with people playing games, modeled in accordance with the original game theory models. In contrast to usual game theory models, where one by deduction only obtains one single solution, the outcome in these experiments often vary greatly between each other, often depending on the types of participant, and one hence needs many experiments when studying a particular game.

Since these experiments are time consuming and costly, there has arisen an interest in running computer simulations, where one in the simulation model can introduce various behavioral assumptions. The behavioral assumptions are for both models of this paper based on evidence from experimental games with Swedish students. One can also allow for various types of uncertainty, e.g. about demand and costs.

We shall in this paper present two very simple simulation models, based on fundamental game theory. They both deal with duopoly, i.e. a market situation with two sellers, and homogeneous products, i.e. both firms produce identical products, like wheat or oil. The game theory behind these two models are taught in most micro-economic courses, often in the same class session, as the starting point of oligopoly theory. We shall after having presented our simulation models, discuss their relationship to earlier literature and how they can be expanded.

The first model deals with the oldest of all game theory models, the duopoly model of Cournot from 1838 (Cournot 1838). The distinguishing assumption is here that the two firms just decide on the quantity that they supply to the market and that price is dependent on the total supply. The second game example is the second oldest duopoly model, namely that of Bertrand from 1883 (Bertrand 1883). The main distinction from the Cournot game is that in the Bertrand game both firms set a price, and the quantity sold by each firm is dependent on the price it sets and the price set by the competitor. Furthermore, while there is no uncertainty in our Cournot model, the demand in our Bertrand model is characterized by stochastic variations. For both simulation models we assume that both firms have constant unit costs.

We shall simulate both models in a simple Discrete Event Simulation system, aGPSS. aGPSS is a streamlined and very simplified version of GPSS (the General Purpose Simulation System), originally an IBM product. aGPSS is mainly used in business schools in shorter courses in Management Science (Born and Ståhl 2013, Ståhl 2007). In contrast to earlier GPSS systems, it has a GUI for building the model and built- in graphics. Both models, as well as the aGPSS system, can be downloaded from www.aGPSS.com. The two models have been used at the Norwegian School of Economics in Bergen in a simulation course focused on applications in the energy sector.

THE COURNOT GAME

Our first example deals, as mentioned, with a game of duopolies producing and selling homogenous goods, like wheat, oil or coal. Although the game is very small, our model contains the main characteristics of a game based on Discrete Event Simulation. To make the model very simple we assume that demand can be described by a linear demand function $p = a\text{-}bQ$, where $Q = \Sigma q$. Both

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

producers have the same constant unit cost *c*. We shall first study the analytical solution. For this purpose, we can write firm *i*'s profit as

$$v_i = q_i(p\text{-}c) = q_i(a\text{-}bQ\text{-}c) = q_ib((a\text{-}c)/b\text{-}Q) = q_ib(A\text{-}Q), \text{ where } A=(a\text{-}c)/b$$

To simplify further we write this profit on a new scale as $V_i=v_i/b=q_i(A\text{-}Q)$

For this duopoly case, with $Q = q_1+q_2$ we hence have

$$V_1 = q_1(A - q_1 - q_2) \qquad (1)$$

$$= A\, q_1 - q_1^2 - q_1\, q_2 \text{ and } V_2 = Aq_2 - q_2q_1 - q_2^2$$

The standard analytical solution of the game, the non-cooperative solution, found already by Cournot, is obtained when each firm regards the competitor's quantity as given. The optimal quantities are then determined by setting

$$V_1'(q_1) = A - 2\, q_1 - q_2 = 0 \quad \text{and}$$

$$V_2'(q_2) = A - q_1 - 2q_2 = 0$$

This implies in turn that $2\, q_1 + q_2 = A = q_1 + 2q_2$, i.e.

$$q_1 = q_2, \text{ i.e. } A\text{-}3\, q_1 = 0, \text{ i.e.}$$

$$q_1 = q_2 = A/3 \qquad (2)$$

The optimization decisions behind this equilibrium presupposes that both parties know the parameters *a* and *b* of the demand function and also the unit cost *c*.

In line with behavioral game theory (Camerer 2003), we now want to simulate a market, where the two firms know neither demand nor costs, but only their own profits and own quantities offered in each period. In experiments, done e.g. by Grubbström (1972) and Edman (2005), one has found that, under certain conditions, the Cournot solution will be obtained after some time also under these assumptions of no knowledge about demand and costs. The question is under what general conditions this can occur. In order to investigate this, a simulation model is necessary, since one in contrast to costly experiments can afford to test a great number of conditions by inexpensive simulations.

The aGPSS model used here for these simulations is quite small and we shall present it in four parts.

Table 1: Part 1 of Cournot model

```
      INPUT   x$oneOld
      INPUT   x$oneNew
      INPUT   x$twoOld
      INPUT   x$twoNew
      INPUT   x$aval
      INPUT   x$chg
      LET     x$size=1
prof1 VALUEOF x$one*(x$aval-x$one-x$two)
prof2 VALUEOF x$two*(x$aval-x$one-x$two)
```

The first part deals with the starting values, before the game begins. For each of the two firms, we first input two starting quantities, *oneOld* and *oneNew* for firm 1, and *twoOld* and *twoNew* for firm 2.

We also input the intercept of the linear demand function *aval* (*A = (a-c)/b*) and a change factor *chg*. We set the initial value of *size* to 1. *chg* and *size* are both, as discussed below, used to try to ensure that we get some convergence of the quantities. We note that variables are preceded by x$ in aGPSS.

At the bottom of table 1 we define two expressions. First, *prof1=one*(aval-one-two),* which corresponds to equation (1) $V_1 = q_1(A - q_1 - q_2)$. Likewise we define the profit function *prof2* for firm 2.

We next proceed to part 2, where the simulation starts.

Table 2: Part 2 of Cournot model

```
      GENERATE  ,,0,1 !  Start conditions
      LET       x$one=x$oneOld! 1’s start q
      LET       x$two=x$twoOld! 2’s start q
      GRAPH     cl,x$one,x$two
      LET       x$qch1=x$oneNew-x$oneOld
      LET       x$qch2=x$twoNew-x$twoOld
       TERMINATE
*
       GENERATE  ,,1,1      !  Time 1
       LET       x$one=x$oneNew ! 1’s first move
       GRAPH     cl,x$one,x$two
       TERMINATE
```

We here have two segments. A first single event is generated at time 0. This puts the starting quantities of the two firms into a graph, after first having given them to the variables *one* and *two*. Next we calculate the initial changes in quantity, *qch1* and *qch2.* Another single event is then generated, at time 1, which puts *oneNew,* by the way of *one*, into the graph at time 1.

Part 3 contains the most important actions of firm 1.

Table 3: Part 3 of Cournot model

```
     GENERATE  2,,2 !  1 moves in 2,4,6...
     LET       x$one=x$oneOld! 1's earlier q
     LET       x$two=x$twoOld! 2's earlier q
     LET       x$oldP1=v$prof1 ! 1’s old profit
     LET       x$oneNew=x$oneOld+x$qch1
     LET       x$one=x$oneNew ! 1’s new q
     LET       x$newP1=v$prof1!  1's new profit
     IF        x$newp1>x$oldp1,inc
     IF        x$newp1<x$oldp1,dec!
     GOTO      dec,0.5 ! If equal,50 % to DEC
inc  LET       x$qch1=(x$oneNew-x$oneOld)*x$size
     GOTO      join
dec  LET       x$qch1=(x$oneOld-x$oneNew)*x$size
join LET       x$size=x$size*x$chg! Change size
     GRAPH     cl,x$one,x$two ! q1+q2 to graph
     LET       x$oneOld=x$oneNew! Oldq now newq
     TERMINATE 1
```

Firm 1 makes its moves in periods 2, 4, 6, etc., while firm 2, as seen in Table 4 below, makes its moves in periods 3, 5, 7, etc. aGPSS will with the block GENERATE 2,,2 in Part 3 produce a loop where the segment of firm 1's moves will be gone through every other period. The parties will hence take turns making their moves. This sequential quantity setting is fundamental for the model, since it ensures that each party can see the effect of one's own latest move.

We start by calculating firm 1's old profit *oldP1*. For this, we call on the *prof1* value expression by *V$prof1* at the bottom of table 1. To use this expression we first have to set `one=oneOld` and `two=twoOld`. We next calculate the new q_1 for firm 1, *oneNew*, by adding *qch1*, determined in a preceding period, to the old q_1. On the basis of this new q_1 we calculate the profit *newP1* of firm 1 to be obtained in this period. If this new profit is higher than the previous profit, we calculate the next change of q_1, *qch1* (at the line called *inc*), as an increase, i.e. as *(oneNew-oneOld)*size*, where *size* is a factor going over time from 1 to eventually be close to 0. If the new profit is lower than the previous profit, we calculate instead the next change of q_1 (at the line *dec*) as a decrease *(oneOld-oneNew)*size*. In case of equal, i.e. unchanged profits, we go with 50 percent probability to *inc* and with 50 percent probability to *dec*.

Finally in each period 2, 4, 6, etc., we first update the value of *size*. The variable *size* is used when calculating the change at *inc* and *dec*, as seen in the previous paragraph. *size*, which is initially set at 1 (see Table 1), is here multiplied by *chg*. If *chg* is input e.g. as 0.9, then *size* will take the values 0.81, 0.729, etc. This is meant to make convergence possible. We also put the new value of q_1 into a graph and give the value of *oneNew* to the variable *oneOld* to update this for the next period.

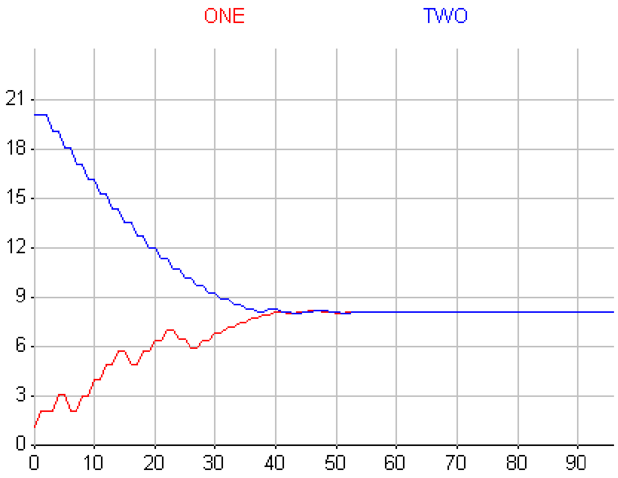
There is a similar segment for the actions of firm 2 in periods 3, 5, 7, etc. This is mainly obtained by letting the variables of firms 1 and 2 in Table 3 switch names. This segment is presented in table 4 below.

Table 4: Part 4 of Cournot model

```
        GENERATE 2,,3! 2 moves in 3, 5, 7..
        LET   x$one=x$oneOld
        LET   x$two=x$twoOld
        LET   x$oldP2=v$prof2
        LET   x$twoNew=x$twoOld+x$qch2
        LET   x$two=x$twoNew
        LET   x$newP2=v$prof2
        IF    x$newp2>x$oldp2,inc2
        IF    x$newp2<x$oldp2,dec2
        GOTO  dec2,0.5
inc2    LET   x$qch2=(x$twoNew-x$twoOld)*x$size
        GOTO  join2
dec2    LET   x$qch2=(x$twoOld-x$twoNew)*x$size
join2   LET   x$size=x$size*x$chg
        GRAPH c1,x$one,x$two
        LET   x$twoOld=x$twoNew
        TERMINATE 1
        START 100
        END
```

Besides the blocks similar to those of Part 3, Part 4 also contains a statement START 100, allowing for a total of 100 periods, and a finishing END statement.

When a user runs this model, aGPSS will first ask for the values of *oneOld, oneNew, twoOld, twoNew, aval* and *chg*. Let us exemplify with the values 1, 2, 20, 19, 24 and 0.995. In this case, with the initial values of q_1 and q_2 fairly far apart, we get the following graph.



Figures 1: Example of Cournot Game Graph

We see that we after around 40 periods get a convergence on the quantity 8. In this case with *aval* = *A*=24, this is equivalent to the theoretical Cournot equilibrium of 24/3 = 8 of equation (2).

If we input other values, we can see that convergence might not be so fast and in several cases there will not be any convergence. We first keep the value of *aval* as 24 and of *chg* as 0.995, but change the starting quantities. With e.g. 5, 6, 15, 14 or 6, 7, 10, 9, we get convergence a little later and somewhat more oscillations than in Figure 1.

If we, however, next keep the originally input values of 1, 2, 20, 19 and of *chg*=0.995, but set *aval* to 30 or 42, we still get convergence, but with *aval* = 45 or 48 there is no convergence. We next keep all the originally input values except for *chg,* which we set to 0.9 or 0.99. In both cases we do not get convergence. If we set *chg* to 1, we get repeated oscillations between 7 and 9, even if we run for e.g. 300 periods.

By downloading this Cournot model from aGPSS.com one can test out any combination of values. One can then possibly get ideas for how to construct experiments, e.g. as regards the number of periods to be played, but one can also change the parameter *aval* of the demand function. It would also be very simple to change the demand function, e.g. to the constant elasticity function of the Bertrand model below.

THE BERTRAND GAME

Also in the Bertrand game we have two firms that sell identical products. If, as in original model from 1883, demand is deterministic, marginal costs are constant, production is made to order, there are no limits to production capacities and no inventories, then the firm with the lower unit cost will according to theory drive the price down to just below the unit cost of the competitor, who will then not be able to sell anything.

If one, however, assumes that demand is stochastic and that the firms produce to inventory, this conclusion does not hold. This was seen when playing a small DES game, partly similar to the one presented below, in a set of experiments, run with Swedish business students (Ståhl 1993, Ståhl 2010).

The firm with the higher unit cost could survive by being able to sell at a higher price, since the other producer would run out of inventories from time to time, due to the stochastic demand variations, and the buyers would then purchase from the firm with inventories on hand, even if it charged a higher price.

The model presented here is different to the model used for these experiments, by the fact that values are only input once at the beginning and the prices are changed by a simple "robot" algorithm in the program, where prices are changed automatically depending on the size of the inventory. Another assumption is that prices are never set below cost. These behavioral assumptions were influenced by experience from the experimental game runs with the Swedish business students.

The aGPSS Bertrand model is also very simple. We shall present it in four parts.

The first part of the Bertrand program is shown in Table 5 below. We here first input the initial prices and the unit costs of the two firms. We next input the two constants of the demand function, the size factor *a*, here called *aSize*, and the absolute value of the price elasticity *b*, here called *elast*.

Table 5: Part 1 of Bertrand model

```
        SIMULATE 1
        INPUT x$price1
        INPUT x$price2
        INPUT x$cost1
        INPUT x$cost2
        INPUT x$asize
        INPUT x$elast
        INPUT x$priLow
        INPUT x$priAdd
        INPUT x$protim
        LET   x$lowpr=x$price1
stock1  CAPACITY ! 2,000,000,000
stock2  CAPACITY ! 2,000,000,000
demand  VALUEOF asize*x$lowPr^(-x$elast)
```

We also input the two constants *priLow* and *priAdd*, used for changing the prices based on stocks. We finally input *proTim*, the production time.

Since we name the firms 1 and 2 so that firm 1 is the firm with the initially lower price, we also set the temporary lower price, *lowrPr*, equal to *price1.*

Next the program contains the definition of the capacities of the inventories of the firms. Since we do not want to limit these, we give them a very high capacity.

Finally in part 1, we define the demand function, *demand*, that a firm faces, if it charges a lower price, *lowrPr,* than the competitor. This is equivalent to the formula $q = ap^{-b}$. We believe that this demand function in general is more realistic than the linear demand function of the Cournot model above.

The Bertrand model next contains a small segment on the initial values of sales and production quantities shown in Table 2 below.

Table 6: Part 2 of Bertrand model

```
GENERATE  ,,0,1,1  !  Initial  decisions
LET x$salLa1=v$demand/52!Sold last week
LET x$salLa2=v$demand/52
GRAPH cl,x$price1
GRAPH cl,x$price2
TERMINATE
```

The events are generated once, at time 0, with priority to assure that they precede those of the first week in part 3. In order to determine production in period 1, done in part 3, we set for each firm the sales of the preceding week (before the start), *salLa1* and *salLa2*, as possible annual sales divided by 52. We then put the initially read-in prices into graphs.

The next part of the Bertrand program is shown in Table 7 on the next page. We here generate a report and decision event each week of 7 days, starting at time 0. We here first print the number of the week, CL/7, where the clock time CL is the number of days since simulation start. Then the profits as well as then number of units in stock of each firm are printed. All printing is done with 0 decimals, set at the first PRINT block.

Next we check if firm 1 has no stocks. If so, the price is too low and *price1* is increased by *priAdd.* If not, i.e. stocks are not empty (NE), price has been too high and we decrease *price1,* at *lower1,* by *priLow* times the number of units in stock, so that price is lowered more if the stocks are large. The price may, however, not be lower than unit cost. In that case *price1* is set to *cost1.*

In the blocks started with the address *price2* we execute the changes in price done by firm 2, which are similar.

Table 7: Part 3 of Bertrand model

```
month   GENERATE 7,,0 ! At start of every week
        PRINT 'Week',cl/7,0
        PRINT 'Profits firm 1',x$sales1-x$tcos1
        PRINT 'Profits firm 2',x$sales2-x$tcos2
        PRINT 'Stocks firm 1',s$stock1
        PRINT 'Stocks firm 2',s$stock2
        IF stock1=NE,lower1 ! Stocks empty?
        LET+ x$price1,priAdd
! Price 1 up if everything sold
        GOTO price2
lower1  LET- x$price1,s$stock1*x$priLow!
! Price 1 reduced due to remaing stocks
        LET x$price1=fn$max(x$price1,x$cost1)
! Set price to cost if reduced price < cost
price2  IF stock2=NE,lower2 !Same as for firm 1
        LET+ x$price2,priAdd
        GOTO prod1
Lower2  LET- x$price2,s$stock2*x$priLow
        LET x$price2=fn$max(x$price2,x$cost2)
Prod1   ADVANCE x$proTim      ! Production time
        LET x$ordq1=x$salLa1   ! q =last sales
        IF s$stock1>=x$ordq1,prod2! Stocks large?
        ENTER stock1,x$ordq1   ! Into stocks
prod2   LET x$ordq2=x$salLa2   ! q =last sales
        IF s$stock2>=x$ordq2,finish
        ENTER stock2,x$ordq2
finish  GRAPH cl,x$price1
        GRAPH cl,x$price2
        LET x$salLa1=0 ! Set last sales = 0
        LET x$salLa2=0
        x$lowrPr=fn$min(x$price1,x$price2)
! New lowest price set
next    TERMINATE 1
```

Then at the address *prod1* we deal with production. Firm 1 here decides on a production quantity. We assume that it is ignorant about the price set by firm 2, so firm 1 just sets its production *ordq1* equal to the amount it sold in the preceding period, i.e. *salLa1*, determined in part 4 (see Table 8). Firm 2 production at *prod2* is similar.

After a potential production period of *proTim* days, the products are ready to be put into inventory to be available for sales. However, if firm 1's stock, *stock1*, is already larger than its planned production, then the blocks dealing with putting production into stocks are skipped, implying that no production takes place this month. Similar conditions apply to firm 2.

We next put the possibly new prices into graphs and next we set the accumulated sales of the week, *saLa1* and *salLa2*, to 0, so that they can be updated correctly in part 4 (see Table 8).

We also determine what is now the lower price of *price1* and *price2*, and set this as *lowrP*r. Finally, at the end of each week we will with TERMINATE 1 decrease the termination counter, initially set to 52 in Part 4, by 1. In week 52 the counter becomes 0, which stops the simulation.

The rest of the Bertrand model is shown in Table 8.

Table 8: Part 4 of Bertrand model

```
        GENERATE fn$xpdis*52*7/v$demand
* Mean IAT = 364/annual sales
        IF x$price1<x$price2,sltes1 ! p1<p2
        IF x$price1>x$price2,sltes2 ! p2<p1
        GOTO sltes2,0.5! If p1=p2 50/50
sltes1  IF stock1=NE,sell1 ! 1 sells if stocks
        IF stock2=NE,sell2 ! Else 2 if stocks
        GOTO sell1
* If 2 also 0 stocks, go to firm 1 to wait
sltes2  IF stock2=NE,sell2 ! 2 sells if stocks
        IF stock1=NE,sell1 ! Else 1 if stocks
        GOTO sell2
* If 1 also 0 stocks, go to firm 2 to wait
sell1   WAITIF stock1=E! Wait if 1 has 0 stocks
        LEAVE stock1 ! Take 1 unit from stocks
        LET+ x$salLa1,1 ! 1 more sold for 1
        LET+ x$sal1,x$price1 ! Increase revenue
        LET+ x$tcos1,x$cost1 !Incr. total costs
        TERMINATE
sell2   WAITIF stock2=E
        LEAVE stock2
        LET+ x$salLa2,1
        LET+ x$sal2,x$price2
        LET+ x$tcos2,x$cost2
        TERMINATE
        START 52
        END
```

In the GENERATE block we generate every single order from customers wanting to buy at the lower price. The average time between two such orders, measured in days (with 52*7 days a year), is 364 divided by the number units demanded annually, *V$demand*, obtained from the *VALUEOF* in Table 1. The actual time between two orders varies, however, stochastically, since we multiply this average time by *fn$xpdis*, representing the negative exponential distribution. This usually gives a value between 0 and 8, with values below 1 in 67 percent of the times.

If *price1<price2*, i.e. firm1 has the lower price, we go to the address *sltes1*, where we test if firm 1 has any stocks. If so, firm 1 can sell and we go to *sell1*. If *price1>price2*, i.e. firm 2 has the lower price, we go to the address *sltes2*, where we test if firm 2 has any stocks. If so, firm 2 can sell and we go to *sell2*. If the firms have equal prices, we proceed by random with 50 percent chance to *sltes2* and with 50 percent chance to the next block *sltes1*

At the address *sell1* we might first have to wait until there is some unit in firm 1's stocks. We then take one unit out of firm 1's stocks and increase the number of sold units this week by 1 as well as firm 1's revenues by the price of the product. We also add up firm 1's costs of goods sold by the unit cost of the product and then terminate this sales event. At the address *sell2* the corresponding happens to firm 2.

This completes the blocks of the Bertrand model. The model is then, as seen at the bottom of Table 8, finished by START 52 and END.

When this model is run, it produces a table and two graphs. We exemplify with the following input values: *price1*=24, *price2*=26, *cost1*=12, *cost2*=13, *asize*=30000, *elast*=1.5, *priLow*=0.1, *priadd*=0.5, and *proTim*=5.

In table 9 we see an excerpt of the table on weekly data on profits and stocks that is then obtained. .

Table 9: Example of Profits and Stocks

```
Week          1
Profits firm 1     50
Profits firm 2      0
Stocks firm 1       0
Stocks firm 2       4
Week          2
Profits firm 1    102
Profits firm 2     13
Stocks firm 1       0
Stocks firm 2       3
Week          3
Profits firm 1    115
Profits firm 2     51
Stocks firm 1       3
Stocks firm 2       3
```

We see that also firm 2 with higher unit costs can make profits, which is in contrast to the original Bertrand solution, according to which firm 2 would get 0 profits

In Figure 2 we present one of the graphs, namely that of the prices of firm 2, i.e. the firm with the higher costs. The graph of the prices of firm 1 is not very different.

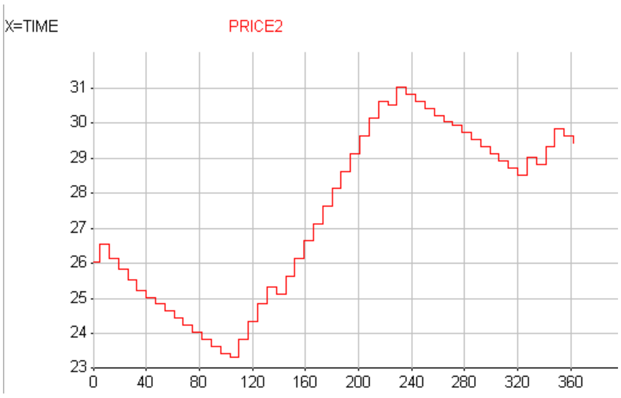


Figure 2: Example of Bertrand Price 2 Graph

We here see that development is quite different from that of the original theory, where the price of the firm with lower cost would go down to be just below the unit cost of the competitor, i.e. in this case 13. Here the price also goes up and, when going down, never to 13.

SIMULATION, EXPERIMENTS AND THEORY

We shall finally comment briefly on the relationship between the two simulation models presented above and the experiments and original theories inspiring the models to try to bring out what is special with the two models above.

We have as a background for our thoughts in this area glanced at some 60 entries on Google, under the headings Cournot and Bertrand, with the subheadings theory, experiments and simulations.

It should first be noted that both regarding the Cournot and the Bertrand game there are lot of issues and complications, like more than two players, that have not at all been touched upon in this paper.

As regards the literature items on Cournot, it should first be noted that many of them deal with the issue of whether there will a cooperative solution, i.e. collusion, or a non-cooperative solution, i.e. a Nash equilibrium, which in this case is the original Cournot solution. In particular many of the experiments deal with this issue (e.g. Thorlund-Petersen 1990). It seems that in many experiments with two and also three players the results are closer to the cooperative solution, but with four and more firms the solution is closer to the non-cooperative one (Huck *et al.*, 2004).

As regards the Bertrand game, there is both much theory and also some experiments dealing outright with capacity restraints, which can lead to the result that also the firm with higher costs an sell (e.g. Brown Kruse *et al.* 1994). It should be stressed that in our Bertrand model there are no capacity restraints. Our result with also the higher cost firm selling is due to our assumption of stochastic demand, an assumption that is rare in the literature.

As regards duopoly simulations in the literature, they mainly refer to the Cournot game and most of them are done in Excel, sometimes with the spreadsheet made available on the web. A few other simulations are stated to have been made in Java and JavaScript, but we have not been able to find the programs. We have not either been able find any simulation of duopoly games done in a discrete event simulation package.

By providing both the simulation system, aGPSS, the two models as well as introductory lessons, free of charge, on the web at www.aGPSS.com, we hope that it will be possible for other oligopoly researchers to extend the two models to cover important aspects left out in our simple models. Extensions to three firms and to different unit costs for the firms in the Cournot model seem like suitable first steps.

REFERENCES

Bertrand, J. 1883. “Théorie Mathematique de la Richesse Sociale.” *Journal de Savants*, 53, 499-508.

Born, R. and I. Ståhl. 2013. *Modeling Business Processes with a General Purpose Simulation System – Part 1 Introduction,* SSE, Stockholm

Brown Kruse, J., Rassenti S., Reynolds, S. and Smith, V. L. 1994. "Bertrand-Edgeworth Competition in Experimental Markets." *Economic Journal,* 113, 495-524.

Camerer, C. 2003. *Behavioral game theory: experiments in strategic interaction*. Princeton University Press, Princeton.

Cournot, A. 1838. *Recherches sur les Principes Mathematiques de la Théorie des Richesses.* L. Hachette, Paris.

Edman, J. 2005. "Capabilities of Experimental Business Gaming." *Developments in Business Simulation and Experiental Learning*. 32,104-109.

Grubbström, R. 1972. *Economic decisions in space and time: theoretical and experimental inquiries into the cause of economic motion*. Gothenburg Business Administration Studies, Göteborg.

Huck, S., H. Normann, H. and J. Oechssler. 2004. "Two are Few and Four are Many. Number Effects in Experimental Oligopolies." *Journal of Economic Behavior and Organization*. 53, 435-446.

Ståhl, I. 1993. "A Small Duopoly Game based on Discrete-Event Simulation." In Pave, A. (Ed.) *Modelling and Simulation ESM 1993*. SCS, Lyon, 295-301.

Ståhl, I. 2007. "Teaching Simulation to Business Students – Summary of 30 Years' Experience. In S. G. Henderson, B. Biller, M.-J. Hsieh. J. Shortle, J.D. Tew and R.R. Barton (Eds.) *Proceedings of the 2007 Winter Simulation Conference.* ACM, Washington, D.C., 2327-2335.

Ståhl, I. 2010. "Discrete Event Simulation in the Study of Energy, Natural Resources and the Environment." In E. Björndal, M. Björndal, P.M. Pardalos and M. Rönnqvist (Eds.) *Energy, Natural Resources and Environmental Economics*. Springer-Verlag, Berlin, 509-521.

Thorlund-Petersen, L. 1990. "Iterative Computation of Cournot Equilibrium." *Games and Economic Behavior*, 2, 61-75.

AUTHOR BIOGRAPHY

INGOLF STÅHL is Professor Emeritus at the Stockholm School of Economics, Stockholm, of a chair in Computer Based Applications of Economic Theory. He has taught GPSS for forty years at universities in Sweden, Norway, Germany, Latvia and the USA. Based on this experience, he has led the development of the aGPSS educational simulation systems. His email address is <ingolf.stahl@hhs.se>. His aGPSS system is at <http://www.agpss.com/>.

Determination of Factors Influencing the Decision on Purchasing Organic Food

Walailak Atthirawong
Department of Statistics, Faculty of Science
King Mongkut's Institute of Technology Ladkrabang, Bangkok10520, Thailand
Email:walailaknoi@gmail.com

KEYWORDS

Environmental issues, Purchase behaviour, Organic food, Logistic regression, ROC analysis

ABSTRACT

Since there has been a rising awareness about health, food safety and environmental issues, the demand for organic food has grown rapidly among consumers. However, information and profiling of consumers in Bangkok, which is the capital of Thailand, are not yet well reported. As such, this study aims at analyzing factors affecting consumers' organic food purchase intention. Cross-sectional data were carried out with the respondents in Bangkok through questionnaires. Data were analyzed using a logistic regression model employed to test the proposed hypotheses. The results revealed that education level and attitude towards place had positively influenced consumer's decision in buying organic food. Although the respondents had high score on attitude about healthfulness and food safety towards organic food, it was found that they had little knowledge about them. Evidence provided in this study could be employed as a reference information for policy makers and marketers regarding such issue.

INTRODUCTION

Over the past decades, it is widely recognized that patterns of food consumption have rapidly changed as a consequence of consumers becoming concerned on environmental sustainability, food safety and health issues. Not only does it affects consumer's health but unsafe food can also lead to great economic impacts on people in the country. Along with this trend, nowadays organic agriculture has been expanding quickly everywhere across the globe. According to the latest survey by FiBL, statistical information on certified organic agriculture is now available from 172 countries (Willer and Lernoud 2016). Thailand, which is known as "The Food Basket of Asia", not only produces agricultural products but is also concerned with food safety standards (Supaphol 2010). Organic agricultural development has been recently included in the top five "urgent agendas" of Ministry of Agricultural and Cooperative (Wai 2016). The total organic agricultural land is expanding as the demand for organic food within the country has grown rapidly. However, similar to other developing countries, local market for organic food is still a tiny market shared. Understanding about attitudes of customers and purchasing behaviors intention towards organic food may help relevant agencies facilitate and develop programs to drive organic consumption which are free from chemical residues, toxic elements and pesticides (Fotopoulos and Krystallis 2002). Since information and profiling of consumers behavior is infant and not yet well reported in Thailand, the primary objective of this research is to investigate purchasing behavior intention towards organic food of consumers in Bangkok. This study focuses on Bangkok as Bangkok is the capital and the largest city in Thailand with about 5.69 million registered residents (https://en.wikipedia.org/wiki/Bangkok). In sum, the findings of this study will be of great contribution for retailers and marketers in providing fruitful information to stimulate organic food purchasing behavior and for producers in developing domestic organic production to meet consumers' requirements.

The structure of the remainder of this paper is, therefore, organized as follows. The following section reports definition of organic food and organic products in Thailand, followed by literature survey adapted to this research. Next, research methodology is presented and then results obtained from the survey are described. Finally, conclusion remarks and discussion of the results as well as practical implications for relevant agencies are then discussed in the last section of the paper.

LITERATURE REVIEW

Definition of Organic Food

The board definition of "organic food" is food which is produced by methods that comply with the standards of organic farming system (Allen and Albala 2007). Organic farming refers to a farming system which enhances agro-ecosystem health, including biodiversity, biological cycles and soil biological activity (United Nations Food and Agriculture Organization 1999). A definition of organic farming varies considerably among countries depending upon regulations. Organic food is

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

produced in a way that the production will not contaminate with any artificial or chemical fertilizers, pesticides synthetic or the use of genetically modified organism. Specifically, organic food are also usually processed using natural production system which does not have effects on the environment. Uma and Selvam (2016) divided organic food into three major categories i.e. organic vegetables and fruits, organic dairy products (such as milk, cheese and ice cream, etc.) and organic fish & meat.

Organic Food in Thailand

Thai organic agriculture has been introduced to Thailand in the 1970s resulting from green revolution. At present, Thailand's organic sector is probably in the growth stage of development. The total organic agricultural land is expanding as the demand for organic food within the country has grown rapidly. Recently statistics has disclosed that production areas under organic farming in Thailand increased from 1,6483 hectares in 2000 (Green Net/ Earth Net Foundations 2013) to 33,600 hectares in 2015 (FIBI & IFOAM-Organic International 2016). Organic agricultural development has recently been included in the top five "urgent agendas" of Ministry of Agricultural and Cooperative (Wai 2016). The emerging attractiveness of organic and environmentally friendly products in Thailand has resulted from a combination of reasons. At the beginning, it starts from people's general concern with healthy living and food safety, but later on the crisis faced in the farm sectors has enforced sustainable development of agricultural production system. Currently, the trend is transforming into a broader scheme covering environment awareness, ecology and biodiversity. Nonetheless, organic products is only accounting for 1 per cent of Thai food market.

In 2014, organic products produced at 71,847 tons accounted for 2,331.55 million, of which about three fourths was for the export market. Most of organic products were exported to Europe and North America (Kongsom and Panyakul 2016). It is revealed that the production process for organic farming in Thailand is still being in a conventional way and uses simple and limited technology. Furthermore, the products to produce are still basic and unprocessed, for instance, rice, vegetable and fruit whereas processed organic produce, as finished consumer products, are quite few in the market due to insufficient raw materials. The supply of organic products is not continuous to support the plants, many of them are imported in unprocessed commodities.

Previous Studies

Several research have examined consumers' behavior towards environmentally friendly products in Thailand. However, there has been limited academic study on consumers' purchasing behavior with regard to organic products (Sangkumchaliang and Pakdee 2012). The followings are the review of previous researches in organic food conducted in Thailand five years ago (during 2012-2016):

Sangkumchaliang and Huang (2012) explored consumers' perceptions and attitudes towards organic food products in Chiang Mai province of Thailand. Questionnaire was employed to gather information from 390 respondents who bought products in different markets i.e. the Multiple Cropping Centre, the Royal Project shop and Top supermarkets. Chi-square test was used to analyze data. An expectation of a healthy and environment friendly was the main reason for purchasing organic food products. There was a significant difference between the groups of buyers and non-buyers in their demographic characteristics. The study also highlighted that information available is a key barrier to increase market share of organic products.

Sangkumchalian and Pakdee (2012) investigated factors affecting consumers intention to purchase agricultural product based on the Theory of Planned Behavior (TPB) model. Data were gathered through a questionnaire conducted in supermarkets and a fresh market in Khon Kaen province where organic or safe food produce is available. Descriptive statistics and Structural Equation Modeling (SEM) were used to analyze data. The results indicated that positive attitude towards organic agriculture had increased consumer intention to buy organic agricultural products. Subjective norm and perceived behavior control directly affected consumers' intention to purchase organic agricultural products. Additionally, demographic characteristics and the reasons whether to purchase an organic product or not were included in this study.

The study of Pomsanam et al. (2014) explored factors driving Thai consumer intention to purchase organic food. The objective of this research is to analyze factors affecting consumers in Sa Kaeo province of Thailand intent to buy organic food. Data were collected via a questionnaire with 400 participants. Factor analysis and multiple regression analysis were employed to analyze data. The findings of this research indicated that factors driving consumers purchase intention were subjective norms, environmental protection, trust in label, food quality and availability and convenience in accessing organic food. The suggestion of this study was to provide knowledge of environment awareness to consumers through social media. Moreover, relevant agencies should provide a more effective organic certification system so that consumers can check whether products are from real organic process.

Songkroh (2015) explored the relationship of determinants of demand for organic agriculture products. Data were gathered from respondents in

Chiang Mai by using questionnaires both in Thai and English versions. Multiple regression analysis was applied to analyze information. The study showed that determinants of demand for organic agriculture products were price, price of substitution product, price of complimentary product and income level of consumer. Only price of complimentary product and income level of consumer had positive relationship with demand. It was found that R^2 equaled to 0.231, meaning that the equation can explain the demand in 23.1 per cent.

Recently, Ueasangkomsate and Santiteerakul (2016) presented the relationship between Thai consumers' attitudes and intention to buy organic food. A self-reported questionnaire survey was conducted on 316 respondents across Thailand. Pearson correlation coefficient was conducted to test the relationship between attitude towards organic food and intention to buy. The study showed that the local origin is the highest correlation to buying intention organic food in positive way, followed by animal welfare and environment attribute. The authors also suggested researchers to study more in buying behaviors on organic food of consumers.

In addition, Kongsom and Panyakul (2016) investigated the production and market of certified organic products in Thailand. Data were collected via a survey and in-depth interview. Secondary data from organic agriculture certification body and publications were also explored. The study revealed that the largest exports of certified organic products were processed food accounting for 66.1 per cent of total export value, followed by organic rice accounting for 30.4 per cent. Modern trade was the largest domestic channel (59.48 per cent of total domestic sales, followed by green shop (29.47 per cent) and food establishment (5.85 per cent). In order for Thailand to become a center of organic farming and trading within ASEAN, there is a need to gain more policy supports and appropriate strategies from the government and relevant agencies.

Kongsom and Kongsom (2016) investigated the awareness, knowledge and consumer behavior towards organic products in Thailand. Data were collected across the country with 2,575 consumers over the age of 20 years who intended or made purchases from 1) green shops, 2) supermarkets with branches, and 3) green markets using a purposive sampling technique. Descriptive statistics were employed to analyze data from the questionnaires. The outputs indicated that more than 92 per cent of consumers were aware of organic agriculture production but had less knowledge about it. Almost half of respondents had confused between the food safety logo and the certified organic logo, and whether GMO was allowed in organic agriculture practice or not. Respondents also felt that processed organic products in Thailand were relatively small in quantity.

Even though there have been few studies in Thailand that provided empirical evidence in this area, it is wise to further investigate to a deeper understanding in underlying motive driving purchase intention of organic food products. This type of research will allow richer insights into customers motivation for purchase decision. Likewise, such research will be of assistance in providing more reliable and accurate results. Additionally, it will be of great value especially to policy makers by developing sustainable strategies specific to the target groups.

MATERIAL AND METHODS

Sampling Technique

Data for this research were carried out by means of hand-delivered questionnaire during November and December 2016. The study's scope was with consumers whose ages are more than or equal to 18 years old. The number of respondents in Bangkok (N) was 5,696,409 people in 2015 (https://th.wikipedia.org/). Due to the population being enormous, a total of 384 sample sizes (n) were selected for this study (Yamane 1973).

Research Instruments

Data employed in this study were obtained from structured questionnaires design, which is considered as one of the most common and widely used research tools in the field of survey research. The questionnaire was divided into four sections.

The first part includes socioeconomic and demographic variables (inquiries about general information, socio-economic characteristics of organic food consumption. In the second part, the respondents were asked whether they have knowledge and understanding about organic food products. The third part was the main part of the questionnaire. A question of 33 items was developed regarding literature review and previous studies. All statements were formulated on 5-point Likert-type scale (Wolfer 2007) ranging from "strongly disagree (1)" to "strongly agree (5)". The final part is inquiries about opinions and ideas on how to promote organic food products in practices.

Validity and Reliability

Prior to data collection, the quality of the research instrument or questionnaire was examined by assessing the face validity and the reliability (Hair et al. 2006). Cronbach's Alpha coefficient was employed to evaluate the quality of the survey instrument in Section 3 whether it is appropriate to complete the goal it was used for or not. Thirty respondents took part in pre-test process. As shown in Table 1, the total Cronbach's Alpha coefficient is reached at 0.916 which implies that the tool is sufficient and reliable for being used to collect data in primary source (Creswell 2002).

Table 1: Relaibility Statictics

Conbrach's Alpha coefficient	Cronbach's Alpha Based on Standardized items	Number of items
.916	.926	33

Analytical Techniques

Descriptive statistics i.e. mean, standard deviation and percentage were used to explain demographic characteristics, the level of knowledge of organic food products as well as the level of attitude of the sample. A binary logistic regression model was adopted to determine the extent to which selected demographic characteristics, knowledge on organic food products, attitude and behavior influence customers' intention to buy organic food. Binary logistic regression, also called a logit model, is usually employed when the dependent variable is dichotomous and the independent variables are either continuous or categorical variables. Specifically, it is used to model the relationship between the categorical dependent variable and one or more independent variables by estimating probabilities using a logistic function. Normally, the outcome in logistic regression analysis is coded as 0 or 1, where 1 indicates that the outcome of interest is present, and 0 indicates that the outcome of interest is ignored (Hair et al. 2006). As such, if p is defined as the probability that the outcome is 1, the multiple logistic regression model can be expressed as follows:

$$\hat{p} = \frac{\exp(b_0 + b_1X_1 + b_2X_2 + ... + b_pX_p)}{1 + \exp(b_0 + b_1X_1 + b_2X_2 + ... + b_pX_p)} \quad (1)$$

where $\hat{p}$ is the expected probability that the outcome is present, $X_1, \ldots, X_p$ are independent variables and b_0 through b_p are the regression coefficients. The multiple logistic regression model is sometimes written differently. In the following form, the outcome is the expected log of the odds that the outcome is presented in equation (2).

$$\ln\left(\frac{\hat{p}}{(1-\hat{p})}\right) = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p \quad (2)$$

An iterative likelihood methods is normally employed to estimate the regression coefficient (Hair et al. 2006). In order to discover the effect of the explanation variables on the decision to buy organic food, descriptions of each variable can be explained in Table 2.

Table 2: Dependent and Independent Variables and Scale Used in the Model

Variables	Description	Variable scale
X_1	Gender	1 = male; 0 = otherwise
X_2	Age	1 = above or equal to 41 years old; 0 = otherwise
X_3	Marital status	1 = married; 0 = otherwise
X_4	Education level	1= undergraduate level or above; 0 = otherwise
X_5	Income per month	1 = more than 30,000 baht; 0 = otherwise
X_6	Employment	1 = employment; 0 = Student/ unemployment
X_7	Type of resindent	1 = home; 0 = otherwise
X_8	Type of information	1 = social network; 0 = otherwise (e.g. television/radio,etc.)
X_9	Shopping place	1 = heath shop; 0 = otherwise
X_{10}	Attitude toward healthfulness	1 = strongly disagree; 5 = strongly agree
X_{11}	Attitude toward food safety	1 = strongly disagree; 5 = strongly agree
X_{12}	Attitude towards taste	1 = strongly disagree; 5 = strongly agree
X_{13}	Precieved value	1 = strongly disagree; 5 = strongly agree
X_{14}	Promotion	1 = strongly disagree; 5 = strongly agree
X_{15}	Image	1 = strongly disagree; 5 = strongly agree
X_{16}	Attitude towards place	1 = strongly disagree; 5 = strongly agree
Y	Decision to buy organic food	1 = yes; 5 = no

RESULTS

384 surveys were distributed and 349 questionnaires were returned during the data collection period. 320 questionnaires in total were completed and included in the analysis, which produced a high response rate of 83.3 per cent. The results of demographic profile reveal that 62.5 per cent of the respondents were female. The majority of the residents (69.1 per cent) consists of single people. Half of them (49.4 per cent) had achieved undergraduate level, followed by lower undergraduate level (29.4 per cent). Over half of the sample (61.6 per cent) lived in their own houses. The study found that a majority (67 per cent) of the respondents had received information about the organic food products from social network, followed by from television /radio (17.8 per cent).

Knowledge on Organic Food Products

Results from Table 3 show that a mean score of knowledge on solid waste management equals to 4.77 (S.D.=1.56) out of 10. It indicates that the respondents still had less knowledge on organic food products.

Table 3: Knowledge on Organic Food Products

Level of knowledge in organic food products	n	$(\overline{X})$	S.D.
Total	320	4.77	1.56

Consumers'Attitudes and Influecning Factors toward Puchasing Organic Food Products

The mean score of each variable of attitudes and influencing factors toward organic food products are showed in Table 4. Compare to 7 variables, the table indicates that promotion (X_{14}) has the highest score, followed by attitude toward healthfulness (X_{10}) and attitude toward food safety (X_{11}), respectively. While attitude towards taste (X_{12}) receives least concerns from the respondents. However, all of them are ranged in the class interval of 3.48 to 4.18. It is implied that the respondents had high level on attitudes and influencing factors toward organic food**.**

Table 4: Consumers's Attitudes and Purchase Behaviour toward Organic Food

Varibles	Mean	S.D.	Rank
Attitude toward healthfulness (X_{10})	4.092	0.616	2
Attitude toward food safety (X_{11})	3.977	0.729	3
Attitude towards taste (X_{12})	3.478	0.924	7
Precieved value (X_{13})	3.940	0.613	4
Promotion (X_{14})	4.178	0.671	1
Image (X_{15})	3.644	0.744	6
Attitude towards place (X_{16})	3.659	0.634	5

Model Results

a.Partial Test and Model Building

Out of 318 respondents about 320 were involved into the model as other cases are deleted for having missing information.The empirical results using *forward stepwise logistic regression model* based on the survey data are displayed in Table 5. The model is employed to predict whether consumers purchase organic food with respect to factors affecting the decision. The results of the survey reveal that education level (X_4) and attitude towars place (X_{16}) were important predictors of decision to buy organic food.

Table 5: The Result of Logistic Regression Analysis

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	X_{16}	.737	.190	14.999	1	.000	2.090
	Constant	-3.005	.713	17.755	1	.000	.050
Step 2[b]	X_4	.328	.101	10.529	1	.001	1.388
	X_{16}	.669	.195	11.755	1	.001	1.952
	Constant	-3.630	.767	22.433	1	.000	.027

a. Variable(s) entered on step 1: X_{16}

b. Variable(s) entered on step 2: X_4

b.Significance Test Model

Based on Chi Square calculation as demonstrated in Table 6, the significance value is lower than 0.05 which means that H_0 is rejected. It indicates that the model is very meaningful and passes the minimum standard which suggests that education level, attitude towards taste and image significantly influence decision to buy organic food.

Table 6: Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	15.957	1	.000
	Block	15.957	1	.000
	Model	15.957	1	.000
Step 2	Step	11.303	1	.001
	Block	27.260	2	.000
	Model	27.260	2	.000

c.Odds Ratio Interpretation

Table 5 also displays odd ration value which is an important imformation to explain the influence of the independent variables on the increasing or decrasing of probability to occur the event measure by the dependent variable.Based on the statistically significant coefficients, the findings show that the respondents who had high education level were 38.8% more intend to buy organic food products than respondents' who had a lower education level in a positive way.Additionally, consumers who had high concern on attitude towards place for selling organic food products were 95.2% more intent to purchase compared to consumers who were less concerned.

d. ROC Analysis

ROC Curve demonstrates the amount of area covered by the predictive model graphically. It can be seen in Figure 1 how true positive rate (specificity) is plotted against the false positive rate (1-specificity). Although the curve is above base line but is not close to upper left coner which the classifier performance equals to 66.2 per cent.

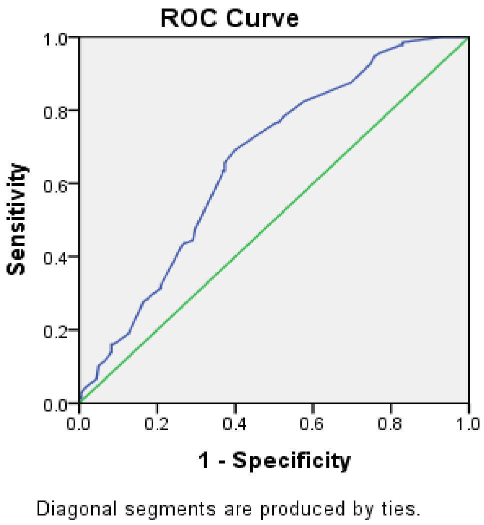


Figure 1: ROC Curve

CONCLUSION AND RECOMMENDATION

The aim of this study was to investigate purchasing behavior intention towards organic food among Bangkok consumers. The empirical findings indicate that the respondents in Bangkok have high attitudes on organic food products; however, they still have little knowledge and understanding about organic food products. The results of the logistic regression analysis demonstrate that the socio-demographic profile of organic food buyers is education level. This findings is consistent with similar studies on organic or green products. For instance, Rezai et al. (2011) found that demographic characteristic i.e. education level significantly influence Malaysian consumers to purchase green produced food. Magnusson et al. (2001) also claimed that people who have higher education are more likely to convey positive attitudes towards organic products and also are willing to pay for organic food. However, it seems that in this study income does not positively correlate to the decision to buy organic food which is in line with the study of Fotopoulos and Krystallis (2002). Generally, having higher income does not necessarily imply that customers have higher likelihood of buying organic food.

In addition, based on the empirical results, it is interesting to note that attitude towards place was a significantly positive relationship with the decision on buying organic food. According to Kotler and Keller (2009), place decision involves activities that make products available to end customers. Convenience location, easy accessibility and comfortable atmospherics could determine consumers' buying behavior resulting in increasing consumers' purchase.

Surprisingly, this empirical study did not show that decision to buy organic food is influenced by other factors, especially attitude toward healthfulness and food safety. These findings may suggest that although respondents had high attitude towards health and safety organic food, they may have fewer health benefits from organic food such as health improvement. The evident shows that the respondents still had little understanding and knowledge about it. Accordingly, the knowledge and public awareness regarding organic food should be created and promoted by policy makers and marketers for stakeholders especially target consumers. More importantly,it is necessary to emphasis on environmental issues, food safety, eco-friendly and perceived value but not for profit purpose in order to increase understanding and raising demand in the future.

Although the study had been conducted only in the capital of Thailand, the findings can be applied or extended in other cities in Thailand to obtain a reliable and more accuracy results. Furthermore, it is wise to carry out the similarly studies in other countries such as in ASEAN region which can enhance the development of organic markets. As suggested by Petrescu and Petrescu-Mag (2015), learning and understanding more about target customers will help marketers and retailers create appropriate strategies in order to sustainable behaviors and encourage their development for further environmental benefits. Finally, the interaction among psychological factors, personal norms, social factors, intention behavior, knowledge and other motives factors is also remarkable to investigate.

ACKNOWLEDGEMENT

The author would like to thank Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, the sponsor of this research funding. Appreciation also goes to all Bangkok respondents who participated in this survey and all survey teams who assisted in administrating surveys.

REFERENCES

Allen, G. and Albala, K.. 2007. *The Business of Food: Encyclopedia of the Food and Drink Industries*. Greenwood Press, USA.

Bellows, A. C. and Onyango, B., Diamond, A., and Hallman, W. K. 2008. "Understanding Consumer Interest in Organics: Production Values vs. Purchasing Behaviour". *Journal of Agricultural & Food Industrial Organization*, Vol 6,1-28.

Creswell, J.W. 2002. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. 2nd ed. SAGE Publications, Inc.

Fawcett, T. 2006. "An introduction to ROC analysis". *Pattern Recognition Letters*, Vol. 27,861–874.

Fotopoulos, C. and Krystallis, A .2002. "Purchasing motives and profile of the Greek organic customer: a countrywide survey". *British Food Journal*, Vol.104, No.4, 735-765.

FIBI&IFOAM-Organic International. 2016. "The World of Organic Agricultural-Statistic and Imerging Trends". https://shop.fibl.org/fileadmin/documents/shop/1698-organic-world-2016.pdf.

Green Net / Earth Net Foundations. 2013. “Situation of Thai Organic Agriculture” (in Thai language).

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E. and Tatham, R.L. 2006. *Multivariate Data Analysis*. New Jersey: Pearson Prentice Hall.

Kotler, P. and Keller, K. L. 2009. *Marketing Management*. Pearson Prentice Hall.

Kongsom W. and Kongsom, C. 2016. “Consumer Behavior and Knowledge on Organic Products in Thailand”. *Engineering and Technology International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, Vol.10, No.8,2524-2528.

Kongsom, C.and Panyakul,V. 2016. “Production and Market of Certified Organic Products in Thailand”. *The Proceedings of 18th International Conference on Organic Agriculture and Food Security,*Venice, Italy.

Krystallis, A. , Marco V., George C. and Toula P. 2008. “Societal and individualistic drivers as predictors of organic purchasing revealed through a portrait value questionnaire (PVQ)-based inventory”. *Journal of Consumer Behaviour,* Vol. 7. 164-187.

Magnusson, M., Arvola, A., Koivisto Hursti, U., Aberg, L. and Sjoden, P. 2001, "Attitudes towards organic food among Swedish consumers", *British Food Journal,* Vol. 103, No. 3, 209-26.

Petrescu, D. C. and Petrescu-Mag, R.M. 2015. “Organic Food Perception: Fad, or Healthy and Environmentally Friendly? A Case on Romanian Consumers”. *Sustainability*, Vol. 7, 12017-12031; doi:10.3390/su70912017.

Pomsanam,P. Napompech, K. and Suwanmaneepong, S. 2014. “Factors Driving Thai Consumers’ Intention to Purchase Organic Foods”. *Asian Journal of Scientific Research,* Vol. 7, 434-446.

Rezai, G., Mohamed, Z., Shamsudin, M. N. and Phuah, K. T. 2011. “Demographic and Attitudinal Variables Associated with Consumers’ Intention to Purchase Green Produced Foods in Malaysia”. *International Journal of Innovation Management and Technology*, Vol.2, No.5, 401-406.

Sangkumchaliang, P. and Huang, W.H. 2012. “Consumers’ perceptions and attitudes of organic food products in Northern Thailand”. *International Food and Agribusiness Management Review* , Vol. 15, No.1, 87-102.

Sangkumchaliang, P. and Pakdee, P. 2012. “Consumers' Intention to Purchase Organic Agricultural Product in Northeast Thailand”. https://ora.kku.ac.th/db.../11279-00000-abstract_file.pdf. [available online] [access 20 November 2016].

Songkroh, M. 2015. “Demand for Organic Agriculture Products in Chiang Mai”. *Proceedings of International Conference on Management Finance Economics,* July 11-12. https://www.innovativeresearchpublication.com/.../pdf%2034.pdf. [available online] [access 20 November 2016].

Supaphol, S. 2010. “Status of Food Safety and Food Security in Thailand: Thai's Kitchen to the World". *Journal of Developments in Sustainable Agriculture*, Vol. 5, 39-46.

Ueasangkomsatea, P. and Santiteerakulb, S. 2016. “A study of consumers’ attitudes and intention to buy organic foods for sustainability”. *Procedia Environmental Sciences,* Vol.34, 423 – 430.

Uma, R. and Selvam,V. 2016. “Customer Attitudinal and Perceptions towards Purchasing Organic Food productions: a Critical Review of Literature from 2005 to 2015”. *International Journal of Applied Business and Economic Research*, Vol 14, No.10,7179-7197.

United Nations Food and Agriculture Organization (FAO). 1999. *Understanding the Codex Alimentarius*. Rome, Italy: FAO.

Wai, O.K. 2016. “The World of Organic Agriculture 2016 : Summary”. *Research Institute of Organic Agriculture (FiBL and IFOAM)-Organic International.* In the World of Organic Agriculture : Statistics and Emerging Trends 2016, 172-188.

Willer. and Lernoud, J. 2016. “The World of Organic Agriculture 2016 : Summary”. *Research Institute of Organic Agriculture (FiBL and IFOAM)-Organic International.* In the World of Organic Agriculture : Statistics and Emerging Trends 2016, 24-32.

Wolfer, L. 2007. *Real Research: Conducting and Evaluating Research in the Social Sciences.* Boston, Pearson/Allyn and Bacon.

Yamane, T. 1973. *Statistics: An Introductory Analysis*. 3rd ed. New York, Harper and Row Publications.

http://www.greennet.or.th/article/411.[available online] [access 9 January 2017].

https://en.wikipedia.org/wiki/Bangkok.[available online] [access 31 Ocober 2016].

AUTHOR **BIOGRAPHIES**

WALAILAK ATTHIRAWONG is Associate Professor of Operations Research at Faculty of Science, King Mongkut’s Institute of Technology Ladkrabang (KMITL) in Thailand. She received doctoral degree from the Univeristy of Nottingham in Manufacturing Engineering and Operations Management. She is actively engaged in research on logistics and supply chain management, simulation, multi-criteria decision making, applied statistics and optimization. Her e-mail address is : walailaknoi@gmail.com**.**

LIFETIME PROBABILITY OF DEFAULT MODELING FOR HUNGARIAN CORPORATE DEBT INSTRUMENTS

Tamás Kristóf and Miklós Virág
Enterprise Finances Department
Corvinus University of Budapest
Fővám tér 8, 1093 Budapest, Hungary
E-mail: tamas.kristof@uni-corvinus.hu, miklos.virag@uni-corvinus.hu

KEYWORDS

IFRS 9, credit risk, probability of default, expected loss, Markov chain

ABSTRACT

The paper attempts to provide forecast methodological framework and concrete models to estimate long run probability of default term structure for Hungarian corporate debt instruments, in line with IFRS 9 requirements.

Long run probability of default and expected loss can be estimated by various methods and has fifty-five years of history in literature. After studying literature and empirical models, the Markov chain approach was selected to accomplish lifetime probability of default modeling for Hungarian corporate debt instruments.

Empirical results reveal that both discrete and continuous homogeneous Markov chain models systematically overestimate the long term corporate probability of default. However, the continuous non-homogeneous Markov chain gives both intuitively and empirically appropriate probability of default trajectories. The estimated term structure mathematically and professionally properly expresses the probability of default element of expected loss that can realistically occur in the long-run in Hungarian corporate lending. The elaborated models can be easily implemented at Hungarian corporate financial institutions.

INTRODUCTION

Credit risk analysis of corporate financial instruments is a central issue of corporate finances from theoretical and empirical points of view. One of the most important research fields, which is at the same time the fundamental credit risk parameter of the debtors, is the probability of default (PD) that can be quantified both from average PDs mapped to rating classes, or by using statistical PD estimation models.

As an industrial standard, PD models have traditionally been elaborated using cross sectional or some years of historical data, applying multivariate statistical classification methods, estimating PD for one year horizon. It has a rich literature and empirical results also in Hungary (see inter alia Kristóf 2008; Virág and Fiáth 2010; Kristóf and Virág 2012; Virág et al. 2013; Virág and Nyitrai 2014, Nyitrai 2015). Static, one-year PD estimation approach has met supervisory authority expectations and professional best practice for a long time.

However, as an aftermath of the recent financial crisis, substantial regulatory pressure has been made on the further development of credit risk models, laying emphasis on the timely recognition of credit losses, underpinning the establishment and implementation of IFRS 9 standards, coming into effect on 1st January 2018 (IASB 2014). The forward looking impairment model of IFRS 9 calls for the quantification of lifetime credit loss, if significant credit risk deterioration happens to the debtors, which requires lifetime PD modeling.

According to naïve approach, the constant annual PD might be extended to multiple periods. However, on the basis of practical experience, it is easy to see that the time behavior of PD is not constant and non-linear, thereby more complex modeling is necessary.

The aim of this paper is to publish forecast methodological framework and concrete models to estimate long run PD term structure for Hungarian corporate debt instruments, in line with IFRS 9 requirements.

METHODOLOGICAL APPROACHES

Lifetime PD modeling has fifty-five years of history in literature. According to our best knowledge, the first lifetime expected loss model was published by Cyert et al. (1962) for accounts receivables, applying the Markov chain method. Consideration behind the application of discrete Markov chain was the fact that accounts receivables month by month migrate among different delinquency states. Movements among delinquency states were described by migration matrices or transition matrices.

The structural approach of corporate default modeling appeared in the 1970s, the theoretical and methodological foundation of which was formulated by Black and Scholes (1973); Merton (1974); Black and Cox (1976) for corporate bonds. The pioneer publications assumed that the behavior of corporate receivables depend on the asset quality under certain conditions (interest rate, capital structure etc.). It was a difference, however, that Merton (ibid.) equated the time of default with the maturity of bonds, whereas according to Black and Cox (ibid.) a company might become defaulted any time before maturity. The default

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

event and its probable time were approached by a perceived incident, when the asset quality of a company first time fell behind a predefined threshold.

Examination of relationships between term and default spreads lead to the appearance of the term structure models (Fama 1986). The three most important findings of these models were that default spreads stand in reverse ratio with term, they depend on economic cycles, and are not necessarily monotonous.

The study of Jarrow et al. (1997) represented a milestone in the literature that elaborated a continuous Markov chain model for corporate bonds, taking into account the credit rating. Changes of credit rating formulated the states of the Markov chain. The transition matrix expressed the probability of remaining in the existing rating class, and the migration to other rating classes.

Within the framework of a comparative analysis Lando and Skodeberg (2002) compared the performance of the continuous multistate Markov model to the traditional, cross sectional, discrete Markov model. The authors concluded that the continuous model outperformed the discrete model.

A problem of applying Markov chain in practice emerged from the observation that the behavior of data modeled by Markov chain is often non-homogeneous. Bluhm and Overbeck (2007) generated PD term structures using homogeneous and non-homogeneous, continuous Markov chains, and compared the results to the fifteen years of cumulated actual default rates published by Standard&Poors. Results with the non-homogeneous model were much better, from which it was concluded that the homogeneity assumption could be set aside.

Since the end of the 1990s – in parallel with the development of retail scoring models – survival analysis models have begun to spread, facilitating the estimation in the function of time, when a client is expected to default (Banasik et al. 1999). Survival time can be estimated with hazard function, which forecasts the magnitude of PD change for any future time.

A great part of PD term structure literature attempted to estimate the term structure from market data (Duffie and Singleton 1999; Jarrow 2001; Longstaff et al. 2005). PDs are often implied from default swap or bond data. Based on practical experience, however, since the majority of loan portfolios contain financial instruments not traded in secondary markets, there is no market data, particularly for loans, from which it would be possible to derive PDs.

A number of studies involved macroeconomic variables and economic cycles into PD modeling to ensure that the relationship between actual economic environment and credit risk is taken into account. Changes in credit risk state are usually explained by industry, location and changes in economic cycle (Gavalas and Syriopoulos 2014).

After studying various literature and empirical models, the Markov chain approach was selected to accomplish lifetime PD modeling for Hungarian corporate debt instruments. The formal description of the method is provided in the next chapter.

MARKOV CHAIN MODELING

A series of random variables formulate a Markov chain, if an observation is in any period in an initial i-th state, and the probability that it migrates to a j-th state in the next period, exclusively depends on the value of i. Let $(X_t)_{t\geq 0}$ denote the series of random variables with {1, 2, …, K} fixed number of classes, where K denotes the default state. The series is a finite first order Markov chain, if:

$$P\left(X_{t+1}=j \middle| X_0=x_0,\dots,X_{t-1}=x_{t-1},X_t=i\right)=P(X_{t+1}=j|X_t=i) \quad (1)$$

for each t, and i, j∈{1, 2, …, K}

$P_t(i,j)=P(X_{t+1}=j|X_t=i)$ means the probability of transition in t-th period from i-th state to j-th state in (t+1)-th period, and represent the element of the K×K size P_t transition matrix.

The Markov chain is stationary, if $P_t=P$ for each t≥0. Then the transition matrices are identical in each time. In this case any multi-period transition matrix can be calculated by raising the annual transition matrix to power:

$$P(X_{t+k}=j|X_t=i)=P^k(i,j) \quad (2)$$

The continuous X_t Markov chain is timely homogeneous, if for each i, j state and t, s≥0 times:

$$P(X_{t+s}=j|X_t=i)=P(X_s=j|X_0=i) \quad (3)$$

In case of continuous Markov chain a transition matrix between 0-th and t-th period can be estimated by exponentiating the generator matrix. G generator matrix is such a K×K matrix, where $P(0,t)=\exp(Gt)$. Gt is scalar product, and the exponential function is:

$$\exp(Gt)=\sum_{n=0}^{\infty}\frac{t^n}{n!}G^n \quad (4)$$

The generator matrix has the following characteristics:

$G_{i,j}=0$ for each i≠j

$G_{i,i}=-\sum_{j\neq i}G_{i,j}$.

The elements of the generator matrix relate to the time spent in each rating class. The remaining time in i-th class can be characterized by exponential distribution having $-G_{i,i}$ parameter. Timely homogeneous probabilities of transitions in any horizon can be expressed in the function of the same generator matrix. However, in case of non-homogeneous transitions, the generator matrix depends on time, and can be formulated as follows:

$$P(0,t)=\exp\left(\int_0^T G(t)dt\right) \quad (5)$$

EMPIRICAL RESEARCH

In Markov chain modeling the first research task is to construct a transition matrix based on observed changes of states. In case of corporate credit risk modeling it generally means an annual transition matrix, reflecting the change in rating. The transition matrix can be assembled from internal or external data. It is important to note, however, that only financial institutions possess appropriate internal data for this modeling purpose. Since it is not possible to publish models using internal banking data, we have considered the long run global corporate annual probabilities of transitions of Fitch and Standard&Poors rating agencies.

In line with the objective of the paper, Hungarian idiosyncrasies should be considered in modeling. Since the credit rating history of the best Hungarian corporations strongly correlate with the sovereign rating of Hungary, and in recent years it was in the BBB and BB classes at both agencies, it is assumed that neither Hungarian company can be better than the credit risk characteristics of companies in the BBB classes. Accordingly all ‘A’ category rating classes were excluded from both transition matrices, thereby the ten remaining classes plus the default class represented the object of analysis. Furthermore it was necessary to handle the problem of withdrawn rating. Assuming that withdrawn rating does not mean upgrading or downgrading, the matrices were normalized by simple scaling. The so constricted and normalized transition matrices showed very corresponding tendencies and results. For further calculations the average transition matrix of Fitch and Standard&Poors was used (Table 1). The PD of each class is reflected by the probability of transition to the D class. If the classification of debtors were already default in the initial period of transition, both annual and lifetime PD of such debtors are 100%. The default class is absorbing state, regardless the fact where the migration is from.

On the basis of the transition matrix a discrete homogeneous, a continuous homogeneous and a continuous non-homogeneous Markov chain model have been elaborated.

Discrete Homogeneous Model

In line with the assumption system of the discrete Markov chain, probabilities of transitions for future terms can be estimated, by raising the transition matrix into power. PD term structure was estimated for twenty years. Matrix multiplication results in cumulated PDs.

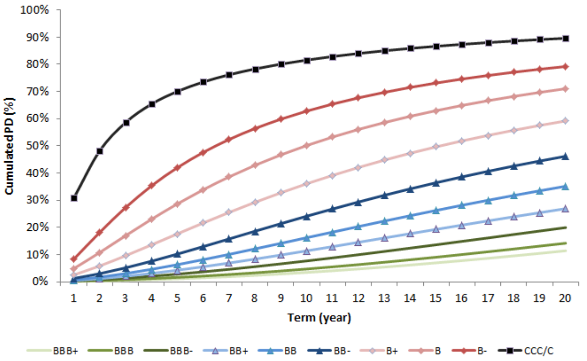


Figure 1: PD Term Structure Estimation for the ten Classes with the Discrete Homogeneous Model

Analyzing the results from practical corporate lending viewpoints, it can be argued that intuitively the PD term structure of the worse classes (B+ and down) might seem to be acceptable, since the trajectories follow a shape with decreasing progress, nevertheless, in particular in the second ten years the rate of growth appears to be exaggerated. However, in case of the better classes (BB- and up) the requirement of decreasing growth in time is not at all met, accordingly the results of discrete Markov chain model must be handled with doubts, since the estimated lifetime PD, as a consequence the expected loss, could be unduly high.

Continuous Homogeneous Model

For continuous Markov chain modeling it is essential to construct a generator matrix. It is easy to see that neither the simple root nor the logarithm of the annual transition matrix is appropriate, because the

Table 1: The applied Annual Transition Matrix

		To rating class										
From		1	2	3	4	5	6	7	8	9	10	Default
rating class		BBB+	BBB	BBB-	BB+	BB	BB-	B+	B	B-	CCC/C	D
1	BBB+	87.35%	9.14%	2.02%	0.43%	0.43%	0.11%	0.21%	0.11%	0.00%	0.11%	0.11%
2	BBB	8.04%	81.80%	6.77%	1.59%	0.74%	0.32%	0.32%	0.11%	0.00%	0.11%	0.21%
3	BBB-	1.38%	10.00%	77.86%	5.75%	2.45%	0.96%	0.43%	0.32%	0.21%	0.32%	0.32%
4	BB+	0.55%	2.18%	13.08%	70.56%	7.20%	3.27%	1.20%	0.76%	0.22%	0.55%	0.44%
5	BB	0.22%	0.67%	2.67%	10.91%	71.05%	8.69%	2.56%	1.34%	0.45%	0.78%	0.67%
6	BB-	0.11%	0.34%	0.46%	2.28%	10.81%	69.49%	9.68%	3.64%	1.02%	0.91%	1.25%
7	B+	0.12%	0.12%	0.12%	0.35%	1.85%	9.12%	70.91%	9.81%	3.00%	2.08%	2.54%
8	B	0.00%	0.11%	0.00%	0.23%	0.34%	1.58%	9.73%	68.87%	9.39%	4.87%	4.87%
9	B-	0.11%	0.11%	0.11%	0.11%	0.22%	0.56%	2.89%	12.24%	62.61%	12.69%	8.35%
10	CCC/C	0.12%	0.12%	0.12%	0.00%	0.23%	0.47%	1.40%	3.38%	10.62%	52.74%	30.81%

Source: calculations based on Fitch (2015) and S&P (2016) characteristics of generator matrix are not necessarily realized and negative results might arise. The empirical transition matrix might in itself possess such properties that exclude the existence of a generator matrix, and the same transition matrix might be resulted starting from more generator matrices (Israel et al. 2001).

Within the framework of this empirical research an approximated generator matrix was elaborated applying the regularization procedure published by Kreinin and Sidelnikova (2001) guaranteeing very good fit to the transition matrix considering Euclidean distance.

The first step of regularization is to take the natural logarithm of the annual transition matrix. Where negative values are resulted apart from the diagonal, they must be substituted with zero, so an initial G matrix is received. To achieve that the generator matrix contains zero sums of rows, non-positive diagonal values and non-negative non-diagonal values, the rows of the matrix must be modified considering the relative contribution of each element (Kreinin and Sidelnikova ibid.), formulating a $\widetilde{G}$ matrix, the elements of which are calculated as follows:

$$\tilde{g}_{ij}=\left|g_{ij}\right|\frac{\sum_{j=1}^{N} g_{ij}}{\sum_{j=1}^{N}\left|g_{ij}\right|} \qquad (6)$$

The difference of the two matrices gives $\widehat{G}$ generator matrix (Table 2), in which the sums of rows are zero:

$$\widehat{G}=G-\widetilde{G} \qquad (7)$$

In line with the assumption system of the continuous Markov chain, probabilities of transitions for – even fractional – terms can be estimated, by exponentiating the generator matrix to the desired power. Figure 2 summarizes the estimated PD term structure for twenty years.

It is visible from the PD trajectories that results are very similar to the discrete model, accordingly the drawn critical observations are also valid for the continuous homogeneous model. Hence, despite the fact that the exponentiation of the generator matrix almost perfectly estimates the annual transition matrix, the forward looking results are disappointing, since the model systematically overestimates the realistically expected default.

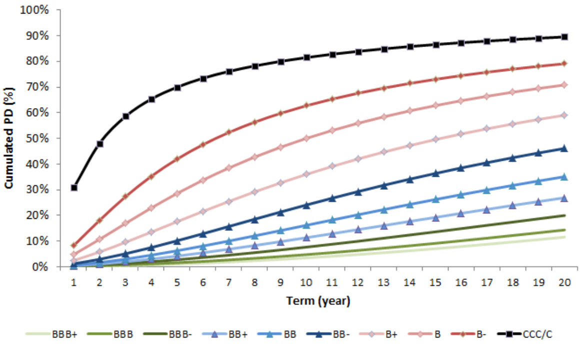


Figure 2: PD Term Structure Estimation for the ten Classes with the Continuous Homogeneous Model

Continuous Non-homogeneous Model

Perceived problems of the homogeneous models are expected to be resolved by giving up the homogeneity assumption. It ensures the flexibility that estimated PD term structure better reflects realistic default trajectories. Again the $\widehat{G}$ generator matrix (Table 2) is the starting point, however, it is no more assumed that the transitions are identical, and a timely dependent generator is applied:

$$\widehat{G}_t=\phi(t)\times\widehat{G} \qquad (8)$$

where × is matrix multiplication and $\phi(t)=(\varphi_{ij}(t))_{1\leq i,j\leq R}$ is such an R×R diagonal matrix, where:

$$\varphi_{ij}(t)=\begin{cases}0 & \text{if } i\neq j\\ \varphi_{\alpha,\beta}(t) & \text{if } i=j\end{cases} \qquad (9)$$

$\varphi_{\alpha,\beta}(t)$ can be formulated in the function of non-negative α and β parameters per rating class as follows (Bluhm and Overbeck 2007):

$$\varphi_{\alpha,\beta}(t)=\frac{(1-e^{-\alpha t})t^{\beta-1}}{1-e^{-\alpha}} \qquad (10)$$

Table 2: The applied Generator Matrix

	BBB+	BBB	BBB-	BB+	BB	BB-	B+	B	B-	CCC/C	D
BBB+	-14.05%	10.72%	1.97%	0.33%	0.44%	0.06%	0.22%	0.10%	0.00%	0.13%	0.07%
BBB	9.50%	-21.17%	8.27%	1.70%	0.72%	0.27%	0.34%	0.08%	0.00%	0.12%	0.19%
BBB-	1.07%	12.47%	-26.25%	7.46%	2.82%	0.93%	0.36%	0.30%	0.22%	0.39%	0.23%
BB+	0.47%	1.71%	17.57%	-36.38%	9.66%	3.93%	1.18%	0.78%	0.11%	0.71%	0.27%
BB	0.19%	0.52%	2.25%	15.31%	-35.93%	11.97%	2.61%	1.32%	0.34%	1.00%	0.43%
BB-	0.09%	0.35%	0.18%	2.06%	15.31%	-38.30%	13.38%	4.15%	0.87%	0.94%	0.98%
B+	0.13%	0.09%	0.08%	0.16%	1.60%	12.90%	-36.29%	13.55%	3.30%	2.44%	2.04%
B	0.00%	0.13%	0.00%	0.26%	0.19%	1.34%	13.71%	-39.71%	13.68%	6.42%	3.98%
B-	0.13%	0.11%	0.12%	0.11%	0.20%	0.40%	2.89%	18.28%	-50.12%	21.73%	6.15%
CCC/C	0.15%	0.14%	0.16%	0.00%	0.29%	0.54%	1.63%	3.81%	18.43%	-66.36%	41.20%
D	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

In case of t=1 the diagonal matrix purely consists of $\varphi_{\alpha,\beta}(1)=1$. In the numerator $(1\text{-}e^{-\alpha t})$ denotes the exponential distribution of the random variable, while $t^{\beta-1}$ serves for convexity or concavity adjustment. Hence both the flexibility of free parameter selection and the application of well known functions from probability theory are met. By proper selection of α and β parameters, the generator matrix can interpolated to empirically given cumulated PD rates, achieving satisfactory estimation accuracy.

To optimize α and β parameters the long-term actual global corporate cumulated default rates of Fitch and Standard&Poors have been considered, for horizons where data was available at both agencies. The first-year rates equal to the probabilities of transitions to the default class in the annual transition matrix.

Table 3: Cumulated PD Calibration Targets

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 10
BBB+	0.11%	0.34%	0.51%	0.80%	1.11%	2.31%
BBB	0.21%	0.58%	0.83%	1.31%	1.77%	3.78%
BBB-	0.32%	1.03%	1.63%	2.34%	3.11%	6.59%
BB+	0.44%	1.69%	3.07%	4.33%	5.45%	8.93%
BB	0.67%	2.42%	3.82%	5.49%	7.00%	12.30%
BB-	1.25%	3.01%	4.82%	6.45%	7.86%	13.18%
B+	2.54%	5.59%	7.83%	10.21%	11.89%	16.22%
B	4.87%	8.26%	10.41%	13.51%	16.12%	19.90%
B-	8.35%	12.55%	15.09%	18.91%	22.31%	25.40%
CCC/C	30.81%	35.64%	37.78%	40.35%	42.70%	45.29%

Source: calculations based on Fitch (ibid.); S&P (ibid.)

During optimization the monotonically increasing cumulated PDs, the accurate estimation of empirical default rates, and the realistic reflection of practical corporate lending experience also played important role. Table 4 summarizes the so optimized parameters.

Table 4: Optimal α and β Parameters for the Classes

	α	β
BBB+	0.40	1.85
BBB	0.08	0.32
BBB-	0.78	0.21
BB+	0.47	0.96
BB	0.35	0.92
BB-	0.09	0.87
B+	0.11	0.59
B	0.05	0.14
B-	0.93	1.43
CCC/C	0.80	0.01

In line with the assumption system of the continuous non-homogeneous Markov chain, probabilities of transitions for – even fractional – terms can be estimated, by exponentiating the timely changing generator matrix to the desired power. Figure 3 summarizes the estimated PD term structure for twenty years.

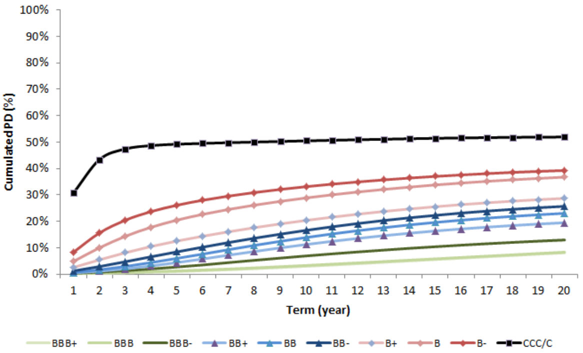


Figure 3: PD Term Structure Estimation for the ten Classes with the Continuous Non-homogeneous Model

The application of continuous non-homogeneus model results in the reduction of marginal default rates in the forecasting horizon, which is most observable in the CCC/C class. It is, however, not surprising, considering that initially this class has the highest PD, and the probability of remaining in the same class is the lowest in the transition matrix.

The non-homogeneous Markov chain intuitively and empirically gives suitable PD term structure in the light of actual default rates published by rating agencies, and also from the viewpoint of practical corporate lending experience. The estimated PD term structure mathematically and professionally properly expresses the PD element of expected loss that can realistically occur in the long-run in Hungarian corporate lending.

CONCLUSIONS

The paper attempted to provide forecast methodological framework and concrete models to estimate long-run PD term structure for Hungarian corporate debt instruments, in line with IFRS 9 requirements.

Lifetime PD modeling has fifty-five years of history in literature. PD term structure can be estimated by various methods. It was concluded that the most viable method to estimate long term PDs for Hungarian corporate debt instruments is the Markov chain approach.

In Markov chain modeling the first task is to construct an annual transition matrix, for which the normalized average of long run global corporate annual probabilities of transitions of Fitch and Standard&Poors rating agencies were considered, diminished to 10+1 rating classes, reflecting the credit risk characteristics of Hungarian corporate debtors.

On the basis of the transition matrix a discrete homogeneous, a continuous homogeneous and a continuous non-homogeneous Markov chain model were elaborated. PD term structures were estimated for twenty years.

Empirical results revealed that both discrete and continuous homogeneous models systematically overestimated the long term corporate PDs. However, the continuous, non-homogeneous Markov chain gave both intuitively and empirically appropriate PD term structure.

REFERENCE LIST

Banasik, J.; J.N. Crook and L.C. Thomas. 1999. “Not if but when will borrowers default”. *Journal of the Operational Research Society*, Vol. 50, No. 12, 1185-1190.

Black, F. and J.C. Cox. 1976. “Valuing corporate securities: Some effects of bond indenture provisions”. *Journal of Finance*, Vol. 31, No. 2, 351-367.

Black, F. and M. Scholes. 1973. “The pricing of options and corporate liabilities”. *Journal of Political Economy*, Vol. 81, No. 3, 81-98.

Bluhm, C. and L. Overbeck. 2007. “Calibration of PD term structures: to be Markov or not to be”. *Risk*, Vol. 20, No. 11, 98-103.

Cyert, R.; H. Davidson and G. Thompson. 1962. “Estimation of the allowance for doubtful accounts by Markov chains”. *Management Science*, Vol. 8, No. 3, 287-303.

Duffie, D. and K. Singleton. 1999. “Modeling term structures of defaultable bonds”. *Review of Financial Studies*, Vol. 12, No. 4, 687-720.

Fama, E.F. 1986. “Term premiums and default premiums in money markets”. *Journal of Financial Economics*, Vol. 17, No. 1, 175-196.

Fitch. 2015. *Fitch ratings global corporate finance 2014. Transition and default study.* Fitch Ratings. Available from Internet: www.fitchrating.com

Gavalas, D. and T. Syriopoulos. 2014. “Bank credit risk management and rating migration analysis on the business cycle”. *International Journal of Financial Studies*, Vol. 2, No. 1, 122-143.

IASB. 2014. *IFRS 9 financial instruments.* International Accounting Standards Board, London

Israel, R.B.; J.S. Rosenthal and J.Z. Wei. 2001. „Finding generators for Markov chains via empirical transition matrixes, with applications to credit ratings”. *Mathematical Finance*, Vol. 11, No. 2, 245-265.

Jarrow, R.A. 2001. “Default parameter estimation using market prices”. *Financial Analyst Journal*, Vol. 57, No. 5, 75-92.

Jarrow, R.A.; D. Lando and S. Turnbull. 1997. “A Markov model for the term structure of credit risk spreads”. *Review of Financial Studies*, Vol. 10, No. 2, 481-523.

Kreinin, A. and M. Sidelnikova. 2001. “Regularization algorithms for transition matrices”. *Algo Research Quarterly*, Vol. 4, No. 1-2, 23-40.

Kristóf, T. 2008. “A csődelőrejelzés és a nem fizetési valószínűség számításának módszertani kérdéseiről [On methodological questions of bankruptcy prediction and PD modeling]”. *Közgazdasági Szemle*, Vol. 55, No. 5. 441-461.

Kristóf, T. and M. Virág. 2012. “Data reduction and univariate splitting. Do they together provide better corporate bankruptcy prediction?”. *Acta Oeconomica*, Vol. 62, No. 2, 205-227.

Lando, D. and T.M. Skodeberg. 2002. “Analyzing rating transactions and rating drift with continuous observations”. *Journal of Banking & Finance*, Vol. 26, No. 2-3, 423-444.

Longstaff, F.; S. Mithal, and E. Neis. 2005. “Corporate yield spreads: default risk or liquidity? New evidence from the credit-default swap market”. *Journal of Finance*, Vol. 60, No. 5, 2213-2253.

Merton, R.C. 1974. “On the pricing of corporate debt: The risk structure of interest rates”. *Journal of Finance*, Vol. 29, No. 2, 449-470.

Nyitrai, T. 2015. “Hazai vállalkozások csődjének előrejelzése a csődeseményt megelőző egy, két, illetve három évvel korábbi pénzügyi beszámolók adatai alapján [Bankruptcy prediction of Hungarian enterprises using one, two and three years of historical annual report data]”. *Vezetéstudomány*, Vol. 46, No. 5, 55-65.

S&P. 2016. *Default, transition and recovery: 2015 annual global corporate default study and rating transitions.* Global fixed income research. Standard & Poors Financial Services. Available from Internet: www.standardandpoors.com/ratingdirect

Virág, M. and A. Fiáth. 2010. *Financial ratio analysis.* Aula Kiadó, Budapest

Virág, M.; T. Kristóf; A. Fiáth and J. Varsányi. 2013. *Pénzügyi elemzés, csődelőrejelzés, válságkezelés [Financial analysis, bankruptcy prediction, crisis management].* Kossuth Kiadó, Budapest

Virág, M. and T. Nyitrai. 2014. “Is there a trade-off between the predictive power and the interpretability of bankruptcy models? The case of the first Hungarian bankruptcy prediction model”. *Acta Oeconomica*, Vol. 64, No. 4, 419-440.

AUTHOR BIOGRAPHIES

TAMÁS KRISTÓF was graduated from Budapest University of Economic Sciences and Public Administration where he obtained his MSc degree in 2001. PhD since 2009. Senior lecturer at Corvinus University of Budapest, Strategic risk management director at MFB Hungarian Development Bank Plc., Member of Management Board at Garantiqa Hitelgarancia Plc., Public Body member at Hungarian Academy of Sciences.

His research fields encompass credit risk modeling, forecast methodology, bankruptcy prediction and futures studies.

His e-mail address is:
tamas.kristof@uni-corvinus.hu
His ResearchGate profile is:
https://www.researchgate.net/profile/Tamas_Kristof2
His Linkedin profile can be found at:
https://hu.linkedin.com/in/tamás-kristóf-82ab9555

MIKLÓS VIRÁG was graduated from Karl Marx University of Economic Sciences where he obtained his MSc degree in 1982. Dr. Univ since 1984, CSc since 1993, Dr. Habil since 2001. University Professor at Corvinus University of Budapest, Director of Business Development Institute, Member of Senate, Member of Business Administration Faculty Board, Member of the Economics Committee of Hungarian Rectors' Conference, Chairman of Supervisory Board at MVM Hungarian Electricity Plc., Public Body member at Hungarian Academy of Sciences.

His research fields encompass corporate finances, financial performance measurement, bankruptcy prediction, optimizing decision structures and financial rating of national economic branches.

His e-mail address is: miklos.virag@uni-corvinus.hu
His ResearchGate profile is:
https://www.researchgate.net/profile/Miklos_Virag
His Linkedin profile can be found at:
https://hu.linkedin.com/in/miklós-virág-79177917

THE USE OF ECONOMETRIC MODELS IN THE STUDY OF DEMOGRAPHIC POLICY MEASURES (BASED ON THE EXAMPLE OF FERTILITY STIMULATION IN RUSSIA)

Oksana Shubat
Anna Bagirova
Ural Federal University
620002, Ekaterinburg, Russia
Email: o.m.shubat@urfu.ru
Email: a.p.bagirova@urfu.ru

KEYWORDS
Econometric models, time series analysis, demographic policy, total fertility rate, maternity capital

ABSTRACT

Russia is experiencing steady population decline. One of the reasons for this is low fertility. The other major problem is insufficient housing availability. In today's political discussion, these two problems are often presented as interconnected. The aim of our research is to analyse the relationship between fertility dynamics and provision of housing in Russia in order to subsequently assess the effectiveness of the most expensive measure for stimulating fertility in the state's history – the so-called "maternity capital". We estimated regression models for the time series of fertility rates and the availability of housing. To assess the strength of relationship between the time series, we analysed correlation between regressions' residuals in two models.A retrospective analysis of the time series showed no correlation between the two in a historical context. Throughout the time that the maternity capital was in place the correlation analysis also revealed no relationship between them. Our analysis showed that these variables were not significantly correlated either in urban or rural Russian areas. We can conclude that the introduction of maternity capital in Russia was not underpinned by profound statistical and demographic analysis. Our results also give reason to question the effectiveness of maternity capital.

INTRODUCTION

Like most European countries, Russia is experiencing steady population decline. One of the reasons for this is low fertility. In 2015, the total fertility rate (TFR) was 1.78, which is 15.2% below the replacement fertility rate (Total Fertility Rate 2016). This is of concern to country's leadership, which is interested in economic growth and a strengthening of Russia's demographic potential.

The other major issue in our country, which has been around since Soviet times, is insufficient housing availability. For example, the average number of rooms shared per person in a dwelling in Russia is almost half that in Germany and France (Housing 2016).. In early 1986, future USSR president Mikhail Gorbachev promised that by the year 2000, every family would live in their own flat or house. The USSR adopted a state programme called "Housing-2000". However, the collapse of the Soviet Union derailed the implementation of the program. People only received the opportunity to become homeowners in 1991 (Federal law 1541-I 1991); until this time, flats were mostly distributed free of charge through a queue-based system, which could last decades.

In today's political discussion, these two problems are often presented as interconnected. At the same time, Russian demographers see insufficient housing as just one of many causes of low fertility (Rotova 2012). According to the theory put forward by V. Borisov, V. Arkhangelsky, A. Antonov et al., there are two sets of factors behind low birth rates: socio-psychological (in other words, a low desire for children) and socio-economic (or poor conditions for actualising the desire for children) (Arkhangelsky 2012). The first group includes factors like the desire to have a particular number of children, widespread social norms regarding the number of children and so on. The second set of factors includes income levels, living conditions, accessibility of kindergartens and the like. Notably, the influence of these groups of factors is interconnected and one cannot talk about fertility being determined solely by economic conditions without accounting for the desire for children. As such, an improvement in the population's economic conditions in and of itself will not lead to a growth in fertility.

Existing state measures for supporting fertility in Russia entail every type of assistance for families spelled out by O. Thevenon and A. Gauthier: assistance to pregnant women; assistance at childbirth; assistance aimed at providing parents with the opportunity to combine childcare and paid employment; payments to parents who look after children (Thevenon and Gauthier 2011). Moreover, in 2007 Russia introduced an unprecedented measure for stimulating fertility – the so-called "maternity capital". This entails a lump-sum payment after the birth of the second (or third and so on) child and can only be received once. The amount (around 7,750 EUR in 2017) can be spent on housing betterment, children's education or on the mother's

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

future pension. The most popular way to spend this payment is to improve living conditions, with some 95% of recipients spending the money on housing (Maternity capital in Krasnoyarsk 2016; Statistical data on the expenditure of maternity capital. 2016). Notably, there is great variability across Russia as regards overall standards of living, including the per square metre cost of housing. Moreover, urban housing is always more expensive than rural residences. Yet the maternity capital amount is in no way modified on the basis of where the mother and child live, and is the same across the country.

The impact of demographic policy measures on overall fertility and its individual determinants has been the subject of extensive research around the world. Balbo, Billari and Mills's work presents a wide spectrum of topical results, where all fertility determinants are grouped into three levels: micro (determinants at the individual and/or couple level); meso (social relationships and social networks) and macro (cultural and institutional settings) (Balbo, Billari and Mills 2013). In Russia, sociologists chiefly study the desire for children among different categories of women, whereas research into socio-economic conditions for actualising these desires is far less prevalent (Sinitca 2012). At the same time, as Sinitsa notes, "Russian research contains extensive recommendations as regards state policy, whereas international studies mostly describe existing processes" (Sinitca 2012: 106).

Undoubtedly, the introduction of a rather expensive mechanism for stimulating fertility should have been preceded by a profound analysis of the demographic and socio-economic situation across different Russian regions. As such, the aim of our research is to analyse the relationship between fertility dynamics and provision of housing in Russia in order to subsequently assess the effectiveness of the most expensive measure for stimulating fertility in the state's history.

We note that maternity capital has been used in Russia as a demographic policy tool for 10 years. The country's leadership regularly praises its effectiveness, but to date, there has been no fundamental scientific research to support these claims.

DATA AND METHODS

1. In the course of our research, we studied the time series for the following indicators:

– We used the average number of square metres of housing per resident to describe the level of housing availability. This is the most accessible and commonly occurring indicator on people's standards of living in Russian statistics. This data is publicly available from 1980 (Living standards data 2016).

– We used Total Fertility Rate to describe fertility, as it can provide an integrated representation of fertility intensity across different age groups. Russian statistics data for this is available over a much more extended period of time. However, for our analysis, we applied a comparable period and only used data from 1980 onwards (Total Fertility Rate data 2016).

2. To explore the correlation between the time series, we tried to to exclude spurious correlation. We estimated regression models for the stated time series and used ordinary least squares as the method for estimating the parameters of the models. In certain cases, we found autocorrelation of the residuals (AR(1)). Since we excluded an incorrect model specification, we removed autocorrelation by estimating such models on the basis of generalized least squares, Cochrane-Orcutt iterative procedure and Prais-Winsten correction. To assess the strength of relationship between the time series, we used Pearson and Spearman correlation. We analysed correlation between regression residuals in two models.

3. To test the hypothesis about the possible influence of maternity capital on growing fertility, we estimated models and tested the relationship between the studied variables separately for two periods of time: 1) for the entire period available for analysis; 2) for the period since the measure was introduced (i.e. between 2007 and now).

4. To test the hypothesis about possibly greater influence of maternity capital on fertility outside large cities, we estimated the correlation between the studied variables separately in urban and rural areas. We considered this an important aspect of the study, because we supposed that the effectiveness of this fertility incentive measure should be greater in rural areas – in parts of the country where housing costs less.

RESULTS

1. A retrospective analysis of the time series characterizing fertility levels in the country and the provision of housing to the population showed no correlation between the two in a historical context. Thus, since the 1980s, the average number of square metres of housing per person in Russia steadily grew (Figure 1).

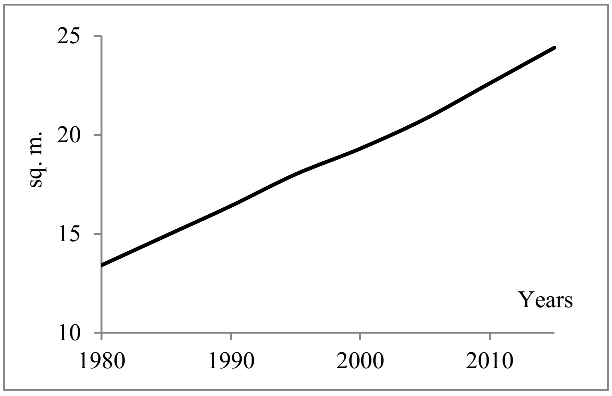


Figure 1: Average Number of Square Metres of Housing per Person in Russia (Living standards data 2016)

Yet over the long run, TFR moved in different directions. Thus, between the mid-1980s and the end of the century, this indicator was declining, only moving into a growth phase in 2000. Since then, TFR grew consistently (except 2005, when fertility fell) – a trend that is observed to date (Figure 2).

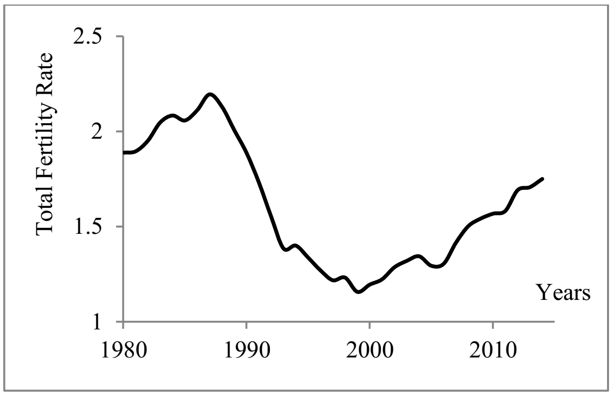


Figure 2: Total Fertility Rate in Russia (Total Fertility Rate data 2016)

Since the two trends became unidirectional in 2000, the subsequent modelling of the possible relationship between the two was done for this period.

2. Modelling trend in the availability of housing between 2000 and 2015 showed that this time series is well approximated by a linear trend. Year-on-year, housing availability grew by an average of 0.34 sq.m. per person in the country (tables 1-3).

Table 1: Model Summary
(dependent variable: model 1 – average number of sq.m. of housing per person; model 2 – TFR)

Model	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	0.995	0.995	0.1190	1.502
2	0.885	0.877	0.0358	1.565

Table 2: ANOVA
(dependent variable: model 1 – average number of sq.m. of housing per person; model 2 – TFR)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	39.41	1	39.41	2781.95	0.000
	Residual	0.19	14	0.01		
	Total	39.60	15			
2	Regression	0.14	1	0.14	107.58	0.000
	Residual	0.02	14	0.00		
	Total	0.15	15			

Table 3: Coefficients
(dependent variable: model 1 – average number of sq.m. of housing per person; model 2 – TFR)

Model		Unstandardized Coefficients		t	Sig.
		B	Std. Error		
1	Constant	-661.754	12.958	-51.071	0.000
	Years	0.340	0.006	52.744	0.000
2	Constant Constant corrected	-38.805 -79.931	3.812	-10.180	0.000
	Years	0.041	0.004	10.372	0.000

Upon modelling trend in the fertility between 2000 and 2015, we identified an autocorrelation of residuals in the initial model. An estimation on the basis of GLS, Cochrane-Orcutt iterative procedure and Prais-Winsten correction enabled us to identify an increasing linear trend in TFR dynamics with an average annual increase of 0.041 (tables 1-3).

3. Correlation analysis of regressions' residuals showed that fertility and housing availability in the examined period were not correlated (table 4).

Table 4: Correlations between TFR and housing availability (from 2000 to 2015)

Indicator	Value
Pearson Correlation	0.366
Sig. (2-tailed) Pearson Correlation	0.164
Spearman's rho	0.279
Sig. (2-tailed) Spearman's rho	0.295

4. Throughout the time that the maternity capital fertility stimulation programme was in place (2007 to 2015), the studied indicators grew. Yet the availability of housing in this period grew by 14%, whereas birth rates grew much more – by 25.5%.

5. Modelling the trends over the stated period and examining the correlations revealed no relationship between the availability of housing and fertility rates since the introduction of the maternity capital programme (table 5).

Table 5: Correlations between TFR and housing availability (from 2007 to 2015)

Indicator	Value
Pearson Correlation	-0.014
Sig. (2-tailed) Pearson Correlation	0.972
Spearman's rho	0.017
Sig. (2-tailed) Spearman's rho	0.966

An analysis with account of a possible lag in changes to fertility in response to improved housing conditions (with a delay of 1-2 years) did not reveal any connection between the evaluated variables either. Analogous analysis with account of a possible lag between improved housing conditions in response to greater fertility (as envisaged by the maternity capital programme) also did not reveal any correlation.

6. An analysis of the dynamics of the studied variables by territory showed that throughout the time that the maternity capital programme has been operating, rural indicators have been higher than urban ones. Thus, housing availability in rural areas each year was on average 7.6% greater than in cities; TFR in rural areas was on average 39.6% higher each year than in urban ones.

On the whole, between 2007 and 2015, housing availability grew both in cities and in non-urban areas. The difference in growth proved insignificant (13.7% and 14%). Yet the difference in fertility growth was marked (29.7% in cities and 17.4% in rural areas).

7. Modelling trends separately for cities and rural areas showed that the evaluated time series throughout the time of the maternity capital programme is well approximated by linear trends. Correlation analysis of regressions'residuals showed that birth rates and housing availability in the examined period were not significantly correlated either in urban or rural areas (table 6). Correlation analysis with lag effects also did not uncover any relationship.

Table 6: Correlations between TFR and housing availability by territory (from 2000 to 2015)

Indicator	Value for urban areas	Value for rural areas
Pearson Correlation	0.500	-0.069
Sig. (2-tailed) Pearson Correlation	0,170	0.859
Spearman's rho	0.383	-0.083
Sig. (2-tailed) Spearman's rho	0.308	0.831

DISCUSSION

The lack of correlation between the dynamics of birth rates and housing availability in Russia gives reason to question the effectiveness of maternity capital, the most expensive state-funded fertility stimulation programme in the country's history. Our results give rise to a number of discussion points.

First of all, we believe there are two main reasons for the registered growth of TFR in Russia in recent years:

1. The introduction of maternity capital could have had a certain effect on the total number of births per woman and the timing of childbearing. Thus, women who had unstable reproductive plans in 2007 may have been prompted to have a child after this programme was implemented (the so-called postponement of childbearing effect). Moreover, one could suppose that some people (particularly rural dwellers) misunderstood information about the rollout of maternity capital, believing that they could spend the money as soon as their second child was born, in whatever way they saw fit;

2 A certain positive effect could be the result of the relatively successful development of the Russian economy between 2007 and 2015. In this period, per capita GDP grew 1.45 times, according to IMF data (Gross domestic product 2016). Despite the observation made by, for example, Sobotka et al, that the relationship between TFR and GDP is contradictory (Sobotka, Skirbekk and Philipov 2011), there is research that clearly registers a positive correlation between these indicators. Thus, Martin identified this connection for Australia (Martin 2004), Santow and Brachner – for Sweden (Santow and Bracher 2001), and so on.

In our view, it is important to note that the growth in TFR happened against a background of a negative influence of structural factors on total births in Russia. Indeed, in this period, there was a significant decline in the proportion of women of fertile age in the total female population. Thus in 2002 it was 43.6%, compared to 39.4% in 2015. The number of women of this age declined by 2.9 million people during this period.

We would also note that despite the growth in TFR, its value during the studied period was below the replacement fertility rate (2.1). Moreover, the total effectiveness of maternity capital measure would have to lead TFR to increase by at least 1.0 (if one is to imagine that women of fertile age would, upon learning of this unprecedented support, choose to use it immediately). However, official Russian statistics for 2007-2008 registered a growth in TFR of just 0.086. This shows that in 2008, only every 12th woman gave birth to one child more than in the previous year (moreover, it is unclear whether that decision was the result of the introduction of maternity capital).

If one is to accept the hypothesis that maternity capital and the potential improvement in living conditions it promises influences the average number of births for a Russian woman, its impact should be greater for rural areas, where living standards and living costs are lower. Statistical data disprove this hypothesis. When it comes both absolute and relative values, the growth in TFR was greater for women from cities.

It is also clear that before the introduction of such a costly fertility stimulation programme, no retrospective analysis of relationships between fertility levels and housing availability was carried out. Our data show that there is no positive correlation between these indicators. This is supported by research by other Russian scientists. For example, Maleva and Sinyavskaya estimated the probability of childbirth depending on various socio-economic indicators between 2001 and 2004 (Maleva and Sinyavskaya 2007). They registered that for women who already have children, the decision to have another child was in no way connected to living conditions . These researchers note that "ceteris paribus, housing availability statistically significantly increases the likelihood of childbearing in the model of all completed births, and with respect to the model for first children most of all. However, it proves insignificant for the birth of the second and subsequent child" (Maleva and Sinyavskaya 2007: 183). As such, it can be asserted that from the outset, maternity capital was not able to influence the very category of women at whom it was aimed (women who have one child and are potentially ready to have another).

Moreover, we consider the content of this programme to be rather ill-thought through for reasons that include:

1) strong variability across Russia as regards standard of living with the same maternity capital amount for different regions. This leads to varying perceptions about the impact of this measure for different groups of the population. As such, the very mechanism was initially aimed at stimulating fertility in the most economically laggard parts of Russia;

2) such significant amounts (in 2016 alone, Russia spent RUB 304.3 billion (Federal law 364-FZ 2015) or EUR 4.77 billion on maternity capital) are allocated solely for the birth of a child, rather than for supporting his subsequent upbringing and development. Thus, we believe that people are presented with a certain "message" that the state cares most about the quantity and not the 'quality' of children. We note that there are many international examples of a more balanced demographic policy. For example, in France, child benefits are paid from the second child onwards, until the child turns 20 and the amount grows with the number of children and their age (Köppen 2006).

CONCLUSIONS

A study of cause-and-effect relationships between variables presented as time series is one of the difficulties of econometric modelling. This could explain the deficit of Russian demographic research that uses such instruments. At the same time, econometric modelling has high heuristic potential and could become the basis for developing more effective demographic policy measures. Firstly, such modelling enables justifying the expedience of a particular demographic policy measure (by studying the potential relationship between variables that are expected to be affected by its introduction). Secondly, the use of econometric methods allows adjusting the substance of the developed measures.

On the basis of the results we obtained, we can suppose that the introduction of maternity capital in Russia as a state measure for stimulating fertility was not underpinned by profound statistical and demographic analysis. Thus, we believe that there were insufficient grounds to tie the introduced demographic measures solely with changes to the provision of housing. Moreover, the differentiation of Russian regions was not taken into consideration, which fully negated the possible impact of introducing the maternity capital programme. For example, our earlier research showed that countries with a large number of constituent parts with high variance in their development require a demographic policy that is differentiated by type of region (Shubat et al 2016). Such development of targeted measures for different types of regions would improve the effectiveness of Russia's overall demographic policy. On the whole we note that such a strong emphasis on one economic measure for stimulating fertility without creating conditions to drive greater desire for children among the Russian population could not, from the outset, lead to higher fertility - the goal of demographic policy in contemporary Russia.

ACKNOWLEDGMENTS

The work was supported by Act 211 Government of the Russian Federation, contract № 02.A03.21.0006. The project was also sponsored by the Graduate School of Economics and Management of Ural Federal University.

REFERENCES

Arkhangelsky, V.N. 2012. Methodological issues in research into determinants of demographic processes. In: Zvereva, N.V. and Arkhangelsky, V.N. (eds.) *Determinants of demographic processes*: compendium. Moscow: Maks Press.

Balbo, N., Billari, F. C. and Mills, M. 2013. "Fertility in Advanced Societies: A Review of Research". *European Journal of Population / Revue Européenne De Démographie,* Vol 29, Issue 1, 1-38.

Federal law 1541-I from 04.07.1991 "On the privatization of housing in the Russian Federation"

Federal law 364-FZ from 14.12.2015 "On the budget of the Russian Federation Pension Fund for 2016"

Gross domestic product per capita, current prices. 2016. International Monetary Fund. URL: http://www.imf.org/external/pubs/ft/weo/2016/02/weodata/weorept.aspx?pr.x=64&pr.y=11&sy=1997&ey=2021&scsm=1&ssd=1&sort=country&ds=.&br=1&c=922&s=NGDPDPC&grp=0&a= (access date 21.01.2017)

Housing. OECD Better Life Index. 2016. URL: http://www.oecdbetterlifeindex.org/topics/housing/ (access date 21.01.2017)

Köppen, K. 2006. "Second Births in Western Germany and France". *Demographic Research*, Vol 14, 295-330. Retrieved January 21, 2017, from http://www.demographic-research.org/volumes/vol14/14/

Living standards data. 2016. The Federal State Statistics Service. Rosstat, Moscow. URL: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/population/level/ (access date 21.01.2017)

Maleva, T. and Sinyavskaya, O. 2007. Socio-economic factors of fertility in Russia: Empirical measurement and social policy challenges. In T. Maleva & O. Sinyavskaya (Eds.), *Parents and Children, Men and Women in Family and Society*. URL:

http://www.socpol.ru/publications/pdf/PiDMiG1_end.indd.pdf (access date 21.01.2017)

Martin, J. 2004. “The Ultimate Vote of Confidence’: Fertility Rates and Economic Conditions in Australia, 1976-2000”. *Australian Social Policy 2002-2003*. Canberra: Commonwealth of Australia.

Maternity capital in Krasnoyarsk and Krasnoyarsk Territory. 2016. URL: http://pro-materinskiy-kapital.ru/regionalniy/v-krasnoyarske-i-krasnoyarskom-krae/ (access date 21.01.2017)

Rotova, R. S. 2012. On the socio-economic determinants of fertility. In: Zvereva, N.V. and Arkhangelsky, V.N. (eds.) *Determinants of demographic processes*: compendium. Moscow: Maks Press.

Santow, G. and Bracher, M. 2001. “Deferment of the first birth and fluctuating Fertility in Sweden”. *European Journal of Population*, Vol 17, Issue 4, 343-363.

Shubat, O., Bagirova, A., Abilova, M. and Ivlev, A. 2016. “The Use of Cluster Analysis for Demographic Policy Development: Evidence From Russia”. *ECMS-2016: 30th European Conference on Modelling and Simulation*, 159-165.

Sinitca, A. L. 2012. Caring for pre-school children as a factor of reproductive behaviour: theoretical approaches. In: Zvereva, N.V. and Arkhangelsky, V.N. (eds.) *Determinants of demographic processes*: compendium. Moscow: Maks Press.

Sobotka, T., Skirbekk, V. and Philipov, D. 2011. “Economic Recession and Fertility in the Developed World”. *Population and Development Review*, Vol 37, Issue 2, 267-306.

Statistical data on the expenditure of maternity capital in Moscow and Moscow Region. 2016. The Pension Fund of the Russian Federation. URL: http://www.pfrf.ru/branches/moscow/news/~2015/05/25/92288 (access date 21.01.2017)

Thevenon, O. and Gauthier, A.H. 2011. “Family Policies in Developed Countries: a ‘Fertility-Booster’ with Side-Effects”. *Community, Work and Family*, Vol 14, Issue 2, 197-216.

Total Fertility Rate data. Single inter-departmental information and statistical system (SIDIS). 2016. Rosstat, Moscow. URL: https://fedstat.ru/indicator/55407 (access date 21.01.2017)

AUTHOR BIOGRAPHIES

OKSANA SHUBAT is an Associate Professor of Economics at Ural Federal University (Russia). She has received her PhD in Accounting and Statistics in 2009. Her research interests include demographic processes, demographic dynamics and its impact on human resources development and the development of human capital (especially at the household-level). Her email address is: o.m.shubat@urfu.ru and her Web-page can be found at http://urfu.ru/ru/about/personal-pages/O.M.Shubat/

ANNA BAGIROVA is a professor of economics and sociology at Ural Federal University (Russia). Her research interests include demographical processes and their determinants. She also explores issues of labour economics and sociology of labour. She is a doctoral supervisor and a member of International Sociological Association. Her email address is: a.p.bagirova@urfu.ru and her Web-page can be found at http://urfu.ru/ru/about/personal-pages/a.p.bagirova/

THE USE OF CLUSTER ANALYSIS TO ASSESS THE DEMOGRAPHIC POTENTIAL OF RUSSIAN REGIONS

Oksana Shubat
Anna Bagirova
Irina Shmarova
Ural Federal University
620002, Ekaterinburg, Russia
Email: o.m.shubat@urfu.ru
Email: a.p.bagirova@urfu.ru
E-mail: i.v.shmarova@urfu.ru

KEYWORDS
Cluster analysis, Demographic potential, Demographic policy, Russian regions.

ABSTRACT

In recent years, Russia has been grappling with a serious economic crisis. The slowing pace of economic development is accompanied by adverse demographic trends. The purpose of our study is to assess the demographic potential of Russian regions and identifying groups that require the implementation of specific measures aimed at its development. We used hierarchical cluster analysis to model Russia's demographic space and segment regions with comparable problems related to forming demographic potential. Clustering was based on the indicators describing the demographic potential at the macro-level (regional) and meso-level (family level). The analysis identified the groups of regions that have the best and the worst conditions for the development of demographic potential. We proposed a set of measures that would be most relevant to the needs of specific groups of regions and could directly drive the development and actualisation of demographic potential. The analysis showed the need to use multi-factor classification in the demographics of countries that have a high level of regional differentiation. Modelling the demographic space on the basis of cluster analysis can be seen as an element of the system of supporting administrative decision-making and the development of effective demographic policy.

INTRODUCTION

In recent years, Russia has been grappling with a serious economic crisis. This manifests through a depreciation of the national currency, a drop in people's real incomes, increased unemployment and so on. The slowing growth and even decline in Gross Domestic Product also points to the adverse economic situation. Such negative dynamics are particularly stark in the context of global trends (table 1).

Table 1: Real GDP growth (annual percent change) (Real GDP growth 2016)

	2012	2013	2014	2015	2016
Russia	3.5	1.3	0.7	-3.7	-0.8
World	3.5	3.3	3.4	3.2	3.1
	2017	2018	2019	2020	
	forecast				
Russia	1.1	1.2	1.5	1.5	
World	3.4	3.6	3.7	3.7	

The slowing pace of economic development in Russia is accompanied by adverse demographic trends. Thus, according to official forecasts, the number of people below working age will fall by 0.7 million people by 2030 (Federal State Statistics Service 2016). A drop in the Total Fertility Rate (hereinafter TFR) is also forecast.

It should be noted that there is a marked difference in the demographic situations across Russian regions. Thus, for example, in 2015 maximum TFR was 3.39 (in the Tyva Republic), while the minimum was 1.29 in Leningrad Region. The highest life expectancy was 80.05 years in Ingushetia, compared to 64.16 years in Chukotka (Single inter-departmental information and statistical system 2016). Such high regional differentiation indicates that while there should be an overall uniformity of approach, specific demographic policy measures need to be tailored to the needs of different regions.

Undoubtedly, it is not objectively possible to account for the full spectrum of diversity of regional situations in demographic policy. However, it is possible to cluster together regions with similar problems in this area. We believe that the ability to identify such types of regions is linked to the use of statistical cluster analysis, which will enable modeling Russia's demographic space.

Cluster analysis is often used for regional segmentation outside Russia. For example, Kronthaler identified groups of German regions based on their

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

economic potential (Kronthaler 2005). Laboutkova, Bednarova and Valentova used cluster analysis to study relationship between regional decentralisation and economic imbalances in Europe (Laboutkova, Bednarova and Valentova 2016). Simpach segmented municipalities in a part of the Czech Republic on the basis of demographic development (Simpach 2013). Mertlova and Prokop used a set of macroeconomic indicators for regional clustering (Mertlova and Prokop 2015). Vahalik and Stanichkova drew out groups of countries with analogous competitiveness characteristics (Vahalik and Stanichkova 2016). Koisova and Haviernikova segmented regions of Slovakia using socio-economic indicators (Koisova and Haviernikova 2016). Zhang and Li used cluster analysis to identify groups of Chinese provinces and examine qualitative characteristics of the respective populations (Zhang and Li 2014).

Russian scientists carry out research that draws on cluster analysis for the segmentation of regions based on certain factors, such as the level of development of human capital (Petrykina 2013), levels of business and demographic activity (Ilyshev and Shubat 2008), migration characteristics (Abylkalikov 2015) and so on. Unfortunately, the results of this research are not used as foundation for regional demographic policies. In our opinion, on the one hand, this is linked to a lack of relevant skills among the official bodies developing demographic policies, and on the other hand – to the country's leadership being unwilling to recognise the regional specifics and attain a more profound understanding of existing problems.

In the course of our study, we modelled the country's demographic space on the basis of the level of development of demographic potential. We believe that precisely the development of demographic potential, rather than growth in fertility (which is today the official target benchmark), should be the main objective of Russian demographic policy.

The very concept of demographic potential only emerged in science relatively recently. It is assessed through various indicators. Thus, for example, Goraj et al suggest use purely quantitative population measures (Goraj, Gwiazdzinska-Goraj and Cellmer 2016), while Dobrokhleb and Zvereva include life expectancy (an indicator of the quality of life) as one of the indicators of demographic potential (Dobrokhleb and Zvereva 2016). We believe that an assessment of regional demographic potential should include both quantitative and qualitative measures. This is linked to the fact that the demographic potential of a particular region and the conditions for its actualisation determine not only the quantity, but also the quality of the future population. The very conditions for the development of demographic potential are formed at different levels. The macro level comprises the set of objective conditions related to the economic and demographic conditions in the region. The meso level encompasses the immediate social environment, including family and significant social groups. The micro level is based on an individual's behavioural determinants. We suppose that the more balanced the conditions for the development of demographic potential at different levels, the better the conditions for its overall actualisation.

The purpose of our study is to assess the demographic potential of Russian regions and identifying groups that require the implementation of specific measures aimed at its development. On the basis of cluster analysis we propose a set of measures that would be most relevant to the needs of specific groups of regions and could directly drive the development and actualisation of demographic potential.

DATA AND METHODS

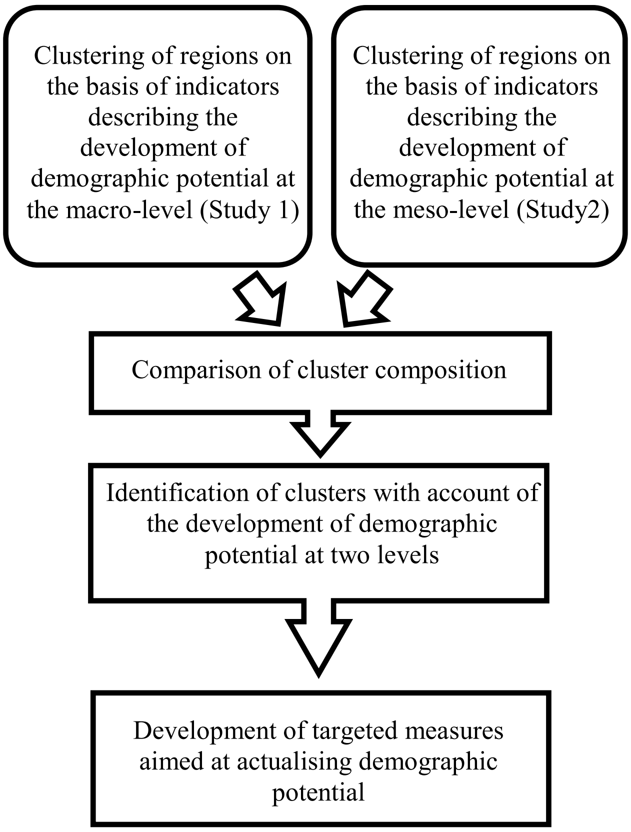
1. We used cluster analysis to model Russia's demographic space and segment Russian regions with comparable problems related to forming demographic potential. In carrying out this analysis, we undertook activities typical for this type of statistical work: the selection and transformation of input variables; the selection of distance measures and linkage rules; the selection of the clustering method; the selection of the number of clusters; the profiling of clusters and interpretation of the attained results.

2. We used hierarchical cluster analysis in our research. We used Euclidean distance as the distance measure and Ward's method to gauge distance between clusters. This decision was made on account of the analytical (discriminatory) abilities of these measures and their effectiveness (supported to the results of multiple studies). Moreover, these measures enabled making the clearest distinctions within the studied body of data, separating out uniform segments. To assess the robustness of the cluster solution, we performed several iterations of the clustering procedure, using different measures of distance between objects. Moreover, we applied partitioning methods of clustering (k-means procedure). For most Russian regions, their allocation into homogenous groups coincided. Some differences in cluster composition did not skew the profile of each identified group of regions. The shared characteristics and relationships identified in the course of the analysis did not change when different distance measures were used.

The decision on the number of identified groups of regions was taken on the basis of:

- Graphical representation of the clustering process (we examined a dendrogram);
- An evaluation of the between-group and within-group variability;
- Cluster size (we tracked the number of regions that form a single cluster to ensure that each group contained a sufficient number of regions).

3. The stages of our research are presented in Figure 1. We called this algorithm targeted clustering.


Clustering of regions on the basis of indicators describing the development of demographic potential at the macro-level (Study 1)
Clustering of regions on the basis of indicators describing the development of demographic potential at the meso-level (Study2)
Comparison of cluster composition
Identification of clusters with account of the development of demographic potential at two levels
Development of targeted measures aimed at actualising demographic potential

Figures 1: Research stages

The clustering of Russian regions on the basis of indicators describing the conditions for the formation of demographic potential at the macro-level – that is, the regional level (Study 1) – was based on the following variables:

- Birth rate;
- Perinatal mortality rate;
- Infant mortality rate;
- Under-five mortality rate;
- Pregnancy rate.

We presented the results of this clustering previously (Shubat et al 2016).

5. The clustering of Russian regions on the basis of indicators describing the conditions for the formation of demographic potential at the meso-level – that is, at the family level (Study 2) – was based on the following variables:

- Number of single mothers with children under the age of 18 (per 1,000 people);
- Number of single fathers with children under the age of 18 (per 1,000 people);
- Number of extramarital births (per 1,000 people);
- Coefficient of marriage instability (number of divorces per 1,000 marriages);
- Number of children without parental care (per 1,000 people).

6. We used both the clustering variables and other variables that describe the demographic situation in a region to interpret the clusters themselves. For this purpose we examined cluster centroids and executed tests of significance in difference between two means (or medians). We applied one-way analysis of variance (ANOVA) to evaluate differences between the means and Levene's test to assess the homogeneity of variances. To test the assumption of normality, we used the Shapiro-Wilks test. If we observed a violation of the assumptions of the one-way ANOVA, we used the non-parametric Kruskal Wallis Test.

7. Our samples included all Russian regions that had complete data for all input variables. Thus, Study 1 included 78 regions and Study 2 had 77 regions.

8. To carry out our research, we used data from our current demographic data, as well as data from the 2010 Census. The need to use data from different periods relates to the fact that not all necessary information is collected in Russia within the same period. We used open-source data provided by Russia's Federal Statistics Service (Federal State Statistics Service 2016).

RESULTS

1. In the course of our study, we identified a high degree of regional differentiation for all variables in both Study 1 and Study 2. There is a manifold difference in the maximum and minimum values of the variables that were used for the cluster analysis. Thus, the maximum value for "Marriage stability" is 5.5 times its minimum value. This ratio is even greater for "Number of children without parental care" - 17.5 times. The identified heterogeneity is evidently reason to use clustering methods for Russian regions.

2. The use of hierarchical cluster analysis for the data in Study 1 enabled identifying 3 clusters of regions:

- Cluster 1 – "Low fertility amid low economic activity";
- Cluster 2 – "Cautious" fertility amid high economic activity";
- Cluster 3 – "High fertility amid economic passivity".

We have previously demonstrated the profiling of these clusters and the assessment of the statistical significance of the obtained model of Russia's demographic space (Shubat et al 2016).

3. The application of hierarchical cluster analysis with respect to the data in Study 2 also enabled us to identify 3 regional clusters. The first cluster included 47 Russian regions; there were 10 regions in the second and 20 in the third. The evaluation of the cluster centroid confirmed the appropriateness of these three groups: the median values of the cluster variables differed significantly between the identified clusters (table 2). Results of the Kruskal Wallis Test are presented in Table 3. The dendrogram illustrates the arrangement of the clusters (Figure 2).

Table 2: Median values for cluster variables (Study 2)

Clustering variables	Cluster 1	Cluster 2	Cluster 3
Number of single mothers with children under the age of 18 (per 1,000 people)	38.4	31.3	34.9
Number of single fathers with children under the age of 18 (per 1,000 people)	3.5	5.3	3.5
Number of extramarital births (per 1,000 people)	4.0	2.5	2.7
Coefficient of marriage instability (number of divorces per 1,000 marriages)	575.6	516.0	558.8
Number of children without parental care (per 1,000 people)	0.8	0.3	0.4

Table 3: Kruskal Wallis Test (Study 2)

Clustering variables	Cluster	Mean Rank	Test Statisics		
			Chi-Square	df	Asymp.Sig
Number of single mothers with children under the age of 18 (per 1,000 people)	1	61.80	28280	2	0.000
	2	30.36			
	3	34.00			
Number of single fathers with children under the age of 18 (per 1,000 people)	1	35.40	16501	2	0.000
	2	65.80			
	3	34.83			
Number of extramarital births (per 1,000 people)	1	62.95	31154	2	0.000
	2	30.00			
	3	34.40			
Coefficient of marriage instability (number of divorces per 1,000 marriages)	1	49.35	8540	2	0.000
	2	24.70			
	3	37.64			
Number of children without parental care (per 1,000 people)	1	62.50	30455	2	0.000
	2	25.60			
	3	31.85			

4. The profiling of the groups of regions showed that they could be identified as clusters with different conditions for the development of demographic potential at the meso-level (family level).

Cluster 1 – "Worst conditions for the development of demographic potential at the meso-level"

This cluster includes around 20 Russian regions. The situation with the development of demographic potential here can be described as extremely poor. Indeed, this cluster has the poorest indicators as regards dissolved marriages, extramarital births, children left without parental care and single mother homes. Additional profiling through variables that were not used in the clustering revealed that this group of regions has the highest abortion indicators. Moreover, the figures for married couples without children are also high here.

Cluster 2 – "Best conditions for the development of demographic potential at the meso-level"

This cluster includes 10 Russian regions. The situation with the development of demographic potential can be described as most favourable. This group of regions had the most stable marriages, the lowest number of extramarital births and the least number of children without parental care. This cluster also had the lowest number of incomplete and single-mother families. Additional profiling through variables not used for clustering showed that compared to the rest of Russia, this group of regions has an above-average number of families with three or more children. One aspect of this group that is dissonant with the cluster's overall profile is the number of single fathers. On this, cluster 2 has the highest value out of all clusters.

Cluster 3 – "Satisfactory conditions for the development of demographic potential at the meso-level"

This cluster includes 20 Russian regions. It shows mid-way values across the analysed variables.

5. A comparison of the clusters obtained in the course of Study 1 and Study 2 allowed us to identify the groups of regions that have the best and the worst conditions for the development of demographic potential both at the macro- and the meso-levels.

We believe that it is possible to identify the most depressed groups of regions by finding overlapping entities in the following clusters: cluster 1 in Study 1 (Low fertility amid low economic activity) and cluster 1 in Study 2 (Worst conditions for the development of demographic potential at the meso-level). By comparing these two clusters, we identified a specific group that includes 11 Russian regions, that has the worst conditions for having and raising children at two levels. We labelled this group of regions the "depressed cluster", since the adverse processes happening at the family level coincide with a poor economic background (low levels of economic activity). This most problematic cluster evidently requires specific measures of demographic regulation.

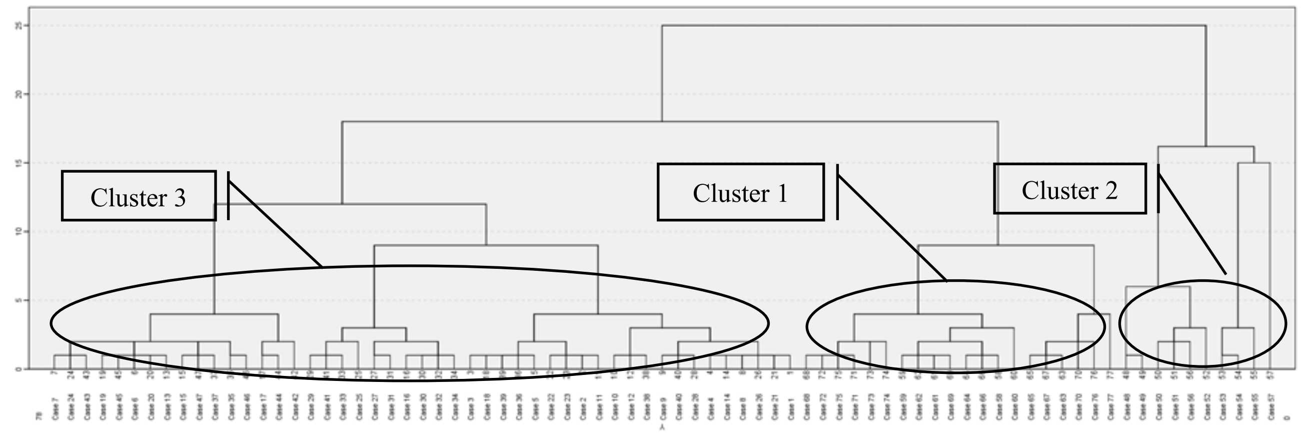


Figure 2: Dendrogram using Ward Linkage (Study 2)

We propose identifying the groups of regions with the most optimal conditions for the development of demographic potential by finding overlapping entities in the following clusters: cluster 3 in Study 1 (High fertility amid economic passivity) and cluster 2 in Study 2 (Best conditions for the development of demographic potential at the meso-level). Unfortunately, only three Russian regions fell within this intersecting zone. We labelled this group the "promising cluster".

DISCUSSIONS AND CONCLUSIONS

In our opinion that amid the adverse economic trends transpiring in Russia, the targeted development of measures for individual types of regions will help optimise budget spending to address such significant demographic policy challenges. The results of our analysis can be used to define priority aims and possible solutions that will reflect the unique needs of each identified regional cluster.

In our view, regions in the depressed cluster would benefit from the following set of measures (table 4).

The measures we propose place a strong emphasis on informational policies and social advertising for the development of demographic potential. We believe it should be segmented by different groups of young people.

Thus, university students can be targeted with social advertising through: 1) video clips on screens in university corridors; 2) advertising in student newspapers (essays, photographs, comic strips); 3) flashmobs; 4) advertising in social networks. The text of advertising materials should be profound and memorable, and convey the value of having and raising children. For example:

- "If I am born, I will bring you joy" (the text is accompanied by a photo of a baby);
- "Family is the meaning of life. Add it to your plans";
- "Plan your future. Plan a family";
- "A child is your future. Do it right"
- An Instagram photo competition around the topic of "This is the family I want".

Table 4: Overall aims and measures for the depressed cluster

Aim	Measures
Reduced perinatal and infant mortality	Improvements in the organisation of medical assistance to pregnant women and newborns, implementation of a "Health of future mothers and babies" programme
Increased pregnancy and birth rates	Informational policies aimed at strengthening young people's reproductive intentions, increasing the desire to have children, establishing a positive image of parenthood (through social networks, billboard advertising, mass media)
Greater marriage stability	Developing psychological support services for families (particularly for young families)
Fewer extramarital births and single-mother households	Informing young people about the legal consequences of unregistered marriages and psychological impact of raising children in incomplete families

Informational measures in regions in the depressed cluster should be supported by organisational and financial measures, aimed at enlarging the pool of state-funded education for children. Given the economic passivity of the region's population, we believe there is a need for a large number of extra-curricular children's activities funded by the local authorities.

Despite the seeming simplicity and obviousness of the described measures, they are hardly implemented in today's Russia. The results of our analysis suggest that the depressed group of regions could be a good 'pilot cluster' for such measures.

We would recommend that the relatively well-off three regions of the 'promising cluster' introduce a reproduction-focused component into educational curricula. At the moment, these regions are experiencing favourable macro- and meso-conditions for the development of demographic potential. As such, there is scope to set strategic aims for these regions, which will not only address quantity objectives, but also seek to improve the quality of the population's future human capital. University curricula can include special modules in disciplines related to parental education, as regards its medical, legal, economic and psychological aspects. A university graduate that has undergone such training will not only be ready to fulfil professional duties, but also to carry out the functions of a parent. They will be more mindful of the parenting process and have a greater understanding of the upsides and challenges it brings. In our view, such serious preparation for parenthood will enable young people – future parents in these parts of the country – to form higher quality human capital in their children.

On the whole, the analysis we carried out showed the need to use multi-factor classification in the demographics of countries that have a high level of regional differentiation. Modelling the demographic space on the basis of cluster analysis can be seen as an element of the system of supporting administrative decision-making and the development of effective demographic policy at the regional and national levels. The results of cluster analysis can be used as a basis for forecasting demographic risks and threats, identifying points of demographic growth and recession zones.

We see scope to further our research by developing a methodology for analysing the conditions for forming demographic potential at the personal (micro-) level. This study did not take this factor into account given its highly subjective nature and also the difficulties in formalising the information required for such analysis. Moreover, a separate study requires research into indicators for single fathers: its variance across Russian regions and the determinants for this rather rare phenomenon for Russia.

ACKNOWLEDGMENT

The work was supported by Act 211 Government of the Russian Federation, contract № 02.A03.21.0006.

REFERENCES

Abylkalikov, S.I. 2015. "Typological analysis of Russian regions by migration characteristics", *Regional economics: theory and practice,* Vol 22(397), 21-30.

Dobrokhleb, V.G. and Zvereva, N.V. 2016. "The Potential of Modern Russian Generations", *Economic and Social Changes-Facts Trends Forecast,* Vol 44, Issue 2, 61-78.

Federal State Statistics Service. 2016. URL: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/en/main/ (access date 01.02.2017)

Goraj, S., Gwiazdzinska-Goraj, M. and Cellmer, A. 2016. "Demographic Potential and Living Conditions in Rural Areas of North-Eastern Poland". *MSED-2016: 10th International Days of Statistics and Economics,* 482-493.

Ilyshev, A.M. and Shubat, O.M. 2008. "Multi-factor statistical analysis of business activity in regional micro-business". *Statistical issues,* Vol 4, 42-51.

Koisova, E. and Haviernikova, K. 2016. "Evaluation of selected regional development indicators by means of cluster analysis". *Actual Problems of Economics*, Vol 184, Issue 10, 434-443.

Kronthaler, F. 2005. "Economic Capability of East German Regions: Results of a Cluster Analysis". *Regional Studies*, Vol 39, Issue 6, 739-750.

Laboutková, S., Bednářová, P. and Valentová, V. 2016. "Economic Inequalities and the Level of Decentralization in European Countries: Cluster Analysis". *Comparative Economic Research*, Vol 19, Issue 4, 27-46.

Mertlova, L. and Prokop, M. 2015. "Cluster analysis as a method of regional analysis". *18th International Colloquium on Regional Sciences*, 56-63.

Petrykina, I. N. 2013. "Cluster analysis of regions of the Central federal district on the basis of the level of human capital development". *VSU gazette. "Economics and administration" series,* Vol 1, 72-80. URL: http://www.vestnik.vsu.ru/pdf/econ/2013/01/2013-01-11.pdf (access date 01.02.2017)

Real GDP growth. Annual percent. 2016. International Monetary Fund. URL: changehttp://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WEOWORLD/RUS (access date 01.02.2017)

Shubat, O., Bagirova, A., Abilova, M. and Ivlev, A. 2016. "The Use of Cluster Analysis for Demographic Policy Development: Evidence From Russia". *ECMS-2016: 30th European Conference on Modelling and Simulation*, 159-165.

Single inter-departmental information and statistical system (SIDIS). 2016. Rosstat, Moscow. URL: https://fedstat.ru/ (access date 01.02.2017)

Simpach, O. 2013. "Application of Cluster Analysis on the Demographic Development of Municipalities in the Districts of Liberecky Region". *MSED-2013: 7th International Days of Statistics and Economics,* 1390-1399.

Vahalík, B. and Staníčková, M. 2016. "Key factors of foreign trade competitiveness: Comparison of the EU and BRICS by factor and cluster analysis". *Society and Economy*, Vol 38, Issue 3, 295-317.

Zhang, X. and Li, Z. 2014. "Application of cluster analysis to western China population quality assessment". *EEE-2014: International Conference on E-Commerce, E-Business and E-Service*, 239-242.

AUTHOR BIOGRAPHIES

OKSANA SHUBAT is an Associate Professor of Economics at Ural Federal University (Russia). She has received her PhD in Accounting and Statistics in 2009. Her research interests include demographic processes, demographic dynamics and its impact on human resources development and the development of human capital (especially at the household-level). Her email address is: o.m.shubat@urfu.ru and her Web-page can be found at http://urfu.ru/ru/about/personalpages/ O.M.Shubat/

ANNA BAGIROVA is a professor of economics and sociology at Ural Federal University (Russia). Her research interests include demographical processes and their determinants. She also explores issues of labour economics and sociology of labour. She is a doctoral supervisor and a member of International Sociological Association. Her email address is: a.p.bagirova@urfu.ru and her Web-page can be found at http://urfu.ru/ru/about/personal-pages/a.p.bagirova/

IRINA SHMAROVA graduated from the Ural Federal University (Russia) in 2008. Now she is a senior lecturer of economics at Ural Federal University (Russia). Her research interests include demographic processes, the study of demographic policy implementation in the Russian regions and its impact on the state of labor resources. Her email address is: i.v.shmarova@urfu.ru and her Web-page can be found at http://urfu.ru/ru/about/personal-pages/i.v.shmarova/

BLIND VS. EMBEDDED INDIRECT RECIPROCITY AND THE EVOLUTION OF COOPERATION

Simone Righi
Department of Agricultural and Food Sciences – University of Bologna
I-40038, Bologna, Italia
Email: s.righi@unibo.it
and
"Lendület" Research Center for Education and Network Studies
Hungarian Academy of Sciences, Centre for Social Sciences
H-1014, Budapest, Hungary

Karoly Takacs
"Lendület" Research Center for Education and Network Studies
Hungarian Academy of Sciences, Centre for Social Sciences
H-1014, Budapest, Hungary
Email: takacs.karoly@tk.mta.hu

KEYWORDS

Agent-based modelling; Evolution of Cooperation; Indirect reciprocity; Forgiveness; Speed of evolution

ABSTRACT

The evolution of cooperation is one of the fundamental problems of both social sciences and biology. It is difficult to explain how a large extent of cooperation could evolve if individual free riding always provides higher benefits and chances of survival. In absence of direct reciprocation, it has been suggested that indirect reciprocity could potentially solve the problem of large scale cooperation. In this paper, we compare the chances of two forms of indirect reciprocity with each other: a blind one that rewards any partner who did good to previous partners, and an embedded one that conditions cooperation on good acts towards common acquaintances. We show that these two versions of indirect reciprocal strategies are not very different from each other in their efficiency. We also demonstrate that their success very much relies on the speed of evolution: their chances for survival are only present if evolutionary updates are not frequent. Robustness tests are provided for various forms of biases.

INTRODUCTION

The evolution of cooperative behaviour within a community of individuals is a largely studied problem (Hoffmann 2000; Sachs et al. 2004). The presence of kinship relationships can explain cooperative behaviour as the willingness to see related individuals - sharing a part of the genetic code - thrive and reproduce (Hamilton 1964; Grafen 1984). More difficult to justify is the cooperation among unrelated individuals.

Direct reciprocity is known to be an powerful mechanism for the evolution of cooperation Axelrod and Hamilton (1981); Axelrod (1984). Direct reciprocity, however, is difficult to justify as the only mechanism behind high levels of cooperation in human societies, as its efficiency strongly relies on direct knowledge and perfect memory of past behaviour of interacting partners. Previous research has uncovered various mechanisms that can contribute to the establishment of cooperation. Among others, these mechanisms include constrained interaction on networks or spatial structures (Hauert and Doebeli 2004; Lieberman et al. 2005), the competition among communities (Gunnthorsdottir and Rapoport 2006; West et al. 2007; Puurtinen and Mappes 2009), the interaction within small populations (Dunbar 1992), the presence of negative relationships among agents on the side of positive ones (Righi and Takács 2014) and more in general the presence of information about peers coming through mechanisms such as social comparison (Whitaker et al. 2016) reputation (Sigmund 2012), image-scoring (Wedekind and Milinski 2000), gossip (Sommerfeld et al. 2007; Wu et al. 2015), and language (Smith 2010).

Most of these mechanisms are related to the idea of indirect reciprocity: help - or retaliation - does not come from the interacting partners but rather from some third individual (Boyd and Richerson 1989; Nowak and Sigmund 1998, 2005). When interactions are constrained on a network structure (Hauert and Doebeli 2004), there can be at least two different operationalizations of the concept of indirect reciprocity, depending on the assumptions made about the informational flow accessible to agents.

A first type of indirect reciprocity is one where an

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

individual can observe the behaviour of the interacting partners with any third individual in the population related to the latter. In this case it is assumed that information can freely flow on the network of interactions so that a joint connection between the indirectly reciprocal agents is not necessary for the information about the behaviour of common peers to pass from the former to the latter. Within this paper we call this blind type of operationalization of the indirect reciprocity concept "unconnected reciprocity" (or UR).

A second type of indirect reciprocity is only concerned with the flow of information coming from peers. Within this paper we call this embedded type of operationalization of the indirect reciprocity concept "connected reciprocity" (or CR).

As this two types of indirect reciprocity rely in different information sets, their strength in sustaining cooperation can differ according to the setup studied. The objective and the innovation of this study is to characterize and compare the effectiveness of these two types of indirect reciprocity strategies. We analyze the chances of the two types of indirect reciprocity to support the evolution of cooperation under different conditions, in particular under network dynamics, in relation to the extent of forgiveness, and to the speed of strategy evolution.

THE MODEL

We consider a model with N agents placed on a non-weighted and non-directed Erdös-Rényi graph with a given density d. Every agent $i \in N$ is characterized by a strategy type and by a set of connections with a subset of the whole population $F_i^t \subset N$. Each time step, the two-person single-shot Prisoner's Dilemma (PD) is played once by each individual with each of her network connections. The agents play the PD, characterized by the classical payoff structure $Temptation(T) > Reward(R) > Punishment(P) > Sucker(S)$, with each of their neighbors. Specifically, for the sake of the simulations hereby discussed we fix the value of the payoffs to $T = 5$, $R = 3$, $P = 1$ and $S = 0$ as in Axelrod (1984). We consider the following four strategies:

- Unconditional Defection (UD): this type of agent always defects regardless of the behaviour or characteristics of the interaction partners.
- Unconditional Cooperation (UC): this type of agent always cooperates regardless of the behaviour or characteristics of the interaction partners.
- Unconnected Reciprocity (UR): this type of agent reciprocates the last action of the interacting partner with a randomly selected connection of the latter (i.e. some $z \in F_j^t$). [1]
- Connected Reciprocity (CR): this type of agent reciprocates the last action of the interacting partner with a randomly selected common connection (i.e. some $z \in F_i^t \wedge F_j^t$). [2]

During every dyadic interaction, agents can observe the past behaviour of the partner with all his connections and play according to their type. One key difference between the CR and the UR strategy is that the former can only get information to condition its behaviour from closed triads, while the latter can also condition its choice on the behaviour of individuals unconnected with him.

Time is divided in discrete periods and simulations run until a stable equilibrium is reached or 10000 periods have passed. At each time step t, each agent i contemporaneously plays the PD with all agents his first order social neighborhood, i.e. with each $j \in F_i$ observing all actions performed by others in $t-1$. Agents of type UR and CR, when observing their partners' actions may observe defection; in this case –with probability P_{for} – they may decide to give the partner another opportunity and observe another of his actions. This simulates the fact that individuals may forgive a defection act. At the end of the interaction phase, each individual computes his average payoff from all his interactions and compares it with that of peers in the direct neighborhood (i.e, with all those he has played with). With probability P_{evo} the individual changes its strategy into the one of the best performing partner.[3] Hence, the evolutionary strategy we apply is the "copy the best" update rule.

Network dynamics. Following hints of previous research (Santos et al. 2006), network dynamics may provide favorable conditions for the evolution of cooperation. For this reason it is important to observe its impact on the effectiveness of indirect reciprocity strategies on cooperation. In this model, each of ties of each individuals F_i^t can be subject to rewiring, with some probability at each time step. We consider three mechanisms of network update.

- With probability P_{ntw} a selected connection is rewired to another agent selected uniformly at random from the population.
- With probability P_{open} a selected connection is rewired so to open a closed triad incident on that particular edge.
- With probability P_{close} the rewiring is done by closing a triad that was open.

The three probabilities are extracted independently – for each link – at each time step. For this reason the sign of the difference $P_{open} - P_{close}$ indicates the propensity of the network clustering to increase (negative signs of the difference) or decrease (positive signs of the difference) over time.

Outcome Measures. We explore our model running a large number of simulations (1000 for the simulations proposed here) and by studying the statistical relationship between the outcome variables and the various parameters. This type of analysis is chosen to explore a complex model with many different parameters extensively. As outcome variables, the number of CR and the number of UR strategies are used alongside

[1] If there is no previous action to observe than the UR acts randomly, with 50% chance of cooperation.

[2] If there is no previous action to observe than the CR acts randomly, with 50% chance of cooperation.

[3] In case of ties, one of the partners with the highest payoff is selected randomly

with the proportion of cooperative acts. Populations of 200 agents are studied for a large number of time periods (10,000) or until the whole population of agents becomes of the same type so that no further evolution is possible. We define this latter moment t_{conv}. Since the model characteristics do not guarantee the convergence of the population to a single strategy, to describe evolutionary success, we define a strategy to gain

- **Absolute dominance**: if each individual in the population ends up playing the strategy or if – when $t_{max} = 10000$ is reached – this strategy has been adopted by more than 90% of the individuals.
- **Relative dominance**: if $t_{max} = 10,000$ is reached and this strategy is adopted by the relative majority of individuals.

Obviously, the second mode of dominance should be computed only for those cases where no convergence to a single strategy has been achieved. Most of our results show a quasi-certain convergence on one type of strategy. Therefore, except when otherwise noted, we will refer to the concept of absolute dominance in the discussion of our results.

In addition, the proportion of cooperative acts is computed at t_{conv} or t_{max} (in case of non-convergence).

RESULTS

Indirect reciprocity strategies and cooperation. We start by running 1000 simulations considering a population that is initially equally divided between the four types of strategies: UC, UD, CR and UR. $N = 200$ agents are laid on a random network with density $d = 0.2$. [4] For each single simulation the values of P_{evo} (the strategy evolution probability), P_{ntw} (probability of random network change), P_{for} (probability of forgiveness), P_{open} (probability to open each triad at each step) and P_{close} (probability to close each triad at each step) are extracted uniformly and independently from uniform distributions between 0 and 1.

Figure 1 shows that cooperation rarely disappears completely. On the opposite, for the majority of the simulations, the final proportion of cooperators is above zero. This implies that, for the largest part of possible parameter combinations, unconditional defection is unable to completely eliminate other types of strategies. The proportion of simulations with a share of cooperating acts larger than 1/2, however, is smaller, indicating that cooperation seldom becomes the action adopted by the majority of agents. It is interesting to note the sharp drop in the frequency of cooperation around a proportion 1/2, indicating that while the strategies of indirect reciprocity survive, cooperation seldom gains relative dominance.

Table I reports that UD was dominant in 83% of the simulations in which a strategy gained absolute dominance. As expected, UC never gained absolute dominance. More remarkable is the significant presence of instances in which CR and UR became dominant in the simulations that resulted in a single winner strategy. Indeed, about 10% of these runs ended up with the dominance of the CR strategy, while 7% of them terminated with the win of the UR strategy. Since the parameter space explored is quite large and contains parameter combinations very hostile to the emergence of cooperation, the relatively small success of indirect reciprocity strategies is noteworthy. Furthermore, the high standard deviation observed around the average number of individuals adopting a strategy at the end of a simulation indicates the potential presence of very different outcomes.

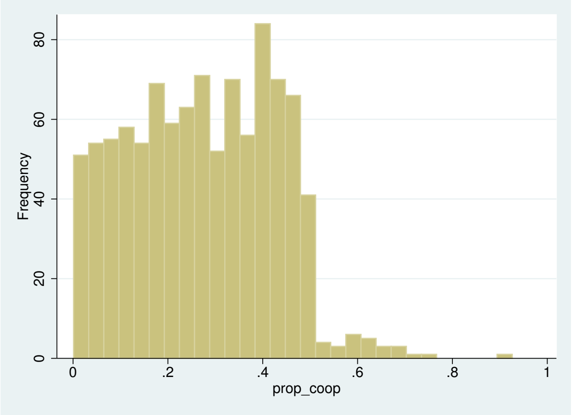


Fig. 1. Distribution of proportion of cooperation actions at t_{conv}. Results from 1000 simulations with populations initially equally divided among the four types of strategies. The value of P_{evo}, P_{ntw}, P_{for}, P_{open} and P_{close} are selected randomly and independently from a distribution between 0 an 1.

Blind indirect reciprocity and cooperation. In order to understand the comparative strength of blind and embedded indirect reciprocal strategies, it is convenient to study separately the cases where only one of them is present alongside the baseline strategies of UD and UC. We run 1000 simulations with initial populations equally divided among UDs, UCs and URs and with P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} extracted uniformly and independently from uniform distributions between 0 and 1. As Figure 2 (Left Panel) shows, the final number of URs has a bimodal distribution. In many cases, UR disappears and the population is progressively dominated by UD. In others, the UR strategy becomes the absolutely dominant one, spreading to the whole population. Simulations characterized by the survival of UR as a non dominant strategy are very rare. The Left Panel of Figure 2 (Left Panel) explores the outcomes for different probabilities of forgiveness P_{for} and clearly shows that UD has the potential to become the absolute dominant strategy under any level of forgiveness. The Right Panel of Figure 2 (Right Panel) shows how the proportion of cooperation increases in cases dominated by the UR strategy with higher levels of forgiveness P_{for}. Notice that not all cases where UR gains dominance are characterized by cooperative behavior. Blind indirect reciprocity often leaves a blind eye for an eye.

[4] The size and density of the network are calibrated throughout the article to resemble that of human ancestors' communities, in line with Dunbar (1992). Moreover from the technical viewpoint the relatively high density almost surely ensures the absence of isolates.

Strategy	%Abs. Dom.	Avg. num.	Std Num	min	max
type					
UD	83%	115.59	98.68	0	200
UC	0 %	0.12	1.35	0	26
CR	10%	13.49	49.40	0	200
UR	7%	10.26	43.31	0	200

TABLE I: Characteristics of simulations absolutely dominated by a strategy. *Notes:* Columns show the proportion of these runs in which each strategy gained absolute dominance, the average number, the standard deviation, the minimum, and the maximum number of agents playing that strategy across these simulation runs. All simulations start from a population equally divided among the four types of strategies. The values of P_{evo}, P_{ntw}, P_{for}, P_{open} and P_{close} are selected randomly and independently from a distribution between 0 and 1.

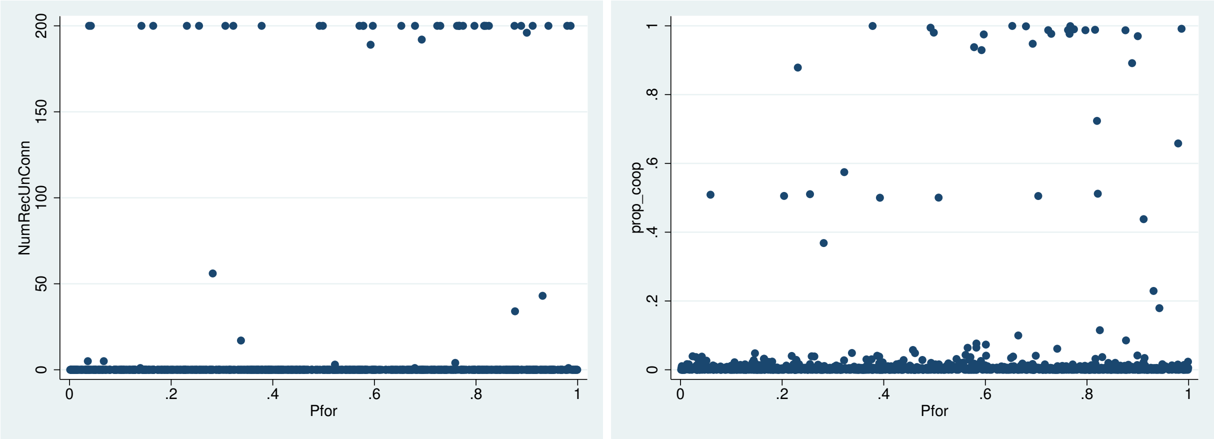


Fig. 2. Left Panel: Number of URs at the end of the simulation. Right Panel: Proportion of cooperative actions at the end of the simulation. Both panels represent the relevant variable against the level of P_{for} associated with the simulation. In all 1000 shown simulations the population is initially equally divided between UD, UC and UR strategies, while P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} are extracted uniformly at random between 0 and 1.

Let us now study the relationship between the perfectness or the speed of evolution and the chances of blind indirect reciprocity. We find that faster evolutionary dynamics is associated with lower chances for the survival of cooperative behaviour. Indeed, the Left Panel of Figure 3 shows that fast evolution completely eliminates the UR strategy: when the speed of evolution is too high the UR strategy cannot adapt fast enough to the defecting behaviour of UDs and disappears consequentially. Observing the Right Panel of Figure 3 it is immediately evident that the maximum level of cooperation achievable by UR strategies decreases in the speed of evolution, after the latter starts to exceed the threshold $P_{evo} = 0.2$. Above this threshold the strategy is unable to sustain the absolute dominance of cooperation.[5]

To conclude our analysis of UR strategies, we now study statistically the level of dominance of this strategy in relation to our key parameter values. For this, we discretize the number of URs at the end of the simulation creating a variable *catUR* that is assigned value 1 if the population of UR strategies is larger than 100, and value of 0 otherwise. We then run a probit regression using *catUR* as the dependent variable and the value of the parameters P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} as independent variables. Results are reported in Table II.

catUR	Coef.	Std. Err.	z	$P > \|z\|$
P_{evo}	-1.351303	.3076922	-4.39	0.000
P_{ntw}	-.4261091	.2804871	-1.52	0.129
P_{for}	.6986143	.2758046	-2.53	0.011
P_{open}	-.2280673	.2728396	0.84	0.403
P_{close}	-.1347793	.2696162	-0.50	0.617
Const.	-1.451655	.3154203	-4.60	0.000

TABLE II: Probit regression of the number of URs at the end of each simulation (discretized as discussed in the main text) on key parameter values.

Table II shows that, as suggested qualitatively by Figure 3, P_{evo} has a strong negative effect (with a p–value lower than 0.01) on the chances of UR to become dominant. On the contrary – and again in line with our qualitative findings, – forgiveness has a positive and significant effect on the chances of UR to become dominant. Finally, P_{ntw}, P_{open} and P_{close} have non-significant coefficients. This implies that the variations in the parameters of network dynamics do not affect significantly the outcomes discussed. The lack of significance of these parameters can be interpreted as a sign of the robustness of the conclusions reached on the conditions that sustain the emergence of cooperation through blind indirect reciprocity.

[5] The value of 1/2 at $P_{evo} = 1$ is due to the constant updating of strategies that start with a 50% probability of cooperation.

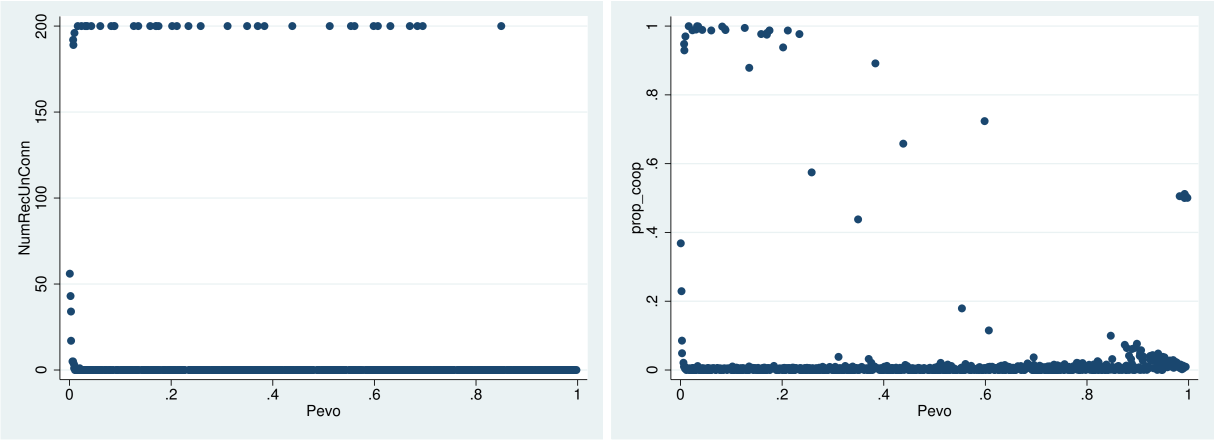


Fig. 3. Left Panel: Number of URs at the end of the simulation. Right Panel: Proportion of cooperative actions at the end of the simulation. Both panels represent the relevant variable against the level of P_{evo} associated with the simulation. In all 1000 shown simulations the population is initially equally divided between UD, UC and UR strategies while P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} are extracted uniformly at random between 0 and 1.

Embedded indirect reciprocity and cooperation. We run a set of 1000 simulations with initial populations equally divided among UDs, UCs and CRs. P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} are all extracted uniformly and independently from uniform distributions between 0 and 1. As Figure 4 (Left Panel) shows, the general relationship between the number of CR strategy registered at the end of the simulation and P_{for} is in line with the one observed in Figure 2 for UR: the majority of simulations ended up with the dominance of UD, but some with the dominance of the CR strategy. The Right Panel of Figure 4 highlights, however, an interesting difference concerning the levels of cooperation that can be sustained. For low levels of the P_{for} parameter, while CR can become dominant, it does so by mimicking the UD strategy, thus the whole population acts as defectors. Only when $P_{for} > 0.4$, cooperation grows significantly, eventually reaching absolute dominance. This is in contrast with what is observed for UR that sustains high levels of cooperation also for lower levels of forgiveness. This difference is due to the underlying difference between the CR and the UR strategies. Indeed, blind indirect reciprocity (UR) has potentially access to a larger number of individuals as a potential source of information for conditioning its behaviour toward the interacting partner. On the contrary, CR is restricted to interact with common partners $(F_j^t \wedge F_i^t)$, and hence forgiveness provides less additional advantages for its success.

Figure 5 shows that the speed of evolution is not favorable for the embedded indirect reciprocity strategy. Despite the broad similarity with Figure 3, there are two notable differences. First, CR can become dominant somewhat more likely than UR with fast strategy evolution. Second, the number of simulations in which a high level of cooperation is sustained is somewhat higher in Figure 5, but drops to zero sharply around $P_{evo} = 0.4$. In contrast, Figure 3 displays a progressive decrease by the speed of evolution, but with more internal variation. The reason for this slight difference is - again - the different set of information accessible to the two indirectly reciprocal strategies. The results show that the larger, but sparser, network of information accessible to UR allows for faster adaptation to defecting behaviour than the tight but smaller network of CR, thus allowing for higher levels of cooperation to be sustained for $0.4 < P_{evo} < 1$.

catCR	Coef.	Std. Err.	z	$P > \|z\|$
P_{evo}	-1.144433	.3030948	-3.78	0.000
P_{ntw}	.1400632	.2657115	-0.53	0.598
P_{for}	.9659768	.2868299	-3.37	0.001
P_{open}	.3467598	.2707578	1.28	0.200
P_{close}	-.0308959	.2661878	-0.12	0.908
Const.	-2.073514	.3350006	-6.19	0.000

TABLE III: Probit regression of the number of CRs at the end of each simulation (discretized as discussed in the main text) on parameter values assumed by each of the model parameters.

We finally run a statistical analysis of the level of dominance of CR in relation to key parameter values. Similarly to the UR case, we discretized the number of CRs at the end of the simulation creating a variable *catCR* that - for each simulation - takes value 1 if the number of CR is larger than 100 and value 0 otherwise. We then run a *probit* regression using *catCR* as dependent variable and parameters P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} as independent variables.

Results are reported in Table III which confirms our qualitative results also for CR. The chances of CR to become dominant are shown to decrease significantly in P_{evo} and significantly increase in P_{for}. Both results have very high levels of significance with $p \leq 0.001$.

Finally, the parameters associated with P_{close}, P_{open}, and P_{ntw} are not significant. This is reassuring about the robustness of our findings under different parametrizations of the network dynamics.

CONCLUSIONS

Against rational interests and equilibrium predictions, humans display a high extent of cooperation, also

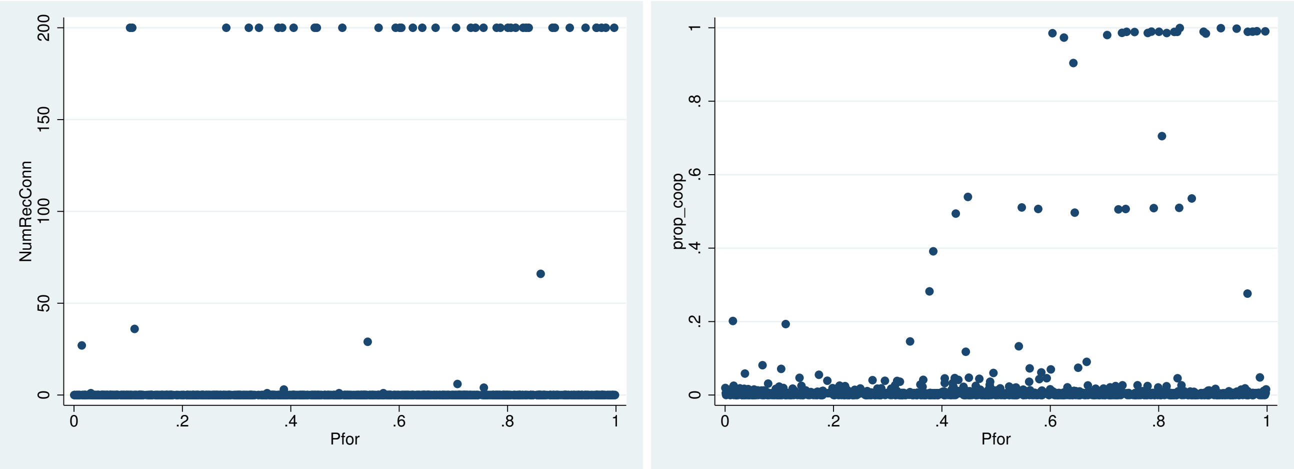


Fig. 4. Left Panel: Number of CRs at the end of the simulation. Right Panel: Proportion of cooperative actions at the end of the simulation. Both panels represent the relevant variable against the level of P_{for} associated with the simulation. In all 1000 simulations the population is initially equally divided between UD, UC and CR strategies while P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} are extracted uniformly at random between 0 and 1.

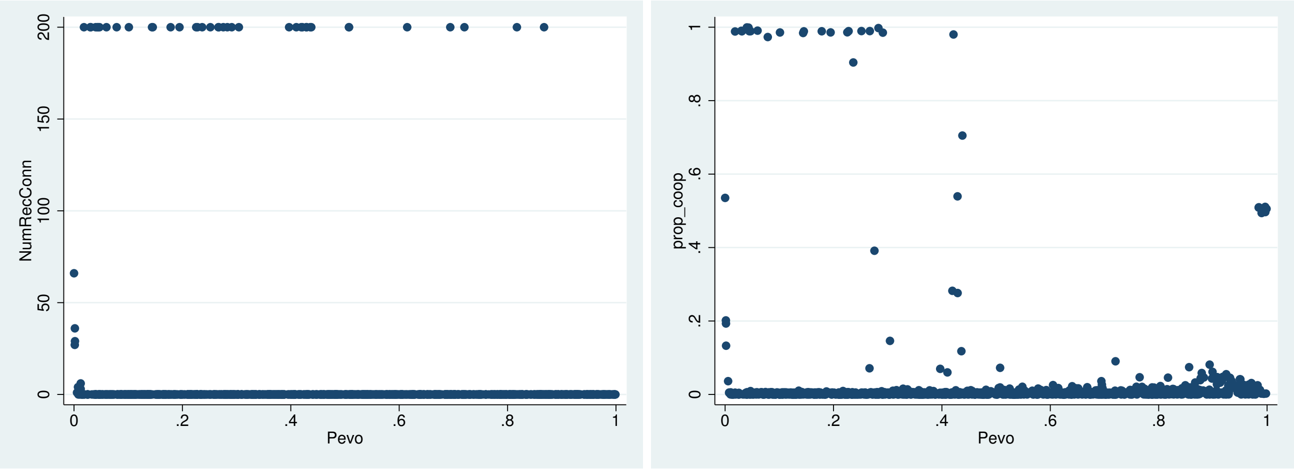


Fig. 5. Left Panel: Number of CRs at the end of the simulation. Right Panel: Proportion of cooperative actions at t_{conv} or at t_{max}. Both panels represent the relevant variable against the level of P_{evo} associated with the simulation. In all 1000 simulations the population is initially equally divided between UD, UC and CR strategies while P_{evo}, P_{ntw}, P_{for}, P_{open}, and P_{close} extracted uniformly at random between 0 and 1.

against strangers. Previous research has suggested that indirect reciprocity might explain substantial cooperation levels. The objective of this study was to assess the effectiveness of two different types of indirect reciprocity strategies in their capacity to sustain cooperation. We modeled interactions between agents with the two-person Prisoner's Dilemma (PD), which is the most puzzling and best studied social dilemma situation. In our model, agents represented humans interacting on a random and developing network. We implemented three mechanisms of network change and examined the chances of cooperation by two different types of indirect reciprocity strategies under a full range of combinations. Precisely, we assumed different combinations of random rewiring of ties, a mechanism of opening triads (creating structural holes), and a mechanism of closing triads (increasing clustering).

Network relations were crucial for our model, because these connections provided agents with the opportunity to condition their actions on the information they received from and about network partners. We studied two forms of indirect reciprocity: a blind one that rewards any partner who did good to previous partners, and an embedded one that conditions cooperation on good acts towards common acquaintances.

Our results show that both strategies sustain similar levels of cooperation in situations characterized by the presence of unconditional defectors. Moreover, for both strategies their success depend strongly and negatively on the speed of evolution for strategies. In contexts characterized by fast evolution, both strategies are dominated by defection and disappear or become functionally equivalent to unconditional defectors, thus surviving by always defecting. Forgiveness – the probability of indirectly reciprocal strategies to give a second opportunity to defecting partners by observing their behaviour with a second partner – is shown to have a positive effect on the capacity of both strategies to sustain cooperation.

While the two types of indirect reciprocity strategies share similar patterns of behaviour, our analysis also uncovered differences in their efficiency. First, the blind indirectly reciprocal strategy (UR) supports relatively higher levels of cooperation even at low levels of forgiveness. The embedded indirect reciprocity strategy needs a higher probability to offer a second chance to sustain the same level of cooperation. Second, the embedded indirect reciprocity strategy (CR) is unable to sustain cooperation when strategies evolve fast while, in some cases, the blind strategy could overcome the pressure provided by the high speed of evolution. The reason for these differences can be attributed to the different set of information accessible to the two indirectly reciprocal strategies. Indeed, the blind strategy can access a broader network to condition its behaviour and thus can adapt its behaviour fast and acquire more information through forgiveness. The embedded strategy relies only on a smaller set of common partners, thus the acquired information is more local and less diverse, making it more difficult to discern pure defectors and other conditional players.

A limitation of our study is that we studied relatively small networks, in which indirect reciprocity might be less important than direct forms of reciprocity. Furthermore, small populations might also be prone to drift-like behavior resulting in local equilibria that would not be feasible in larger populations. However, we selected the relatively small network size because this corresponds better to the natural size of human groups, in the evolutionary past in particular. Our results are shown to be robust to even strong changes in the mechanisms of network dynamics underlying the interactions among the agents.

References

Axelrod, R. (1984). *The Evolution of Cooperation.* Basic Books.

Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489):1390–1396.

Boyd, R. and Richerson, P. J. (1989). The evolution of indirect reciprocity. *Social Networks*, 11(3):213–236.

Dunbar, R. I. (1992). Neocortex size as a constraint on group size in primates. *Journal of human evolution*, 22(6):469–493.

Grafen, A. (1984). Natural selection, kin selection and group selection. *Behavioural ecology: An evolutionary approach*, 2:62–84.

Gunnthorsdottir, A. and Rapoport, A. (2006). Embedding social dilemmas in intergroup competition reduces free-riding. *Organizational Behavior and Human Decision Processes*, 101(2):184–199.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. ii. *Journal of theoretical biology*, 7(1):17–52.

Hauert, C. and Doebeli, M. (2004). Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature*, 428(6983):643–646.

Hoffmann, R. (2000). Twenty years on: The evolution of cooperation revisited. *Journal of Artificial Societies and Social Simulation*, 3(2):1390–1396.

Lieberman, E., Hauert, C., and Nowak, M. A. (2005). Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316.

Nowak, M. and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, pages 1291–1298.

Nowak, M. A. and Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685):573–577.

Puurtinen, M. and Mappes, T. (2009). Between-group competition and human cooperation. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1655):355–360.

Righi, S. and Takács, K. (2014). Emotional strategies as catalysts for cooperation in signed networks. *Advances in Complex Systems*, 17(02):1450011.

Sachs, J. L., Mueller, U. G., Wilcox, T. P., and Bull, J. J. (2004). The evolution of cooperation. *The Quarterly Review of Biology*, 79(2):135–160.

Santos, F. C., Pacheco, J. M., and Lenaerts, T. (2006). Cooperation prevails when individuals adjust their social ties. *PLoS Computational Biology*, 2(10):e140.

Sigmund, K. (2012). Moral assessment in indirect reciprocity. *Journal of theoretical biology*, 299:25–30.

Smith, E. A. (2010). Communication and collective action: language and the evolution of human cooperation. *Evolution and Human Behavior*, 31(4):231–245.

Sommerfeld, R. D., Krambeck, H.-J., Semmann, D., and Milinski, M. (2007). Gossip as an alternative for direct observation in games of indirect reciprocity. *Proceedings of the national academy of sciences*, 104(44):17435–17440.

Wedekind, C. and Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, 288(5467):850–852.

West, S. A., Griffin, A. S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of evolutionary biology*, 20(2):415–432.

Whitaker, R. M., Colombo, G. B., Allen, S. M., and Dunbar, R. I. M. (2016). A dominant social comparison heuristic unites alternative mechanisms for the evolution of indirect reciprocity. *Scientific Reports*, 6:31459.

Wu, J., Balliet, D., and van Lange, P. A. M. (2015). When does gossip promote generosity? indirect reciprocity under the shadow of the future. *Social Psychological and Personality Science*, 6(8):923–930.

SIMONE RIGHI is Assistant Professor in Agricultural Economics at the University of Bologna and Researcher at the "Lendület" Research Center for Education and Network Studies of the Hungarian Academy of Sciences.

KAROLY TAKACS is Researcher and Director of the "Lendület" Research Center for Education and Network Studies of the Hungarian Academy of Sciences

MODELLING THE DEVELOPMENT OF STRATEGIC MANAGEMENT

Nikolett DEUTSCH – Tamás MÉSZÁROS – Lajos SZABÓ
Department of Strategy and Project Management
Corvinus University of Budapest
1093 Budapest, Fővám str. 8.
E-mail:lajos.gy.szabo@uni-corvinus.hu

KEYWORDS

Strategy, development, history, model, SM Cube, paradigm, literature research

ABSTRACT

Nowadays the are several research focusing on different approaches of strategic management. Because of the wide-ranging approaches and interpretations, it is very important to consider it from different perspectives. This paper provides a systematic view of Strategic Management developed by the Strategic Management Researh Group of the Department of Strategy and Project Management at the Corvinus University of Budapest. Using the Strategic Management Cube we are able to examine and model the the effects of environmental components, empirical experiences, and the development of other sciences and disciplines on the evolution of SM as well as the interplay among them.

1. THEORETICAL BACKGROUD

Strategic management (further called "SM") has been a quite popular scientific field. With its more than five-decade history produced significant international journals and books that could fill up whole libraries. Therefore, it is indeed hard to achieve new or novel scientific results for scholars of SM. It is by no accident, that citations in journal articles often already make it difficult to actually grasp the main argument of the article, not to even mention ever growing lists of cited sources are close to one-fifth of article length. Besides a wealth of rich and comprehensive resources, professionals of this fields need to face further difficulties.

First and foremost, taking note that the theory and practice of SM was formed in the second half of the last century, but considering its most flourishing period, it was rather the last third of the same century. The analysis of numerous textbooks and academic literature proved that the above mentioned period brought to life theories, paradigms, schools, models and methods that are considered determinative even today. It is not at all unexpected, that the creators of these theories and schools marched into scientific history as the "gurus" of SM, and that any publication dealing with SM cannot afford to ignore their works.

As follows, the foregoing scholars of SM can present scientific results by partially enriching the existing theories, presenting arguments against them or their practical application or by doing research into the intersections of other scientific fields and SM, their influence and thus produce new or novel results.

Not independently from the above, but deriving from the nature of this scientific field, empirical studies have an important role. Also, from the view of scientific acceptance, a central question is the relationship of strategy and achievement, but other types of surveys and case studies are also subject to these studies.

It could be put that studying the development of this scientific field itself, and its five-decade long process is close to becoming a whole separate scientific branch. The framework within which this study is carried out, what analytical focus points does the process description utilize, could mean a unique and novel approach, scientific result in itself.

2. AIM OF THE RESEARCH

The Strategic Management Research Group of the Department of Strategy and Project Management at the Corvinus University of Budapest started a research project last year. The goal of this research is not to enrich the basis of the above mentioned results, but it is aimed at providing a well-structured view of the half-century development of SM so far, systematically examining and modelling the effects of environmental components, empirical experiences, and the development of other sciences and disciplines on the evolution of SM as well as the interplay among them. The research group made a literature analysis based on which the first results can be introduced and a three dimensional model – the Strategic Management Cube – can be drawn.

Systematisation is guaranteed by the simultaneous utilisation of two interrelated analytical viewpoints, i.e. the intersections between time reference and the alterations of SM's content items form the subject of research. Besides temporality, it is advisable to include to more starting points:

a) Strategy, as a tool of thinking about the future, is the "product" of market competition. Its basics date back to the '60s, when the conjuncture-generating effects of the deferred demand of World War II started to fade. The fast development of capacities produced a demand driven economy built on the mechanism of market competition. A time intersection means the '60s, when it is a period strategic leadership based on process

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

(**process based view**). Its main representatives are Chandler (1962), Ansoff (1965) and the "founders" of Harvard school (Learned et al. 1965).

b) The development of SM in itself can also be studied – what steps and what factors lead from short-term planning focusing on budget to SM, according to some, thus already exceeding the strategy of complex systems (**activity based view**). Factors of generating change, such as globalisation, technological advancement, socio-political changes can also be included in this thought line. Not independently from the above, the influence of other sciences also forms a weighty part in it (e.g. equilibrium theory in economics). This approach concentrates on the view of strategic thinking, strategic behaviour, the emphasis of strategy alignment and realisation, the relationship between strategy and achievement.

However, there exists a more concrete, maybe not independent from the above, but detailed analysis intersection, which involves the study of SM content elements such as theories, paradigms "providing competitive advantage", goal-setting points (foci of interest) and strategic management tools and techniques, as well as the mutual relationship of these elements and their self-development. Descriptive-prescriptive, process- or content-oriented approaches and their scientific background can also be fit into this framework.

Based on the aforementioned statements the main question of the research project is defined as follows: Is SM, as a widely used management approach applied both in theory and practice, in the ascending or descending stage of its "life cycle" – i.e. is there a need for a "new management paradigm"?

In order to answer this question, the research project had been divided into four research stages:

- development and refinement of the applicable research framework;
- literature research and processing;
- practical application, methodology and experiences in connection with the different research streams and paradigms of SM;
- creation of the anthology of academic literature excerpts on strategy.

Due to the fact that the results of the first research phase, i.e. the elaboration of the theoretical and cognitive analytical framework define the unique rules of sorting followed by the research group and form the basis of further steps and examinations, in the next Chapter our Holistic Model of the development of Strategic Management and the fundamental guiding principles, driving forces of its formation will be presented.

3. RESEARCH RESULTS

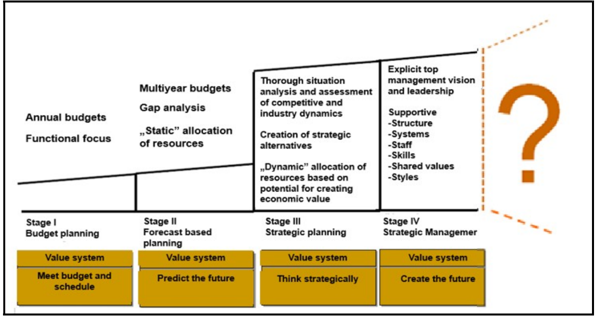
During the examination of the term "strategy" and its relations to the evolution of planning and other fields of science, the different paradigms on strategic planning and management, and the dominant themes of strategic literature two main research hypothesis and additional questions had been emerged.

Thesis 1: The essence of SM has been formed during the second half of the 20th century, however its most flourishing period is concerned to be the last third of it.

How can the process of the evolution and self-development of SM be described (from short-term financial planning to strategic management and beyond)?

McKinsey&Co. provides an internationally accepted presentment of the development process by differentiating four stages: budget planning, forecast based planning, strategic planning and strategic management (Figure 1). Scholars have established that the term "strategic management" was coined by Ansoff, who used it for the first time in 1972 at a conference (Tari 1996). When he realised that strategic planning is only successful in a minority of applications, Ansoff (1965) questioned if strategic planning was a wrong theory or something unfinished that was in need of further development? The answer is the latter, so the unfinished tool, which does not involve the management of change in itself (Bhatia 2014). Grant (2008/c) deals with the series of change that follow each other, obviously supporting the time horizon until the first decade of the third millennium. It is important to note, that this notable scholar of strategic thinking speaks of the evolution of SM and describes the process with the changes "dominant topics" within the system, however he does not establish a new stage that would exceed SM in the development process. The elaboration of studies, articles and their abstracts that were published on the occasion of notable anniversaries in the most renowned European journal of the discipline, Long Range Planning, also belongs to antecedents of this project. It must be mentioned, that the article of former editor-in-chief Taylor (1986) published for the 20th anniversary can serve as a fundamental model for further research, as well as the works of Cummings and Daellenbach (2009) which draw conclusions on changes within SM by an analysis of the abstracts of the first 40 years. It is also worth to invoke a study of O'Shanassy (2001) summarising the results of the past century, and the recently published study of Bhatia (2014).

Figure 1: The evolution of strategic management



Source: Gardener et al (1986, p. 25)

How can the most important driving forces, such as the change of environment and the impacts of the development of other sciences and disciplines, be revealed?

In his 1986 article, titled „Corporate Planning for the 1990s: The New Frontiers", Bernard Taylor - as he states - read about a thousand articles and not only depicts the development of planning divided into periods and the essential features of each period, but also presents – as an analytical model – how the planning, later strategic, features of a given period have emerged or changed under the influence of environment factors. Thus he clearly proved that planning and strategic thinking, or to say, the evolution of SM was not just a spontaneous process, but it is indeed the result of environmental change. With all the above, he encourages further research in at least two more fields. Partially, he encourages the description and understanding of changes within a process (see European Business Forum 2001), and voices a question towards the SM "gurus" and top managers after the blast of the dotcom bubble and 9/11 events in 2001. The question was voiced as follows: „Does strategy still have a meaning?"). Also, detected changes in environment adumbrate the need for change and its content within SM (or exceeding it). Grant does this in the last chapter of his book (2008/c) „Current trends in strategic management", when he describes the summarising features of trends in the external environment of enterprises – third industrial revolution, social pressure, decay of public enterprises- and draws conclusions from these, among many, about the strategy of complex systems or the coming into view of adaptation strategies (Grant 2008/c).

Thesis 2: SM was created by competition, which is a central category of business even nowadays. Although the intensity (e.g. due to hyper competition), extension, and predictability of competition have been modified over time the essence of SM has not changed.

Did the environmental and organisational complexity and change achieve a certain level which calls for a new paradigm (or already has formed one)? And if so, how does that differ from SM?

The contradiction between Thesis 2 and this relevant question is not accidental. There is no definitive answer to this question yet. Rapid changes in the environment raise the need for a new paradigm and some scholars made an attempt to define it.

The works of Prahalad and Hamel (1994), Hermann (2005), Grant (2008/c) and Cummings and Daellenbach (2009) provide much help to the thought about SM's future. A new theory exceeding the SM paradigm can be found in the studies of O'Shanassy (2001) – „strategic thinking"- and Bhatia (2014) - „complex strategic system" – and something similar is outlined in the appearance of „strategy as a practice", (Jarzabkowski et al. 2009). Our review conducted that these "experiments" don't reach the theoretical and methodological credibility of strategic management, there is no justification of their practical applicability.

How can the direct and indirect impacts of the increasing complexity of the world and the development of the related disciplines on SM be identified and analysed?

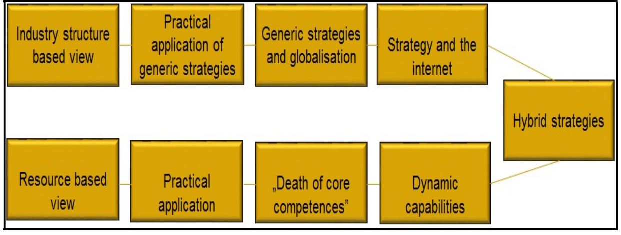
Grant (2008/c) does not question the "survival" of SM, but he does deal with "adaptation strategies" within it. He cites Jack Welch „3S" (Speed – Simplicity – Self-confidence), according to whom break-up with conventions, spontaneity and intuition is essential and has come forward. Even more dynamic development, significant changes in approach, content and techniques can be experienced through the study of the mutual relationship of content elements of SM's system. What are considered content elements? Answers provided to the following questions, more exactly, specific theories, models, schools and paradigms:

- What is the source of competitive advantage? - paradigms [1]
- What are the key foci of interest? - value creation
- What are the key components of the tool system of strategy realisation and alignment? - organisation - business models - methodology (techniques, process)

The general introduction of each element's structure: the basic model - theoretical background, connection - practical application - critique - survival - further development (change, transformation) - exceedance Figure 2 provides an example for this.

[1] In academic literature, one can read about SM as paradigm (O'Shanassy 2001), but one can also find similar notions related to other theories (e.g. industry structure based view) (Evans 2000).

Figure 2: A structure of paradigms' research results


Industry structure based view
Practical application of generic strategies
Generic strategies and globalisation
Strategy and the internet
Hybrid strategies
Resource based view
Practical application
„Death of core competences"
Dynamic capabilities

Source: edited by research group

What is the source of competitive advantage? - one of the basic questions of SM. The industry position that forms the space for competition **(industry structure based view)** or those fundamental resources and capabilities, which create the base of strategy by differentiating itself from others (**resource based view**). In the former case (outside-in analysis) the source of profit is the so-called monopoly rent, while in the latter (inside-out analysis) the Ricardian rent is named (Grant 1991). Both approaches can be considered a paradigm of SM, and is subject to theory and practice until today, including the Porterian industry sector analysis model's modifications, as well as the trend of emphasis shift in resource-based theory into the direction of knowledge-based dynamic capabilities (**knowledge based view**). The basics can be found in the publications of Porter (1980; 1986), Hamel and Prahalad (1990), Grant (1991; 2008/d). Their theoretical background includes transaction cost theory, business economics and resource-based corporate theory. Critiques can be found in the works of Evans (2000), Mintzberg (1998), Tapscott (2001), Grant (2008/b; 2008/e), but Carr's (2013) – "Death to Core Competency" could also be included in this list. Further development can be found in the later works of Porter (Porter 2001/a; 2001/b, Porter and Reinhardt 2006; Porter and Kramer 2007) in the book "Blue ocean strategy" (Kim et al. 2005), McGrath (2013a) and his article published in HBR in 2013 (McGrath 2013b). Requirements and characteristics of dynamic capabilities are linked to the study of Teece and Pisano (1994), Teece et al. (1997). Hybrid strategies aimed at merging two paradigms, and the ever more popular studies on dual-ability enterprises also belong here (Lapersonne et al. 2015; Spanos et al. 2001). A separate big "chapter" of development are the so-called international strategies (Czakó and Reszegi 2010; Luthans and Doh 2014), which expand the system of SM to an international stage while it also produces new mutations.

Whose interests should be reflected in strategy, where should it come from? – another basic question, the answers to which has also invoked a lot from different theories. We would refer to the foci of interest and their clearly stated theoretical background. Customer value and main motive, mutatis mutandis plays an important role in this context. The topic boasts a huge amount of related academic literature, mostly in connection with marketing. It may be an exciting question, how core competence vs. consumer value should be treated in strategy alignment. With the advancement of technology (e.g. big data analysis), it also sheds new light on the application of focus of interest. From the SM view of it, the works of Prahalad and Ramaswamy (2000), Anderson and Rust (1997), and Kordupleski and Simpson (2003) could be considered an important base. Shareholder value is the second possible starting point, to which "principal-agent" theory can be linked. Determining focus of interest in corporate practice can be linked to the work of Rappaport (1986; 2006), though it has received much criticism from scholars. Obvious consequences are corporate accounting scandals (Grant and Visconti, 2008/a). Prahalad and Hamel (1994) also puts its extreme enforcement under criticism, while Porter and Kramer (2011) see shared value creation that appeared instead of it as a new motivation of capitalism. In reality, it is a fundamental goal of enterprises to satisfy both values, which is called "dual value creation" by Chikán (2008). The circle of stakeholders that includes the former and expands the groups which are in a relationship of interest with enterprises, as well as its requirements play an important role in strategic goal setting (Ackoff 1981; Freeman 1984). Strategic thinking built on stakeholder theory underwent multi-directional development in the last decade, examining the mutual relationship of groups of interest in a so-called "power-interest network". This network is based on the classification of dimensions of effect and interest binding on enterprises of those involved. Corporate social responsibility, as a continuously strengthening element that influences strategy and strategic goal-setting, can be linked to involved theories by its own philosophy and practice. Furthermore, the expectations of shared value creation - according to which, value for the enterprise and society has to be created at the same time - has the same roots (Porter and Kramer, 2011).

It can be sensed from the above, that the theoretical background of SM, theoretical base, paradigms and foci of interest "inseminated" each other, changed and developed parallel to each other. Tools (techniques, methods and models) which support strategy alignment (underlying analysis) and decision-making (decision models) – see Bain&Co's Management Tools Survey (Rigby 2013), Becker et al. (2005) and Berényi (2015) - as well as realisation (organisation, culture, motivation and feedback mechanisms) are closely related. One can count on the works of Balaton et al. 2014, Dobák 2008; Bakacsi 2004 in this topic. The systemised picture of the decades of SM development therefore cannot be short of this toolbox, as the follow-up and presentation of the time change an important content element.

4. MODEL DEVELOPMENT FOR ANALYSING THE DEVELOPMENT OF STRATEGIC MANAGEMENT

Taking into account these findings it can be stated that the development of SM can be investigated from three different but interrelated research dimensions. The result of the analysis depends on to what extent the dimensions are highlighted.

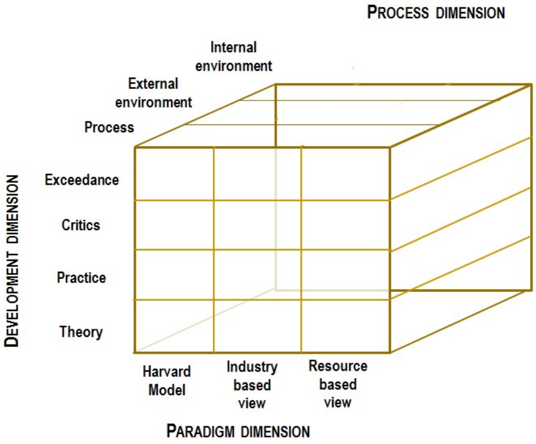
AVP = f (D1, D2, D3)

where
AVP: The dominant analytic viewpoint
D1: Paradigm Dimension
D2: Development Dimension
D3: Process Dimension

As Figure 3 summarises these dimensions are:

- Paradigm dimension: as it was shown above, according to the relevant literature (e.g.: Evans 2000; Herrmann 2005) during the development process of strategic thought three dominant designs or paradigms – business policy, competitive analysis and resource based theory of the firm - had emerged with their own presumptions on the sources of competitive advantage.
- Development dimension: besides the abovementioned disruptive innovations, strategic paradigms had also been developed by incremental innovations which can be identified in the SM literature dealing with the fundamental concepts, their practical uses, their critics and their modifications.
- Process dimension: over the past 50 years different concepts had been appeared regarding the process of strategic planning and implementation accompanied by new and modified strategic models and tools applicable in the particular steps of the process.

Figure 3: Holistic view of the development of Strategic Management



Source: edited by research group

The practical application of the model enables researchers to analyse different segments of Strategic Management. In every analysis there is a well-defined segment which represent the dominant viewpoint of the researcher. Using this model the analitic map of the development of Strategic Management can be created.
The folloving examples indicate the content of these segments.

Example 1.
AVP=f(D13, D21, D33)
where
D13= Resource based view
D21= Theory
D32= Internal environment

Theoretical studies and publications belong to this segment are based on the presumptions of RBV regarding the importance of corporate resources and compenencies, assume that competitive advantage can be derived from Ricardian rent and aim to develop analytical tools or techniques for the easier identification of the sources of permanent competitive advantage, such as the theoretical foundation of the VRIO framework (Barney, 1991).

Example 2.
AVP=f(D12, D23, D32)
where
D11= Industry based view
D21= Critics
D32= External environment
Those articles, researches and works can be classified into this segment which highlight the weaknesses of the analytical techniques developed and utilised by the representatives of the outside-in paradigm. For example the aforedmentioned article of Evans (2000) regarding the usability of Porter's five forces model in the internet era can be mentioned here.

5. FURTHER RESEARCH

The Strategic Management research group of the Department of Strategy and Project Management at the Corvinus University of Budapest is going to start the next phase of this research. The goal of the next phase is to collect case studies, best practices which best represent the evolution of paradigms of strategic management. Furthermore, it aims at identifying and analysing new trends, theories and practices can lead to a new paradigm.

REFERENCES

Ackermann, F. and C. Eden. 2011. “Strategic Management of Stakeholders: Theory and Practice.” *Long Range Planning* 44, No.3, 179-196.

Ackoff, R. 1981. *Creating the Corporate Future*. John Wiley and Sons, New York

Anderson, T. and R. T. Rust. 1997. “Customer Satisfaction, Productivity, and Profitability: Differences between goods and services.” *Marketing Science* 16, No.2, 129-145.

Ansoff, I. 1965. *Corporate Strategy*. McGraw Hill, New York

Bathia, V. K. 2014. “Strategic Management – History and Development”. http:vijaykumarbhatia.weebly.com/strategic-management-history-a … Accessed: 2014.01.14.

Bakacsi, Gy. 2004. *Szervezeti magatartás és vezetés*. Aula Kiadó, Budapest, Hungary

Balaton, K.; L. Hortoványi; E. Incze; M. Laczkó; Zs. R. Szabó; and E. Tari. 2014. *Stratégiai menedzsment.* Akadémiai Kiadó, Budapest, Hungary

Barney, J. B. 1991. “Firm resources and sustained competitive advantage.” *Journal of Management* 17, No.1, 99-120

Becker, P., A. Turner; J. Varsányi; and M. Virág. 2005. *Értékalapú stratégiák: A pénzügyi teljesítmény értékvezérelt menedzsmentje.* Akadémiai Kiadó, Hungary

Berényi, L. 2015. “Lean fejlesztés: értékteremtés vagy veszteségek megszűntetése.” *Marketing & Menedzsment* 49, No.2, 47-60.

Carr, A. 2013. “Death to Core Competency: Lessons from Nike, Apple, Netflix”. http://www.fastcompany.com/3005850/core-competency-dead-lesso... Accessed: 2014.01.14.

Chandler, A. 1962. *Strategy and Structure: Chapters in History of the American Industrial Enterprise*. M.I.T. Press, Cambridge, MA.

Chikán, A. 2014. “Hidak a közgazdaságtan és a gazdálkodástan között: vállalatelméletek.” *Magyar Tudomány* 175, No.8, 914-918.

Chikán, A. 2008. *Vállalatgazdaságtan*. AULA Kiadó, Budapest, Hungary

Cummings, S. and U. Daellenbach. 2009. “A Guide to the Future of Strategy? The History of Long Range Planning.” *Long Range Planning* 42, No.2, 234-263.

Czakó, E. and L. Reszegi. 2010. *Nemzetközi vállalatgazdaságtan*, Alinea Kiadó, Budapest, Hungary

D’Aveni, R. 1994. *Hypercompetition: managing the dynamics of strategic maneuvering.* The Free Press, New York

Dobák, M. 2008. “Átjárhatók-e a szervezeti határok?” *Harvard Business Review* 10, No.6, 31-37.

Evans, P. 2000. “Strategy: The End of the End Game.” *Journal of Business Strategy 21, No.6,* 12-18.

Freeman, R. E. 1984. *Strategic Management: a stakeholder approach.* Pitman, Boston, MA.

Gardner, J. R.; R. Rachlin; and A. Sweeney. 1986. *Handbook of Strategic Planning*. John Wiley and Sons, New York, N.Y.

Gavetti, G. and D. A. Levinthal. 2004. “The Strategy Field from the Perspective of Management Science: Divergent Strands and Possible Integration.” *Management Science* 50, No.10, 1309-1318.

Grant, R. M. and M. Visconti. 2008/a. “A vállalati számviteli botrányok stratégiai háttere”. In *Tudás és stratégia, Siker dinamikus környezetben 2008* Rajk László Szakkollégium (Eds.). Alinea Kiadó, Budapest, Hungary, 205-237.

Grant, R. M. 2008/b. “Stratégiai tervezés turbulens környezetben: a vezető olajvállalatok példája”. In *Tudás és stratégia, Siker dinamikus környezetben 2008* Rajk László Szakkollégium (Eds.). Alinea Kiadó, Budapest, Hungary, 279-325.

Grant, R. M. 2008/c. *Contenporary Strategy Analysis.* Blackwell Publishing Ltd, Oxford, UK

Grant, R. M. 2008/d. “A versenyelőny erőforrás alapú elméletének jelentősége a stratégiaalkotásban”. In *Tudás és stratégia, Siker dinamikus környezetben 2008* Rajk László Szakkollégium (Eds.). Alinea Kiadó, Budapest, Hungary, 9-34.

Grant, R. M. 2008/e. “Úton a tudás alapú vállalatelmélet felé”. In *Tudás és stratégia, Siker dinamikus környezetben 2008* Rajk László Szakkollégium (Eds.). Alinea Kiadó, Budapest, Hungary, 35-58.

Grant, R. M. 1991. “The Resourced-Based Theory of Competitive Advantage; Implications for Strategy Formulation.” *California Management Review* 33, No.3, 114-135.

Hermann, P. 2005. “Evolution of Strategic Management, The Need for New Dominant Design.” *International Journal of Management Reviews* 7, No.2, 111-130.

Jarzabkowski, P. and A. P. Spee. 2009. “Strategy as a practice: A review and future directions for the field.” *International Journal of Management Reviews* 11, No.1, 69 – 95.

Kapás, J. 1998. “A vállalati stratégia elméletei.” *Vezetéstudomány* 29, No.11, 47-55.

Kim, W.C. and R. Mauborgne. 2005. *Blue Ocean Strategy: How to Create Unconteste*. Harvard Business School Publishing Co, Boston, MA.

Kordupleski, R. and J. Simpson. 2003. *Mastering Customer Value Management, The Art and Science of Creating Competitive Advantage*. Pinnaflex Educational Resources Inc., Cincinnati, OH.

Lapersonne, A.; N. Sanghavi; and C. De Mattos. 2015. “Hybrid Strategy, Ambidexterity and Environment: toward an Integrated Typology.” *Universal Journal of Management* 3, No.12, 497-508.

Learned, E. P.; C.R. Christensen; K. R. Andrews; and W. D. Guth. 1965. *Business Policy: Text and Cases*. lrwin, Homewood, IL.

Luthans, F. and J. P. Doh. 2014. *International Management.* McGraw Hill, New York, N.Y.

McGrath, R. 2013a. *The End of Competitive Advantage: How to Keep Your Strategy Moving as Fast as Your*

Business. Harvard Business Review Press, Boston, MA.

McGrath, R. 2013. "Transient Advantage." *Harvard Business Review* 91, No.6, 61-70.

Mintzberg, H.; B. Ahlstrand; and J. Lampel. 1998. *Strategy Safari – A Guided Tour Through the Wilds of Strategic Management.* Free Press, New York, N.Y.

O'Shanassy, T. (2001): "Lessons from the Evolution of the Strategy Paradigm." *Journal of Management & Organization* 7, No.1, 25-37.

Porter, M. and M. Kramer. 2011. "Creating Shared Value. How to Reinvent Capitalism – and Unleash a Wave of Innovation and Growth." *Harvard Business Review* 89, No.1-2, 62-77.

Porter, M. and F. Reinhardt. 2007. "Strategic Approach to Climate." *Harvard Business Review* 85, No.10, 22-26.

Porter, M. and M. Kramer. 2006. "Strategy and Society. The Link Between Competitive Advantage and Corporate Social Responsibility." *Harvard Business Review* 84, No.12, 78-92.

Porter, M. 1980. *Competitive Strategy: Techniques for Analysing Industries and Competitors.* Free Press, New York, N.Y.

Porter, M. 1986. *Competition in Global Industries*. Harvard Business School Press, Boston, MA.

Porter, M. 2001/a. "Now Is the Time to Rediscover Strategy." *European Business Forum* 8, 20-21.

Porter, M. 2001/b. "Strategy and the Internet." *Harvard Business Review* 79, No.3, 63-79.

Prahalad, C.K. and V. Ramaswamy. 2000. "Co-opting Customer Competence." *Harvard Business Review* 78, No.1, 79-87.

Prahalad, C. K. and G. Hamel. 1990. "The Core Competence of the Corporation." *Harvard Business Review 66*, No.3, 79-91.

Rappaport, A. 1986. *Creating Shareholder Value*. The Free Press, New York, N.Y.

Rappaport, A. 2006. "10 Ways to Create Shareholder Value." *Harvard Business Review* 84, No.9, 66-77

Rigby, D. K. 2013. *Management Tools Survey 2013*. An executive's guide, Bain & Company, Boston, MA.

Spanos, Y.E. and S. Lioukas. 2001. "An examination into the casual logic of rent generation: Contrasting Porter's competitive strategy framework and the resource-based perspective." *Strategic Management Journal* 22, No.10, 907-934.

Tapscott, R. 2001. "Rethinking Strategy in a Networked World, or Why Michael Porter Is Wrong about the Internet." *Strategy and Business* 24 (July), 1-8.

Tari, E. 1996. "A stratégiai menedzsment elméletének és gyakorlatának fejlődéstörténete a piacgazdaságokban". In *Fejezetek a stratégiai menedzsment témaköréből 1996*, Z. Antal-Mokos; K. Balaton; E. Tari; and Gy. Drótos (Eds.). Budapesti Közgazdaságtudományi Egyetem, Budapest, 1-33.

Taylor, B. 1986. "Corporate Planning for the 1990s: The New Frontiers." *Long Range Planning* 19, No.6, 13-18.

Teece, D. J.; G. Pisano; and A. Shuen. 1997. "Dynamic capabilities and Strategic Management." *Strategic Management Journal* 18, No.7, 509-513.

Teece, D. J. and G. Pisano. 1994. "The dynamic capabilities of firms: an introduction." *Industrial and Corporate Change* 3. No.3, 537-556.

INTERMEDIARY ACTIVITIES ON DECENTRALIZED FINANCIAL MARKETS

Dániel Havran and Balázs Árpád Szűcs
Department of Finance
Corvinus University of Budapest
1093, Budapest, Fővám tér 8, Hungary
E-mail: daniel.havran@uni-corvinus.hu

March 31, 2017

KEYWORDS

Decentralized markets, inter-dealer trading, risk allocation, market microstructure

ABSTRACT

Financial intermediary institutions often compete and cooperate with each other at the same time. These financial actors provide services to their investors, and enter into transactions with them. Moreover, these players very often trade with each other to mitigate their market risks related to their exposures against their clients. Decentralized inter-dealer markets differ from the Walrasian textbook markets in three characteristics: transactions are bilateral, market players form a network, market players possess diverse bargaining power. We develop and simulate a single-period model to describe the benefits of those dealers on the network, who not only mitigate the risk, but connect other dealers on these markets.

1 INTRODUCTION

A significant part of the turnover on financial markets is transacted by financial intermediaries. These players often act as market makers on the market formed by their own investors. The investment banks and brokerages, however, are also in contact with each other. Trading between intermediaries has been drawing increasing attention since the bankruptcy of Lehman Brothers, and rightfully so, since the bank was a major player on the market, and its collapse endangered the entire financial system. The counterperty risk induced systemic risk has been significantly diminished since trading was forced through the clearing houses. Trade among actors of the financial market, however, still remains an important issue.

In this paper we examine the behaviour of intermediaries in a situation where they trade not only with their clients, but also with each other. Their goal with the latter is to decrease the risk of their positions that arise through client trading. Such structure can be observed on the interbank loan-deposit and IRS market, as well as on the market of US treasury bonds, municipal bonds, or credit derivatives (especially CDS).

The most important attribute of the inter-dealer market is that finding a counterparty is costly, therefore different players meet each other with possibly different probabilities. It is important to understand how market makers find an appropriate exchange partner under such friction and how they determine the price of the exchanged assets in the bilateral trades.

We investigate the bargaining and trade mechanisms of the decentralized financial markets. Our simulation based analysis extends the theoretical model of Havran and Szűcs (2016). We use simulations to explain the connection between the profits gained from risk mitigation and intermediation, and the costs of accessing and bargaining with other dealers. We set up several core-periphery networks to show that gains from these intermediary activities are highly influenced by the network centrality of the players.

2 LITERATURE

The first theoretical paper to focus on the trading that is driven by a risk mitigation motive was Borch (1962), investigating the mechanisms of the reinsurance market. In his model, risk averse players enter the market to trade away their risky positions. Borch deduces the market equilibrium as well as the market price, but ignores the network aspect. Transaction prices may differ within the network. The question is how the players determine the transaction prices. The mechanism proposed by Borch is one way to allocate risk, but there are alternatives to it, such as in Csóka et al. (2009).

The past few years have seen numerous theoretical papers that explain trading on financial networks. Atkeson et al. (2013) model CDS markets in such a framework. In their model, intermediary banks that enter the market must face a fixed cost, and all banks have counterparty limits. Banks diminish their exposure on the inter-dealer

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

market. When two players trade they must agree on the price and quantity. Zawadowski (2013)) relates to this, examining the systemic risk of over-the-counter markets. Malamud and Rostek (2012) builds a more general model where risk averse players trade on a fixed network, but, instead of searching and bargaining, their players use the network to trade simultaneously. These players have different price impact and liquidity, which they consider during the trading process. The price impact of a player is, however, independent of their endowment. In the equilibrium players maximize their utilities by trading on the network. In the market model of Babus and Kondor (2013) differently informed risk neutral players are bargaining on a fixed network. The authors present the dispersion of information in the network, and also deduce the market equilibrium.

In this paper we apply the search and bargain approach with risk averse players, similarly to Atkeson et al. (2013), Viswanathan and Wang (2004) and Zhong and Kawakami (2016). The logic of our inter-dealer market model, however, also relates to the theory of Malamud and Rostek (2012). Building on our formerly developed one-shot model (Havran and Szűcs, 2016), we extend this approach by adding new structures to the trading network and re-defining the bargaining process.

3 MODEL

Following the foundations of Havran and Szűcs (2016), we employ a certain set of assumptions to describe the daily routines of an over-the-counter market. All trades are made within a single time period. It means that players must decide in advance what other players to approach and how much to trade with them. However, before the trade there is a tatonnement process. The price is determined during the procedure through bilateral bargaining, and it is a function of the transacted quantity. To sum up, in a bilateral relation, each traded asset is subject to the same price, but the price may vary between different counterparties.

Theory

Let us assume there is a single risky asset on the market, with unknown end of day value. The distribution of this value is common knowledge among all the players: $v \sim N\left(\mu, \sigma^2\right)$. Players can also hold a risk free asset that serves as money in the trading process. We define the current wealth of a player as $w \doteq v\,(x+y)+c$, where $x \in \mathbb{R}$ is the amount of risky assets obtained from customers, $y \in \mathbb{R}$ is the quantity traded with fellow intermediaries, while $c \in \mathbb{R}$ is the amount of money payed or received during the trading. There are K rational, risk averse players on the markets. Players have a single period mean-variance utility function:

$$U_i\left(y_i\right) = \mathrm{E}\left(w\left(y_i\right)\right) - \frac{1}{2}\lambda \cdot \operatorname{var}\left(w\left(y_i\right)\right) \qquad (1)$$

The λ risk aversion coefficient may differ among players.

Each i player has a $\gamma_i \in \mathbb{R}^K$ searching preference vector, that shows the ratios in which the player intends to contact the other players (the reciprocal of the searching friction). Contacting another player does not necessarily result in trading. The sum of the weights is unity: $\sum_{j=1}^{K} \gamma_{ij} = 1$. Players do not trade with themselves, therefore: $\gamma_{ii} = 0$. The searching preference matrix consists of the vectors of all the players as follows:

$$\Gamma \equiv \begin{bmatrix} \gamma_1' \\ \gamma_2' \\ \cdots \\ \gamma_K' \end{bmatrix} \equiv \begin{bmatrix} 0 & \gamma_{12} & \cdots & \gamma_{1K} \\ \gamma_{21} & 0 & \cdots & \gamma_{2K} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{K1} & \gamma_{K2} & \cdots & 0 \end{bmatrix} \qquad (2)$$

The equation describing the trading process is the following. Let $\tau \in \mathbb{R}^K$ denote the vector of transactions initiated by a certain market maker. The balance of initiated and incoming transactions must equal the final transaction goal, namely $\tau - \Gamma'\tau = y$. The τ quantity of transactions initiated on the searching network must satisfy the equation below:

$$\begin{pmatrix} I - \Gamma' \\ \underline{1}' \end{pmatrix} \tau = \begin{pmatrix} y \\ 0 \end{pmatrix} \qquad (3)$$

Where I in the matrix on the left side is the identity matrix of order K, while $\underline{1}$ is a column vector with all K elements being ones, and 0 denotes the scalar zero. Using the τ vector, the quantity of the actual transactions between players i and j can be expressed as follows $t_{ij} = \gamma_{ij}\tau_i - \gamma_{ji}\tau_j$.

Players both i and j may gain extra profit via any transaction. The sum of the traded volume conducted by player i is $\sum_j t_{ij} = y_i$. The extra profit they achieve combined is allocated between them through Nash bargaining. Depending on their bargaining power, player j gives d_{ij} amount of money to player i based on the following sharing rule:

$$\begin{aligned} d_{ij} = \theta_{ij} &\left\{\phi_{ij}\left[U_j\left(y_j\right) - U_j\left(0\right)\right]\right\} - \\ &- \left(1-\theta_{ij}\right)\left\{\phi_{ji}\left[U_i\left(y_i\right) - U_i\left(0\right)\right]\right\} \end{aligned} \qquad (4)$$

Where we define $\phi_{ij} \in [0,1]$ as the individual utility increment allocation rule, that shows the ratio of the total utility increment of player i is from the trade with player j . The total contribution of the players to dealer i is $\sum_k \phi_{ik} = 1$. The bargaining power between the two players is denoted by $\theta_{ij} \in (0,1)$. Greater values of θ_{ij} mean larger payoffs to player i assuming non-negative profits. d_{ij} is a signed variable, which means that player i may as well be the one paying to player j.

The money transacted between players i and j is made of two components. First, player i pays qt_{ij} amount to player j, where q is the commonly known *fair value of the asset* on the internal market. Second, player i receives a d_{ij} signed amount from player j depending on their bargaining powers. Therefore the cash flow between players i and j becomes $c_{ij} \doteq -qt_{ij} + d_{ij}$.

Player i intends to trade a net amount of y_i and receive a net (signed) amount of $\sum_{k=1}^{K} c_{ik}$ in return. The utility function of player i considering the transactions is:

$$U_i\left(y_i\right) \doteq \mu\left(x_i+y_i\right)-\frac{1}{2}\lambda_i\sigma^2\left(x_i+y_i\right)^2+\sum_{k=1}^{K} c_{ik}\left(y_i\right) \quad (5)$$

Under a certain q internal market fair value, the net demand of player i is the quantity y_i where the marginal utility of player i is zero, and at the same time maximizes the utility function

$$y_i^d\left(q\right) \doteq \left\{ y_i \left| \frac{\partial u_i}{\partial y_i}\left(y_i,q\right)=0; \frac{\partial^2 u_i}{\partial y_i^2}\left(y_i,q\right)<0 \right.\right\} \quad (6)$$

where $u_i\left(y_i\right)$ is the ith element of the explicit utility vector. The net demand does not specify the trading partners, only the intended amount of transactions for the specific player. On a certain market defined by $(x,\lambda,\Gamma,\theta)$ tuple, the net demand function of player i becomes:

$$y_i^d\left(q\right)=-x_i+\frac{1}{\lambda_i\sigma^2}\mu-\frac{1}{\lambda_i\sigma^2}q \quad (7)$$

By summing up these demands, one can derive the y^* equilibrium allocation for player i , resulting

$$y_i^*=\frac{\frac{1}{\lambda_i\sigma^2}}{\sum_{k=1}^{K}\frac{1}{\lambda_k\sigma^2}}\sum_{k=1}^{K}x_k-x_i \quad (8)$$

The *fair value* that clears the market (sum of the net demands becoming zero) is:

$$q^*=\mu-\frac{\sum_{k=1}^{K}x_k}{\sum_{k=1}^{K}\frac{1}{\lambda_k\sigma^2}} \quad (9)$$

With the help of the market clearing fair value, one can easily express the price of all bilateral transactions as $p_{ij}^*=q-d_{ij}/t_{ij}$.

Examples

We show some basic examples for illustrating how the market model works. Let us assume the following setup:

- Number of players is $K=4$;
- The expected value is $\mu=1$ and the variance of the asset equals to $\sigma^2=1$;
- We suppose the same risk aversion parameter for each dealer, $\lambda=2$;
- We suppose that the bargaining power is $\theta=1/2$ for each player, hence they have equal power in the bilateral bargains;
- We define Γ trade preference matrix as it captures some special network structures: a full graph, star-like network (where player 2 is a middleman), and a circle structure.

Table 1: Initial asset distribution, utilities and equilibrium trade amounts

	x	$U(0)$	y^*
1	-1	-2	1
2	0	0	0
3	0	0	0
4	1	0	-1

Table 1. summarizes the initial distribution of the x_i assets, the initial level of individual $U_i(0)$ utilities, and the net trade amounts in equilibrium. In these settings, the market clearing fair value equals to $q=1$.

Figure 1 illustrates the trade in equilibrium. On the first network, the asset transfers from player 4 to player 1 in many ways: once directly, then through player 2 or player 3 itself, or through both of these actors.

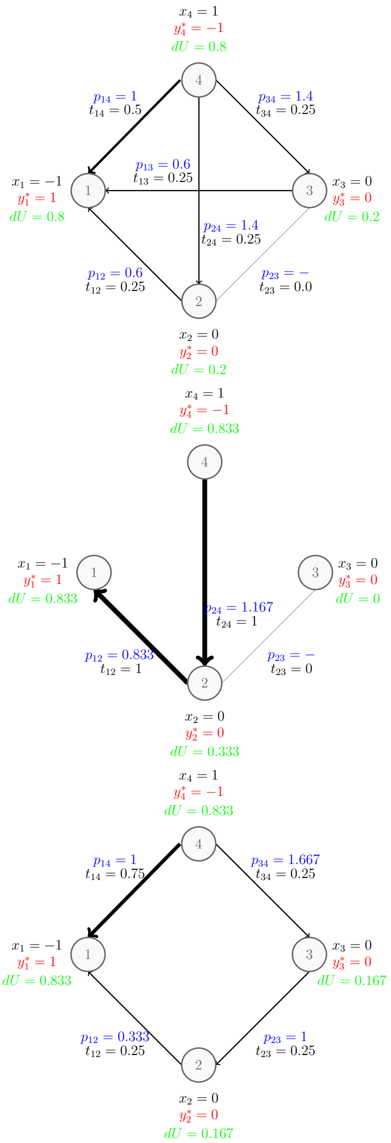


Figure 1: Elementary examples of the markets

Considering the first case, all market players split their trade. Dealers 1 and 4 increase their utilities through risk mitigation, while dealers 2 and 3 attain some benefits by helping them to trade.

The second case illustrates a player who acts a purely intermediary role. Dealer 2 receives the assets from player 4 and she passes it along to player 1. However, dealer 2 does not modify her starting inventory at the end of the day, she gains from connecting to the others. Furthermore, player 3 is able to reach player 2, but they have no motives for trading.

The third case is about a possible long trade chain. Although, dealer 4 is able to connect to dealer 1 directly, she chooses to share some of her trade between players 1 and 3, because of her trade preferences (e.g. counterparty limit considerations). Hence, dealers 2 and 3 realize gains from trading as intermediary players. The literature refers to this phenomenon of multiple exchanges as *hot potato trading*.

Comparing the increments of the end of day utilities, dealers gain different profits according to their initial inventories and positions in the network structure. Players may obtain benefits by diminishing their risks through trading in the opposite direction of their initial positions. However, they have to share these gains in order to motivate the others for cooperation. Distribution of the prices in the bilateral trades depends on the inventories and the network topology as well.

4 SIMULATION METHODOLOGY

In this section we briefly present the improvements to the original model as well as the model settings that we employ. We detail the building blocks of the simulated networks, and expose the concept of the asymmetric bargaining powers.

Network construction

There are many definitions of core-periphery networks. According to a stylized, egde-based definition provided by van der Leij et al. (2016), a core-periphery network satisfies three properties: the core agents form a completely connected clique; there are no links between periphery agents; each core agent is connected to at least one periphery agent and vice versa.

We use a more general approach suggested by Rombach et al. (2012) (on page 9): "a $CP(N,d,p,k)$ network has N nodes, where dN of the nodes are core nodes, $(1-d)N$ of the nodes are peripheral nodes, and $d \in [0,1]$. The edges are assigned independently at random. The edge probabilities for periphery-periphery, core-periphery, and core-core pairs are p, kp, and k^2p, respectively, where $p \in [0,1]$ and $k \in [1,(1/p)^{1/2}]$."

We generate the Γ matrix of our model by elements as $\gamma_{ij} = g_{ij}/\sum_j g_{ij}$; by producing g for all $i > j$ such as

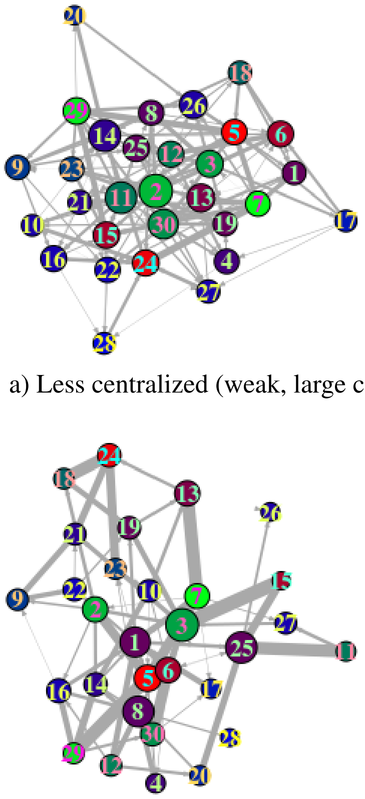
$$g_{ij} = \begin{cases} 1 & \text{if } \xi > \omega_{ij} \\ 0 & \text{otherwise} \end{cases}$$

and for $i \leq j$ $g_{ji} = g_{ij}$, where $\xi \sim U(0,1)$ is a uniformly distributed random variable. Furthermore, ω_{ij} thresholds are defined by

$$\omega_{ij} = \begin{cases} k^2p & \text{if } i \in \mathscr{C} \wedge j \in \mathscr{C} \\ kp & \text{if } i \in \mathscr{C} \wedge j \in \mathscr{P} \\ p & \text{if } i \in \mathscr{P} \wedge j \in \mathscr{P} \end{cases}$$

by using notations $\mathscr{C}$ and $\mathscr{P}$ for the set of nodes in the core, and in the periphery respectively. Thus, the matrix of g_{ij} is symmetric, with zero diagonal and one or zero non-diagonal elements.

For illustration of the characters of the core-periphery networks, we plot graphs of two networks: a core-periphery network with weak and large core ($d = 0.5$, $p = 0.4$, $k = 1.2$), and a core-periphery network with strong and small core ($d = 0.2$, $p = 0.1$, $k = 1.8$). Figure 4 shows the traded amount among the nodes. The larger nodes have more edges. Green nodes possess positive, red nodes possess negative initial positions. Blue nodes have near to zero initial inventories.



a) Less centralized (weak, large core)

b) More centralized (strong, small core)

Figure 2: Trade networks (simulated)

Negotiation technology revised

The bargaining power in the model was exogenously specified in the bilateral bargains. We extend this approach for examining the bargaining effects on the transactional prices and gains. Let us assume that the counterparties play a bargaining game according to the alternating-offers model of Rubinstein (1982), but with certain modifications.

According to this game, the two players are bargaining across several periods of time on the same day. If they cannot make a deal, the counterparty can always make another offer after some $\Delta > 0$ time, and so on. We assume that after Δ units of time player i loses $\psi_i \in \mathbb{R}^+ \in [0,1]$ portion of their payoff, which is the cost of skipping a bargain and entering a new one with the same counterparty. This ψ_i cost of accessing the market can also be interpreted as the impatience attitude of a dealer.

Accordingly, Δ periods of delay discounts the present value of the original payoff of player i by a factor of $(1-\psi_i\Delta)$. Muthoo (2001) shows that, in a subgame perfect equilibrium, the bargaining power of player i against player j in an infinitesimally short period of time (when $\Delta \to 0$) follows:

$$\theta_{ij} = \frac{\psi_j}{\psi_i + \psi_j} \tag{10}$$

This means that in such a bargaining process the player with cheaper access to the market (the more liquid or the more patient one) may claim a larger portion of the overall profit on a certain day.

We suppose furthermore that there are economies of scale on the costs of accessing market. The average cost of entering a bargain is a decreasing function of the number of edges. It represents that a player with many nodes has a cost advantage of accessing its neighbours on the network. The cost is

$$\psi_i(n) = \psi n_i^{\beta-1} \tag{11}$$

for player i, where ψ is the cost of entering into a bargain if the player has only one neighbour, n_i is the number of edges for player i, and $0 < \beta < 1$ is the scale parameter.

We remark that ψ is the same for everyone, thus it does not appear in the θ coefficients.

Setup

We use the following setup for the simulations.

- There are $K = 30$ players on the market;
- Asset value follows $v \sim N(\mu = 1, \sigma^2 = 1)$;
- Homogeneous risk aversion, $\lambda = 2$ for all players ;
- Random initial inventories, $x_i \sim N\left(\mu_x = 1, \sigma_x^2 = 1\right)$
- There are two networks investigated:
  - core-periphery network with weak and large core: $d = 0.5$, $p = 0.4$, $k = 1.2$;
  - core-periphery network with strong and small core: $d = 0.2$, $p = 0.1$, $k = 1.8$;
- No particular preference for trade. We define the Γ trade preference matrix so that a player who has more nodes is willing to trade with all its neighbours in equal amounts. Hence, we determine $\gamma_{ij} = 1/n_i$ for all j.
- We set the individual utility increment allocation rule as $\phi_{ij} = \gamma_{ij}$. It means that the utility increment that player i reaches by trading with player j is proportional to the degrees.
- We use three cases for the scale parameters on the cost of bargaining: $\beta = 0$, 0.5, 1 (strong, weak and no economic of scales).

In total, we have 6 different cases (two networks and three scale parameters). We generate 50 runs for each setting and analyze the results, which seems to be enough for describing the asymptotic behavior of our model.

5 EVALUATION OF THE RESULTS

Inter-dealer prices

First, we investigate the distribution of the transactional prices. Let us define the *average spread* as the difference between the volume weighted average price of buys and the volume weighted average price of sells as the following:

$$s_i = \sum_j \left(\frac{t_{ij}^+}{\sum_j t_{ij}^+} p_{ij} \right) - \sum_j \left(\frac{t_{ij}^-}{\sum_j t_{ij}^-} p_{ij} \right) \tag{12}$$

For identifying the players who are only motivated in the one-sided trade and the players who rather transmit the assets, we introduce the trade balancedness measure. This indicator calculates the relative net trade position to the total amount of individual sells and buys:

$$tb_i = \frac{\left| \left| \sum_j t_{ij}^+ \right| - \left| \sum_j t_{ij}^- \right| \right|}{\left| \sum_j t_{ij}^+ \right| + \left| \sum_j t_{ij}^- \right|} \tag{13}$$

In Figure 3, we explain the average spreads by the trade balancedness indicators. Combining the results of the 50 independent runs of a particular setup, we plot the average spreads on the less and more centralized networks. Each average spread indicator belongs to a player in a particular run. *Beta 0* indicates that the bargaining power is proportional to the number of neighbours on the network. *Beta 1* cases describe the situations where the bargaining powers are symmetric. On the legend, *c.btw* stands for betweenness centrality measure and the colors of the plots indicate the individual betweennes centrality of the actors.

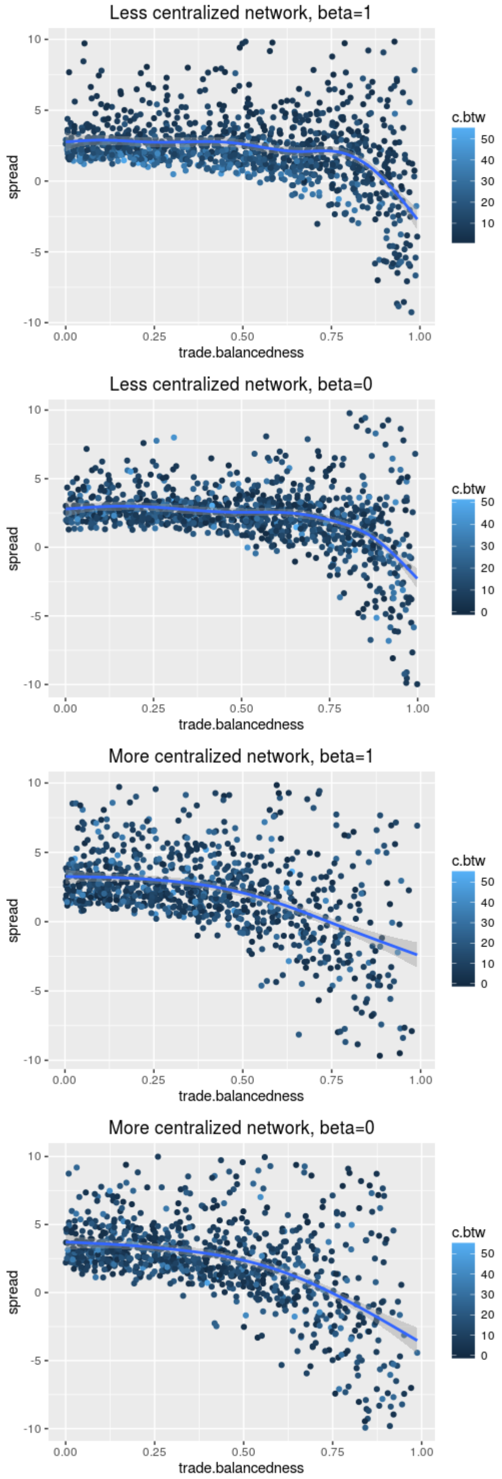


Figure 3: Spreads and trade balancedness

Considering the distribution of the average spreads, one can spot some interesting differences among the effects caused by the network structures.

First, on the less centralized network the core is larger (half of the players) and the core-to-core interlinkage is less dense. The *average spread* is positive even for some of the players who have unbalanced trade positions. The dispersion of the average spread is higher when the unbalancedness is higher. These actors usually possess large positions in one direction of the trade. Sometimes their spread goes negative, which suggests that they are willing to pay more for trading in a given direction. We remark that these players are able to cover this trade from their benefits from the risk mitigation.

Second, on the more centralized network, the core members are able to enforce more favorable transactional prices, thus they can calculate with higher average spread. Unbalanced trade positions are punished intensely by the market, because the periphery players have less linkage to look for an appropriate counterparty.

We capture the centrality positions of the players by the degree centrality measures, and use a simple colour scheme to identify the network positions. The centrality does not play a crucial role in the spreads, they are rather determined by the initial positions of the actors and their neighbours.

Profit distributions

Second, we focus on the benefits from the risk mitigating actions and the gains from the dealer-to-dealer intermediary businesses. To separate these benefits from each other, we introduce a benchmark theoretic utility level that lacks the benefits or losses of the bargains. We define the *total utility of trading at fair price* as the utility of player i if the agent performs all of her transactions at price q.

$$\bar{U}_i\left(y_i\right) \doteq \mu\left(x_i+y_i\right)-\frac{1}{2}\lambda_i\sigma^2\left(x_i+y_i\right)^2-qy_i \tag{14}$$

The $U_i(y_i)-U_i(0)$ difference shows the gains from the sum of the activities, and $U_i(y_i)-\bar{U}_i(y)$ indicates the bargaining gains or costs related to the intermediation. The difference $\bar{U}_i(y_i)-U_i(0)$ refers to the gross gain of entering the market with non-zero initial position (in other terms, the gross gain of the hedge). Although, the gross gain is important to understand the motives of the players, they sometimes fail to reach this theoretic level, because of the incurring costs of the exchange. Hence, we apply the concept of the pure (or net) benefits of the hedge and the pure benefits of the intermediary activities. The net benefits of the hedge equal to

$$B_i^{Hedge}=\frac{1}{2}(\bar{U}_i(y_i)+U_i(0))-U_i(0) \tag{15}$$

In other terms, it is the non-negative gain that a player surely realizes by entering the market. We remark that dealers with non-zero initial assets have some monopolistic power in this model. The net benefits of risk mitigation do not depend directly on the network structure, rather on the initial asset positions. We define the net benefits of the intermediation as

$$B_i^{Inter}=U_i(y_i)-\frac{1}{2}(\bar{U}_i(y_i)+U_i(0)) \tag{16}$$

which is the non-negative gain that a player can realize by its transmission provided by its own position in the dealer network.

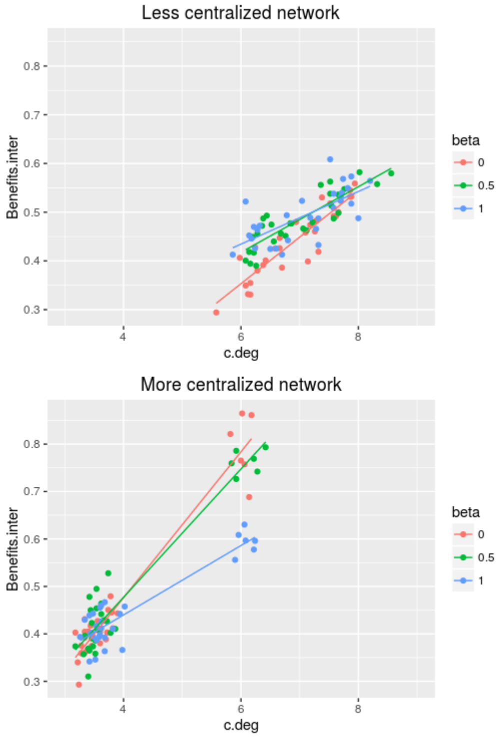


Figure 4: Gains from the intermediary activities

Figure 4 shows the net benefits from passing through the assets by the players' degree centrality. We calculated the average gain over the 50 independent simulations, and the average degree centralities. Hence, we are able to draw 30 points (players) for each setup by their degree centrality (c.deg). Three settings of the less centralized and three settings of the more centralized networks are presented on the two subplots. The most obvious implication of the model is that the core members acquire more benefits from the trade if the core in the network is heavily connected and the number of core members is relatively low. Moreover, different costs of bargaining imply further diversity of gains, this effect is higher on the more centralized markets.

6 SUMMARY

We explored the relationship between the possible benefits of trading on decentralized markets and two attributes of the traders: status in the network and the initial asset position. Based on a theory of Havran and Szűcs (2016), we constructed an extended model of the decentralized markets, which is able to represent the core-periphery network structure, and offers a bargaining mechanism describing asymmetric relationships in bilateral deals.

REFERENCES

Atkeson, A., Eisfeldt, A. L., and Weill, P.-O. (2013). The Market for OTC Derivatives. CEPR Discussion Papers 9403.

Babus, A. and Kondor, P. (2013). Trading and information diffusion in OTC markets. CEPR Discussion Papers 9271.

Borch, K. (1962). Equilibrium in a Reinsurance Market. *Econometrica*, 30(3):pp. 424–444.

Csóka, P., Herings, P. J.-J., and Kóczy, L. Á. (2009). Stable allocations of risk. *Games and Economic Behavior*, 67(1):266–276.

Havran, D. and Szűcs, B. Á. (2016). Árjegyzői viselkedés belső kockázatelosztás mellett (Market Maker Behavior with Hedging on Inter-dealer Markets). *Szigma*, 47(1-2):1–30.

Malamud, S. and Rostek, M. (2012). Decentralized Exchange. Working Papers 12-18, NET Institute.

Muthoo, A. (2001). The Economics of Bargaining. In *Knowledge for Sustainable Development: An Insight into the Encyclopedia of Life Support Systems*. EOLSS Publishers Co. Ltd.

Rombach, M. P., Porter, M. A., Fowler, J. H., and Mucha, P. J. (2012). Core-Periphery Structure in Networks. *CoRR*, abs/1202.2684.

Rubinstein, A. (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica*, 50(1):97–109.

van der Leij, M., Veld, D. i. t., and Hommes, C. (2016). The Formation of a Core Periphery Structure in Heterogeneous Financial Networks. Tinbergen Institute Discussion Papers 14-098/II, Tinbergen Institute.

Viswanathan, S. and Wang, J. J. D. (2004). Inter-Dealer Trading in Financial Markets. *The Journal of Business*, 77(4):987–1040.

Zawadowski, A. (2013). Entangled Financial Systems. *Review of Financial Studies*, 26(5):1291–1323.

Zhong, Z. and Kawakami, K. (2016). The Risk Sharing Benefit versus the Collateral Cost: The Formation of the Inter-Dealer Network in Over-the-Counter Trading. (822).

AUTHOR BIOGRAPHIES

DÁNIEL HAVRAN, PhD is an Associate Professor of Finance at Corvinus University of Budapest. His research interests cover Market Microstructure, Corporate Finance and Credit Risk Management. E-mail address: `daniel.havran@uni-corvinus.hu`.

BALÁZS ÁRPÁD SZŰCS, PhD is an Adjunct Professor at Corvinus University of Budapest. His research topics include Market Microstructure, and Intraday Forecasting of Stock Volumes. E-mail: `balazsarpad.szucs@uni-corvinus.hu`

INDEXED BONDS WITH MEAN-REVERTING RISK FACTORS

Attila A. Víg

Ágnes Vidovics-Dancs, PhD, CIIA

Department of Finance
Corvinus University of Budapest
H-1093, Fővám tér 8, Budapest, Hungary
E-mail: attila.vig@uni-corvinus.hu

KEYWORDS

Inflation-indexed bond, Monte Carlo simulation, risk neutral pricing, mean-reverting stochastic process

ABSTRACT

In this paper, we focus on the value of inflation-indexed bonds in an extended short rate model, which is a specific case of the general framework provided by Jarrow and Yildirim (2003). In the model, we assume mean-reverting stochastic dynamics under the risk neutral measure for both the short interest rate and the instantaneous inflation rate. We define the zero-coupon inflation-indexed bond, and first estimate its value by Monte Carlo simulation, then deduce an analytical formula as well. We briefly touch on the yield and inflation curves the model is able to produce.

INTRODUCTION

Standard nominal bonds are considered one of the simplest financial products, serving as building blocks for complex ones as well. While government bonds are usually considered riskless – apart from default risk which we ignore in this paper –, it is important to note that this property means no risk in the future nominal cash-flow, not in the present value of the security itself.

The main subject of our examination is a bond with a similar idea: the inflation-indexed bond. This security can also be considered riskless, but in a different sense: instead of the nominal values, the real values of future cash flows are fixed. To achieve this, the face value of an inflation-indexed bond is adjusted according to a price index, and the coupon and principal payments are based on this adjusted amount.

An example for the cash flows of nominal and inflation-indexed bonds is given in Table 1. While the future cash flows of the nominal bond are deterministic, those of the indexed bond are random variables, since we do not know the inflation rates of in advance.

However, if we care more about the purchasing power rather than the exact dollar figures, the inflation-indexed bond suddenly becomes deterministic, while the future cash flows of the nominal bond can be considered random variables. The indexed bond described in Table 1 will provide the holder with (5,5,105) units of the basket of goods that is used for calculating the price index.

	0	1	2	3	
inflation		2%	4%	3%	
notional	100.0	100.0	100.0	100.0	nominal bond
cash flow		5.0	5.0	105.0	
notional	100.0	102.0	106.1	109.3	indexed bond
cash flow		5.1	5.3	114.7	

Table 1: Cash flows of standard nominal and inflation-indexed bonds. Coupon rate: 5%. Frequency: annual.

An example – and perhaps the most important one in terms of gross market value – of such an indexed bond is the Treasury Inflation-Protected Security [TIPS] issued by the Treasury of the United States. Issued at maturities of 5, 10, and 30 years, adjustment of the principal amount is based on the Consumer Price Index [CPI]. Real coupon rates are fixed, and are paid semi-annually. TIPS take up more than 6% of total public debt of the United States (United States Department of the Treasury, 2017).

The main holders of TIPS are pension funds, who are especially motivated to hedge against inflation risk. Pension funds' future obligations are tied to inflation, so by investing in TIPS, inflation risk will be present on both sides of their balance sheet, thus cancelling each other. For more information about inflation-indexed securities in general see Deacon et al. (2004). Affine models for inflation-indexed derivatives and related securities are discussed thoroughly by Ho, Huang and Yildirim (2014). Our research follows the path of Mercurio (2005), who considers a specific case of such affine models.

The rest of the paper is organized as follows. First we introduce two basic financial instruments and describe

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

the stochastic model within which we can price them. Afterwards, we detail a simple but computationally expensive estimation of this price by Monte Carlo simulation. Then, we deduce an analytical formula for this price by applying stochastic calculus and pricing theory. Finally, we investigate the yield and inflation curves the model is able to produce, with some emphasis on the model's most prominent parameter.

For detailed analysis of the mathematical background of our paper see Cairns (2004), Medvegyev (2007), Medvegyev and Száz (2010).

THE MODEL

We assume both the short interest rate and the instantaneous rate of inflation to follow an Ornstein-Uhlenbeck process under the risk neutral measure (Q), described by the following stochastic differential equations:

$$\begin{aligned} dr(t) &= \alpha_r[\bar{r} - r(t)]dt + \sigma_r dW_r^Q(t) \\ di(t) &= \alpha_i[\bar{\iota} - i(t)]dt + \sigma_i dW_i^Q(t) \end{aligned} \tag{1}$$

The process for the short interest rate is of course the classic mean-reverting process of the Vasicek model (Vasicek, 1977). Mean reversion is present in most short interest rate models, and it seems reasonable to assume mean reversion for the inflation rate as well.

Intuition behind the parameters is the following: $\bar{r}$ and $\bar{\iota}$ are the long-term means of the processes. Whenever the process is below (above) its long-term mean, the drift becomes positive (negative), thus tending towards the long-term mean. The factors $\alpha_{\{r,i\}}$ represent the strength of this mean-reverting property, and $\sigma_{\{r,i\}}$ are the volatility parameters.

There are two Wiener processes present on this market, so we must define a correlation parameter through their cross quadratic variation:

$$d\langle W_r^Q(t), W_i^Q(t)\rangle = \rho dt$$

It is important to note that neither of these processes have independent increments: their increments are dependent on the current value of the processes themselves, which stems from their essential mean-reverting property. Thus, the parameter ρ does not perfectly describe the correlation of the increments of the two processes, but only the correlation of the stochastic shocks represented by the Wiener-processes.

We assume two financial products to be driven by these processes: the bank deposit $B(t)$ and the price index $I(t)$:

$$\begin{aligned} dB(t) &= r(t)B(t)dt \\ dI(t) &= i(t)I(t)dt \end{aligned}$$

The solution of these differential equations takes the well-known form for $t < T$:

$$\begin{aligned} B(T) &= B(t)e^{\int_t^T r(s)ds} \\ I(T) &= I(t)e^{\int_t^T i(s)ds} \end{aligned} \tag{2}$$

We are investigating the zero-coupon inflation indexed bond [ZCIIB]. This theoretical financial security – together with the standard zero-coupon bond – is going to define the term structure of inflation rates (and interest rates).

The ZCIIB has the following parameters: it is issued at time T_0, it matures at time T, and its single payoff at maturity is $I(T)/I(T_0)$, thus the payoff is indexed by the increase of the price index during the bond's term. We are looking for the value of the ZCIIB at time t, where naturally $T_0 \le t \le T$. Notation: $P_i(t, T_0, T)$.

It should be noted here that while ZCIIBs are mostly theoretical in nature, a real world inflation-indexed coupon bond like the TIPS described in the introduction can be thought of as a portfolio of ZCIIBs with maturities that coincide with the coupon and notional payments of the coupon bond.

Since the model is a simple extension of the Vasicek model to include inflation-indexed bonds, it does not lose standard nominal bonds either. Thus, we also define the zero-coupon bond [ZCB]: a security with a single payoff of 1 (notional) at maturity T. The notation for the value of this security at time $t \le T$: $P_r(t, T)$.

NUMERICAL APPROACH

First, we will calculate the value of a ZCIIB (issued at $T_0 = 0$) at time $t = 0$ by Monte-Carlo simulation.

Without violating generality, we can assume $B(0) = I(0) = 1$. The value of a ZCIIB maturing at time T is thus given by the pricing formula:

$$P_i(0,0,T) = \mathbb{E}_Q\left(\frac{I(T)}{B(T)}\middle|\mathcal{F}(0)\right) \tag{3}$$

Since the stochastic processes detailed above are given under the martingale measure, we can estimate this expectation by Monte Carlo simulation directly.

The two processes are correlated in their stochastic shocks, but most software can only generate independent pseudo-random variables, so we will use Cholesky-decomposition to exchange $W_i^Q(t)$ of equation 1 for two independent Wiener processes:

$$\begin{aligned} di(t) = \alpha_i[\bar{\iota} - i(t)]dt \\ + \sigma_i\left(\rho dW_r^Q(t) + \sqrt{1-\rho^2}d\widetilde{W}_r^Q(t)\right) \end{aligned}$$

where $W_r^Q(t)$ and $\widetilde{W}_r^Q(t)$ are independent.

A given pair of trajectories for the short interest rate and the instantaneous inflation rate will be simulated recursively as a first order approximation of the dynamics given in equation 1:

$$\begin{aligned} r(t_{n+1}) &= r(t_n) + \alpha_r[\bar{r} - r(t_n)]\Delta t + \sigma_r\sqrt{\Delta t}Z_{n+1} \\ i(t_{n+1}) &= i(t_n) + \alpha_i[\bar{\iota} - i(t_n)]\Delta t \\ &\quad + \sigma_i\sqrt{\Delta t}\left(\rho Z_{n+1} + \sqrt{1-\rho^2}\tilde{Z}_{n+1}\right) \end{aligned}$$

where $0 = t_0 < \cdots < t_N = T$ is an equidistant partition of the interval $[0, T]$, $\Delta t = t_{n+1} - t_n$ is the length of a subinterval, and $\{\{Z_n\}_{n=1}^N, \{\tilde{Z}_n\}_{n=1}^N\}$ are independent standard normal random variables.

Representing a single event from the sample space, Figure 1 shows generated trajectories for the short interest rate and the instantaneous inflation rate. We chose $T = 10$ and $N = 1000$ (thus $\Delta t = 0.01$) for the simulation. The strong negative correlation parameter can be traced visually: while both processes fluctuate around their long-term mean, they tend to move in the opposite direction.

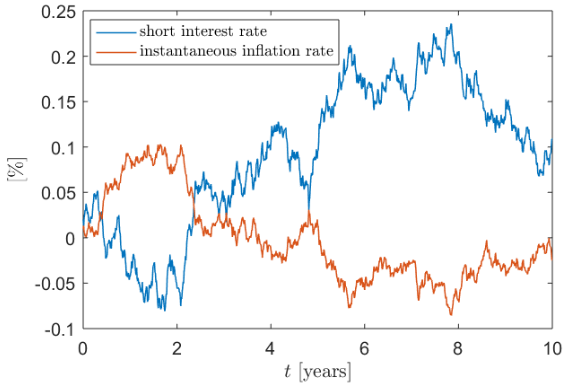


Figure 1: Generated trajectories for the interest rate and inflation rate. Parameters: $\alpha_r = 0.4$, $\alpha_i = 0.4$, $\bar{r} = 0.06$, $\bar{\iota} = 0.04$, $\sigma_r = 0.06$, $\sigma_i = 0.04$, $r(0) = 0.02$, $i(0) = 0.01$, $\rho = -0.9$.

Given the trajectories for the interest rate and the inflation rate, the value processes for the bank deposit and the price index can be calculated by linear approximation of the dynamics given in equation 2:

$$B(t_{n+1}) = B(t_n)(1 + \mathrm{r}(t_n)\Delta t)$$
$$I(t_{n+1}) = I(t_n)(1 + i(t_n)\Delta t)$$

where $B(0) = I(0) = 1$.

While the processes $r(t)$ and $i(t)$ are of unbounded variation, the value processes $B(t)$ and $I(t)$ are defined by the integral of these, thus becoming of bounded variation. This “smoothness” property can be seen in Figure 2.

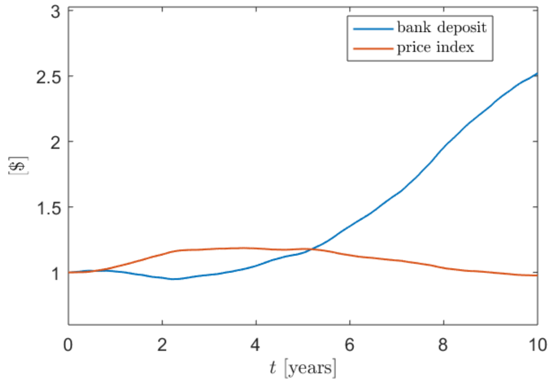


Figure 2: Value processes for the bank deposit and the price index.

The processes of Figure 2 correspond to the instantaneous rate processes in Figure 1: compounding the interest (inflation) rates of Figure 1 will result a bank deposit (price index) seen in Figure 2.

Finally, we calculate the ratio of these value processes, thus receiving a single observation for $I(T)/B(T)$. It is important to note that a single observation is a trajectory itself, since we calculate the ratio for $\forall \{t_n\}_{n=1}^{N} \in [0, T]$. We produce a multitude of such observations in order to get an estimate for the expectation of equation 3.

$$\hat{P}_i(0,0,T) = \frac{\sum_{m=1}^{M} \frac{I_m(T)}{B_m(T)}}{M}$$

We chose a sample size of $M = 10{,}000$ for this simulation. Figure 3 shows the 95% confidence interval of the empirical mean, and it also shows the exact theoretical value, which was calculated by using the closed formula detailed in the next section.

It is important to note the computational costliness of the Monte Carlo approach: a single pair of trajectories required $2 \times N = 2{,}000$ standard normal random variables, and we generated $M = 10{,}000$ such pairs, thus requiring 20 million in total.

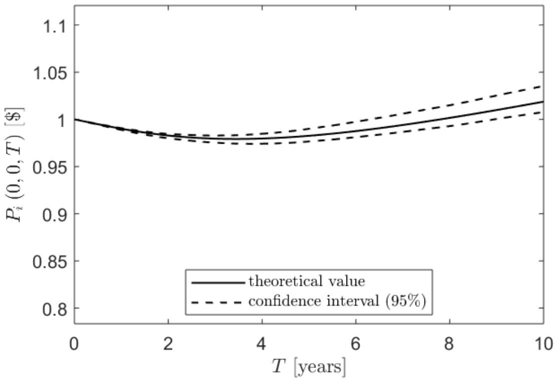


Figure 3: Price of a ZCIIB issued today for maturities $T \in [0,10]$. Parameters used are the same as in Figure 1.

An important property of the ZCIIB can be traced in Figure 3: the value of such a bond does not necessarily have to be decreasing in maturity. While ZCBs’ simple payoff of 1 at maturity means that their price does have to be decreasing in maturity almost by definition because of the time value of money, ZCIIBs’ payoff is indexed by inflation, thus a longer bond can be worth more than a shorter one if market participants expect high inflation in the future.

CLOSED FORMULA

In this section, we will deduce a closed formula for the price of the ZCIIB. The value of a security at time t with the payoff $I(T)/I(T_0)$ is given by the pricing formula:

$$P_i(t, T_0, T) = B(t)\mathbb{E}_Q\left(\frac{I(T)/I(T_0)}{B(T)}\middle|\mathcal{F}(t)\right)$$

After substituting the solutions for $B(T)$ and $I(T)$ given in equation 2, and taking the $\mathcal{F}(t)$-measurable factors out of the expectation, we arrive at:

$$P_i(t, T_0, T) = \frac{I(t)}{I(T_0)}\mathbb{E}_Q\left(e^{\int_t^T r(s) - i(s) ds}\middle|\mathcal{F}(t)\right)$$

The value of this conditional expectation could be calculated directly, but we follow another route. Since both $r(t)$ and $i(t)$ are Markov processes, the condition $\mathcal{F}(t)$ is equivalent to the condition $\{r(t), i(t)\}$. The conditional expectation thus becomes a function with the following arguments:

$$P_i(t, T_0, T) = \frac{I(t)}{I(T_0)} V(r(t), i(t), t, T) \tag{4}$$

Since the Q martingale measure is the specific measure under which any value process divided by the bank deposit (our numeraire) becomes a martingale, $P_i(t, T_0, T)/B(t)$ must also be a martingale, thus its drift must equal 0. We calculate this drift by applying Ito's lemma, and arrive at a partial differential equation for $V(r(t), i(t), t, T)$. We omit the arguments for $V(r(t), i(t), t, T)$, $r(t)$, $i(t)$, and indicate partial derivatives in the subscript:

$$\begin{aligned} V_t + V_r\alpha_r(\bar{r} - r) + V_i\alpha_i(\bar{\imath} - i) + V_{ri}\sigma_r\sigma_i\rho \\ + \frac{1}{2}V_{rr}\sigma_r^2 + \frac{1}{2}V_{ii}\sigma_i^2 + Vi - Vr = 0 \\ V(r(T), i(T), T, T) = 0 \end{aligned} \tag{5}$$

The boundary condition arises from the ZCIIB's payoff at maturity: it pays $I(T)/I(T_0)$ at maturity, thus the bond's value at time T must also equal this amount, which means from equation 4 we get:

$$\frac{I(T)}{I(T_0)} = P_i(T, T_0, T) = \frac{I(T)}{I(T_0)} V(r(T), i(T), T, T)$$

which implies the boundary condition of equation 5: $V(r(T), i(T), T, T) = 1$.

We are going to solve the PDE of equation 5 by assuming the solution to be of a specific – so-called affine – form:

$$V(r(t), i(t), t, T) = e^{A(t,T) - C(t,T)r(t) + D(t,T)i(t)} \tag{6}$$

After substituting this solution form into equation 5, the boundary condition becomes:

$$A(T, T) - C(T, T)r(T) + D(T, T)i(T) = 0$$

We are looking for a solution for $\forall \{r(T), i(T)\} \in \mathbb{R} \times \mathbb{R}$. When $r(T) = i(T) = 0$, the boundary condition becomes $A(T, T) = 0$. When $r(T) \neq i(T) = 0$, we get $C(T, T) = 0$; and similarly, when $i(T) \neq r(T) = 0$, we get $D(T, T) = 0$. Since the boundary condition has to hold true for $\forall \{r(T), i(T)\}$, it has fallen apart into three separate boundary conditions:

$$\begin{aligned} A(T, T) = 0 \\ C(T, T) = 0 \\ D(T, T) = 0 \end{aligned}$$

After substituting the solution form of equation 6 into the PDE of equation 5 itself, we can apply the same reasoning, which results in it falling apart into three separate PDEs as well:

$$\begin{aligned} A_t - \alpha_r\bar{r}C + \alpha_i\bar{\imath}D - \sigma_r\sigma_i\rho CD + \frac{1}{2}\sigma_r^2C^2 + \frac{1}{2}\sigma_i^2D^2 = 0 \\ -C_t + \alpha_rC - 1 = 0 \\ D_t - \alpha_iD + 1 = 0 \end{aligned}$$

The solutions for $C(t, T)$ and $D(t, T)$ are the following:

$$\begin{aligned} C(t, T) = \frac{1 - e^{-\alpha_r(T-t)}}{\alpha_r} \\ D(t, T) = \frac{1 - e^{-\alpha_i(T-t)}}{\alpha_i} \end{aligned} \tag{7}$$

From here, the solution for $A(t, T)$ can be derived by simple integration. We will omit the arguments of $A(t, T)$, $C(t, T)$ and $D(t, T)$ in the result:

$$\begin{aligned} A = C\left[\bar{r} + \frac{\sigma_r\sigma_i\rho}{\alpha_r\alpha_i}\left(1 - \frac{\alpha_r}{\alpha_r + \alpha_i}\right) - \left(\frac{\sigma_r}{2\alpha_r}\right)^2 (\alpha_rC + 2)\right] \\ + D\left[-\bar{\imath} + \frac{\sigma_r\sigma_i\rho}{\alpha_r\alpha_i}\left(1 - \frac{\alpha_i}{\alpha_r + \alpha_i}\right) - \left(\frac{\sigma_i}{2\alpha_i}\right)^2 (\alpha_iD + 2)\right] \\ + (T - t)\left[\bar{\imath} - \bar{r} - \frac{\sigma_r\sigma_i\rho}{\alpha_r\alpha_i} + \frac{\left(\frac{\sigma_r}{\alpha_r}\right)^2 + \left(\frac{\sigma_i}{\alpha_i}\right)^2}{2}\right] \\ + CD\frac{\sigma_r\sigma_i\rho}{\alpha_r + \alpha_i} \end{aligned}$$

Thus, the value of a ZCIIB (issued at T_0 with maturity date T) at time t is:

$$P_i(t, T_0, T) = \frac{I(t)}{I(T_0)} e^{A(t,T) - C(t,T)r(t) + D(t,T)i(t)}$$

where the functions $A(t, T)$, $C(t, T)$ and $D(t, T)$ are defined above.

Figure 3 shows results of this closed formula for the current price of ZCIIBs for different maturities issued today: $P_i(0, 0, T)$, $T \in [0,10]$. The theoretical values fall well inside the 95% confidence interval of the Monte Carlo simulation.

A logical control point for our result arises here: if we choose the parameters in a way that eliminates the effect of inflation, we should get the value of the standard nominal zero-coupon bond $P_r(t, T)$ of the Vasicek model.

To achieve this, first we set the price index at the issue date and current time to unity: $I(T_0) = I(t) = 1$. Next, we eliminate the volatility of the inflation rate: $\sigma_i = 0$. Finally, we make sure the inflation rate starts at zero and stays at zero by setting: $i(t) = \bar{\imath} = 0$. With these parameters, our result reduces to:

$$P_r(t, T) = P_i(t, T_0, T) = e^{A(t,T) - C(t,T)r(t)}$$

where $A(t, T)$ takes the reduced form:

$$\begin{aligned} A(t, T) = C(t, T)\left[\bar{r} - \left(\frac{\sigma_r}{2\alpha_r}\right)^2 (\alpha_rC(t, T) + 2)\right] \\ + (T - t)\left(\frac{\sigma_r^2}{2\alpha_r^2} - \bar{r}\right) \end{aligned}$$

and $C(t, T)$ is given in equation 7. This result does concur with the price of the zero-coupon bond of the Vasicek model.

TERM STRUCTURE OF INFLATION RATES

The yield curve of a given issuer – typically a government – shows the current market conditions for annual yields at which the issuer could refinance its debt at different maturities. There is a one-to-one correspondence between the yield curve and the price of ZCBs, thus the formula of the Vasicek model for $P_r(0,T)$ produces a yield curve:

$$y_r(T) = \frac{1/P_r(0,T)}{T}$$

where $y_r(T)$ is the nominal yield for maturity T. The yield curve typically – although not necessarily – has an upward sloping shape, which the Vasicek model is able to reproduce, as seen in Figure 4.

We go one step further: since our model includes both nominal and inflation-indexed zero-coupon bonds, we will be able to deduce the market-projected annual inflation rate (Dodgson and Kainth, 2006), which we will simply refer to as inflation curve:

$$y_i(T) = \frac{\log \frac{P_i(0,0,T)}{P_r(0,T)}}{T} = y_r(T) - \frac{1/P_i(0,0,T)}{T}$$

where $y_i(T)$ shows the annual inflation rate which is projected by the market for maturity T.

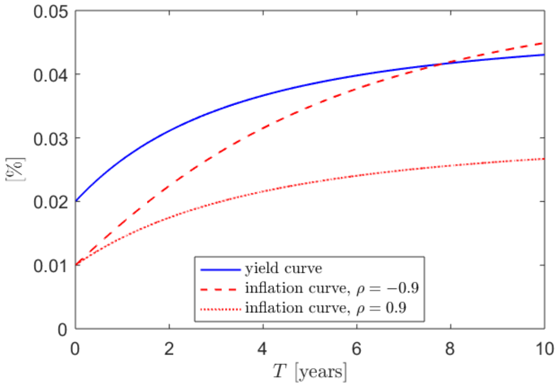


Figure 4: Possible yield and inflation curves described by the model. Parameters are the same as in Figure 1, apart from the correlation parameter as noted in the legend. Note that the mean reversion level of the interest rate process is well above that of the inflation rate process.

Figure 4 shows two inflation curves defined by the model, which differ only in their correlation parameter. Under typical market conditions we would expect the inflation curve to be below the yield curve, since investors usually demand a positive ex ante real yield from their investments, including their nominal bonds. While this does hold true for the case of strong positive correlation parameter, the strong negative correlation parameter makes the inflation curve steeper, even going above the yield curve for longer maturities.

This implies a negative connection between the value of the zero-coupon inflation-indexed bond and the correlation parameter of the stochastic processes for the short interest rate and the instantaneous inflation rate. The exact mechanics of this connection offers ground for further research.

CONCLUSION

In this paper, we presented the main features of inflation-indexed bonds. We defined the zero-coupon inflation-indexed bond, which can serve as building blocks for coupon bonds as well. We set up a mean-reverting stochastic model for the short interest rate and the instantaneous inflation rate. We simulated pairs of trajectories for these processes, and estimated the value of the zero-coupon inflation-indexed bond by Monte Carlo simulation. Then we presented an analytical solution for this problem as well, which corresponded with the results of the simulation. Finally, we used the analytical formula to graph typical inflation curves, and discussed the possible effect of the model's correlation parameter.

REFERENCES

Cairns, A. 2004. *Interest Rate Models: An Introduction*. Princeton University Press. Princeton-Oxford.

Deacon, M., Derry, A. and Mirfendereski, D. 2004. *Inflation-Indexed Securities: Bonds, Swaps and Other Derivatives*. John Wiley & Sons, Chicester. 2nd edition.

Dodgson, M., Kainth, D. 2006. "Inflation-linked derivatives". *Royal Bank of Scotland Risk Training Course, Market Risk Group*.

Ho, H.W., Huang, H.H. and Yildirim, Y. 2014. "Affine model of inflation-indexed derivatives and inflation risk premium". *European Journal of Operational Research* 235, No. 1, 159-169.

Jarrow, R. and Yildirim, Y. 2003. "Pricing Treasury Inflation Protected Securities and Related Derivatives using an HJM Model". *Journal of Financial and Quantitative Analysis* 38, No. 02, 337-358.

Medvegyev, P. 2007. *Stochastic Integration Theory*. Oxford University Press, Oxford-New York.

Medvegyev, P. and Száz, J. 2010. *A meglepetések jellege a pénzügyi piacokon*. Nemzetközi Bankárképző Központ, Budapest.

Mercurio, F. 2005. "Pricing inflation-indexed derivatives". *Quantitative Finance* 5, No. 3, 289-302.

United States Department of the Treasury. 2007. "Monthly Statement of the Public Dept of the United States". Accessed February 9, 2017. https://www.treasurydirect.gov

Vasicek, O. 1977. "An equilibrium characterization of the term structure". *Journal of Financial Economics*, No. 5, 177-188.

AUTHOR BIOGRAPHIES

Attila András VÍG is a PhD student at the Department of Finance at Corvinus University of Budapest. His main research area is pricing theory of financial derivatives, with emphasis on interest rate derivatives. He received his bachelor's degree in finance from Corvinus University of Budapest, and his master's degree in financial mathematics from Eötvös Loránd University. His e-mail address is: attila.vig@uni-corvinus.hu

Ágnes VIDOVICS-DANCS, PhD, CIIA is adjunct professor at the Department of Finance at Corvinus University of Budapest. Her main research areas are government debt management in general and especially sovereign crises and defaults. She worked as a junior risk manager in the Hungarian Government Debt Management Agency in 2005-2006. Currently she is risk manager of a Hungarian asset management company. Her e-mail address is: agnes.dancs@uni-corvinus.hu

STRESS TEST MODELLING OF PD RISK PARAMETER UNDER ADVANCED IRB

Zoltán Pollák
Department of Finance
Corvinus University of Budapest
1093, Budapest, Hungary
E-mail: ZPollak@bankarkepzo.hu

Dávid Popper
International Training Center
for Bankers (ITCB)
1011, Budapest, Hungary
E-mail: DPopper@bankarkepzo.hu

KEYWORDS
Stress test, Default rate modelling, Migration matrix, Probability of Default, IRB.

ABSTRACT

The 2008 crisis highlighted the importance of using stress tests in banking practice. The role of these stress tests is to identify and precisely estimate the effect of possible future changes in economic conditions on the capital adequacy and profitability of banks. This paper seeks to show a possible methodology to calculate the stressed point-in-time PD parameter. The presented approach contains a linear autoregressive distributed lag model to determine the connection between the logit of default rates and the relevant macroeconomic factors, and uses migration matrices to calculate PDs from the forecasted default rates. The authors illustrate the applications of this methodology using real credit portfolio data.

INTRODUCTION

The European Banking Authority (EBA) requires banks under Advanced Internal Rating-Based (AIRB) approach to prudently measure their risk profile and apply more accurate risk management tools.
According to Basel Capital Requirements Regulation (CRR), institutions “shall regularly perform a credit risk stress test to assess the effect of certain specific conditions on its total capital requirements for credit risk” (575/2013/EU, Article 177(2)).
Banks should have a comprehensive stress testing framework, which is based on the forecasting of the capital adequacy, balance sheet and the P&L statement along different stress scenarios. As a part of this framework, credit institutions should estimate stressed PD parameters that serve as important input factors for the calculation of the bank’s performance under the defined scenarios.
One of the ways to forecast PD parameters are econometric models with default rate time series as dependent variable. In this methodology, the estimated stressed default rates are transformed into stressed segment PDs using migration matrices.
The main object of this paper is to reveal connections between various macroeconomic variables and the default rate using econometric methodology on real data, and estimate the stressed PD parameters based on the results. First, we present the used database and the chosen methodology, then we show the results of the default rate model. In the last part, we will transform the stressed default rates into stressed PDs.

INPUT DATA

Default rates

In our presented case, stress test modelling of default rates was carried out based on quarterly default rate time series. The portfolio segment used for modelling is the micro segment (micro-enterprises of a commercial bank’s credit portfolio).
For the stress test modelling, yearly default rates were calculated for each quarter during the observation period (from Q1 2007 to Q1 2016).

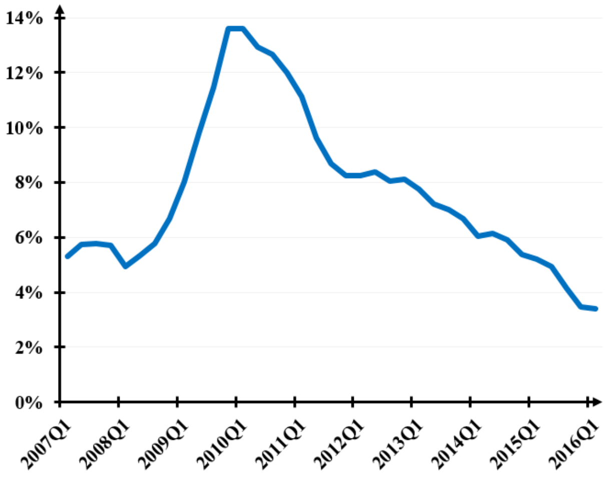


Figure 1: Default rate time series

Figure 1 illustrates the default rate time series for the mentioned micro segment.
To forecast default rates along the stress scenarios, we used different macroeconomic variables. Most of the selected macro variables are available on a quarterly basis, therefore we chose the quarterly frequency for the calculations.
Instead of modelling the default rates directly, we used the logit of them in the equations to avoid estimations outside the [0,1] interval.
Following the Box-Jenkins time series analysis philosophy (Box and Jenkins 1976), as a first step, we checked the stationarity of the dependent variable.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

The results of the unit root test (Augmented Dickey-Fuller and Phillips-Perron) for logits of the default rates are as follows:

Table 1: Unit root tests of the default rate's logit

Level	ADF tau	ADF p-value	PP tau	PP p-value
Not differenced	-0.0794	0.6505	-0.1324	0.6342
First difference	-2.3364	0.0205	-2.4976	0.0133

As the second row of Table 1 shows, p-values for the logits of the default rates (not differenced) are high, so we cannot reject the null hypothesis of the unit root both in the case of Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. Therefore the logits are not stationary at the usually used significance levels (1%, 5% or 10%).

According to the final row of Table 1, the first differences are already stationary on 5% significance level, so variables are first order integrated according to the tests. As a result, we used the first differences of the logits for modelling.

Explanatory macro variables

We chose the following key macroeconomic variables and used during our stress test:

Table 2: Macroeconomic variables used for modelling (the large list)

Variable name	Description	Source
Real GDP (SA) volume growth	GDP volume index, basis period = same period in previous year, seasonally and calendar adjusted, expressed in percent	KSH
Real GDP (SA) gap (percent)	Computed using HP filter on seasonally adjusted real GDP (cyclical component, percentage deviation from „equilibrium” path)	own calculation
Consumer Price Index	Consumer Price Index (3 month average as quarterly data)	MNB
Employment rate (SA)	Employment rate, seasonally adjusted	KSH
Investments (national)	Investments, YoY change	KSH
Investments (companies)	Investments, YoY change	KSH
Industrial Production Index	Industrial Production index, same period in previous year = 100	KSH
BUX index	BUX index	BÉT
DAX index	DAX index	YahooFinance
HUF base rate	HUF base rate	MNB
BUBOR 3M	BUBOR, 3 month	MNB
CHF base rate	CHF base rate, call money rate	SNB
EUR base rate (lending)	EUR base rate, Marginal lending facility	ECB
EUR/HUF exchange rate	EUR/HUF exchange rate	MNB
CHF/HUF exchange rate	CHF/HUF exchange rate	MNB
Housing price index (FHB)	Housing price index (FHB), 2000 =100 (nominal)	FHB
Housing price index (Eurostat)	Housing price index (YoY change)	Eurostat
Retail deposits	Retail deposits (index, December 2001 = 100)	MNB
Leverage ratio	(Domestic corporate loans / Nominal GDP)·100	own calculation, MNB

For time series analysis the explanatory variables are required to be stationary (Hamilton 1999). If stationarity is not satisfied, then the estimated models are not considered reliable, and the risk of spurious regression incurs. This is the reason for why the stationarity of the explanatory variables is inspected. Augmented Dickey-Fuller and Phillips-Perron unit root tests were carried out first for the level of the variables. Regarding the level, only the GDP growth rate and the GDP gap proved to be stationary. For the other explanatory variables, the differences were also examined.

We examined the results for both the first ("d") and the yearly ("d4") difference of the variables. The choice between "d" and "d4" depends on their explanatory power on the default rates. (A preselection of the explanatory variables and their differenced forms is based on the analysis of the cross-correlation matrices.)

According to the unit root tests, all differenced variables can be regarded as stationary.

METHODOLOGY

As we previously mentioned, instead of modelling the default rates directly, we used the logit of them by the estimations. The logit of the default rate is calculated by the following transformation:

$$logit\ of\ the\ default\ rate = \log\left(\frac{defrate}{1-defrate}\right) \quad (1)$$

The connection between the explanatory variables and the dependent variable is estimated using linear autoregressive distributed lag model (using Maximum Likelihood estimation method).

The dependent variable is the first difference of the logit of the default rate. The explanatory variables are the previously described (already stationary transformed) macroeconomic variables and their lags of 1-4 quarter.

Furthermore, the autoregressive component of the dependent variable is also included in the regression (only the first lag proved to be correlated with the current value).

The applied autoregressive linear regression can be described by the following formula:

$$Y_t = \alpha_t + \beta_0 \cdot Y_{t-1} + \sum_{i=1}^{k} \boldsymbol{\beta_i} \cdot \boldsymbol{X}_{t-i} + \varepsilon_t \qquad (2)$$

Where

Y_t	differenced logit of the default rate of given segment in time period *t*
Y_{t-1}	autoregressive component
$\boldsymbol{X_{t-i}}$	macro variables and their lagged values (where *k* = 4 quarters)
ε_t	random error term
α_t and $\boldsymbol{\beta_i}$	coefficients estimated by the linear regression

In case of the dependent variable, the autocorrelation function can provide information about the numbers of lags to be included in the model. Generally, for the explanatory variables and their lags, the statistical significances of the estimated coefficients determine their inclusion or exclusion in the final model equation.

We made the estimation of the regressions with the help of SAS 9.4 software. The stepwise procedure is an iteration process where we start by including a wider range of explanatory variables into the model. During every step, the explanatory variable with the highest p-value is omitted if this p-value is above the pre-given threshold. If there are no more explanatory variables with p-values higher than the threshold, then the iteration is finished. The final model at the end of the procedure is used to carry out the estimation for the period 2016Q2-2019Q4.

The initial (wide) circle of the explanatory variables was determined with the help of the correlation matrices. The correlation matrices help to choose which differences (either "d1" or "d4") and lags of the possible explanatory variables should be included in the initial regression.

If we set the p-threshold at a relatively low level, then the estimation of the coefficients will be highly reliable as they are significantly different from zero even at a strict confidence level. However, the number of the explanatory variables will be strongly limited. The larger the threshold is, the more macro variables the estimation is based on. So we have more possibility to capture the link between the macro environment and the default rates. The coefficient estimations are however less reliable as less significant variables are also kept in the model.

RESULTS

We started the procedure with a wider range of explanatory variables; then the statistically insignificant variables were omitted during the steps. The variables in the initial large model were chosen based on the cross-correlation matrices. The following variables were included („D" indicates the order of difference, „L" indicates which lag is used):

Table 3: The initial model

Dependent variable	Logit_default rate (D1, L0)
Explanatory variables	Logit_default rate (D1, L1)
	Real GDP (SA) growth (D0, L0)
	Leverage ratio (D4, L1)
	Employment rate (D, L0)
	BUBOR (D4, L1)

After running the stepwise procedure, we got a final model (*** denotes that the variable is significantly different from 0 at 1% significance level; ** indicates 5% and * indicates 10% significance level):

Table 4: The final model

Explanatory variables	**Coefficient**	***t* value**	***p* value**
Intercept	0.0137	1.13	0.3270
Logit_default rate (D1, L1)	0.2543	1.82	0.1657
Real GDP (SA) growth (D0, L0)	-0.0151	-3.01	0.0514*
Leverage ratio (D4, L1)	3.4005	2.48	0.0157**

R^2	**Adjusted R^2**	**Number of obs.**	**Applied *p* threshold**
0.6884	0.6689	36	0.2

As presented in Table 4, the p threshold of 0.2 was applied, and two of the initial explanatory variables were omitted. Real GDP growth is significant at 10%, while Leverage ratio is significant even at 5% significance level.

Residual correlation diagnostics

In this part, we examine whether the final model is correctly specified by checking the residuals of the estimation. In the case of a correct specification, no autocorrelation remains in the residuals. The autocorrelation can be checked with the graphs of ACF and PACF functions, furthermore by the Durbin-Watson test.

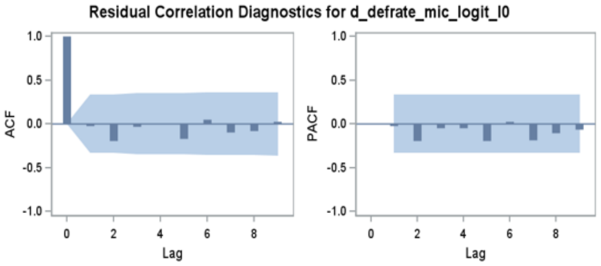


Figure 2: ACF and PACF functions of the residuals

The ACF and PACF functions show that the residual is not strongly correlated with its own lagged values. The Durbin-Watson test with its value close to 2 confirms the conclusion that the residual contains no autocorrelation (DW = 2.1584).

The forecasted default rates

Figure 3 plots the forecast results for the four estimated macroeconomic scenarios (baseline, adverse, severely adverse and the crisis; Crisis(7) in Figure 3 means that the crisis scenario was calculated based on the worst 7 quarters of the 2008 financial crisis period):

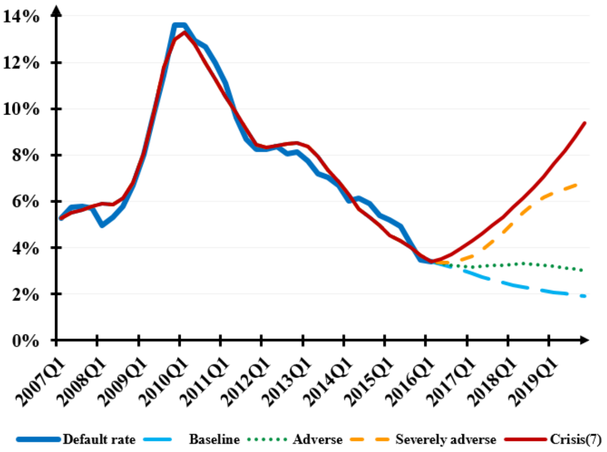


Figure 3: Default rate forecast along the scenarios

Table 5 summarizes the results numerically. Beyond the average yearly default rates (calculated as a simple average of the 4 quarters) it also presents the default multipliers. The multiplier is a ratio comparing the default rate of the adverse, severely adverse and crisis scenario to the respective default rate of the baseline scenario. The multiplier is expressed in percent. It tells how severe the certain scenario is, compared to the baseline forecast.

Table 5: Summary table of the model forecasts

	2017	**2018**	**2019**
Baseline default rate	2.50%	2.14%	1.90%
Adverse default rate	3.26%	3.26%	3.02%
Adverse multiplier	**130.56%**	**152.62%**	**159.15%**
Severely adverse default rate	4.65%	6.18%	6.86%
Severely adverse multiplier	**186.06%**	**289.13%**	**361.04%**
Crisis default rate	5.33%	7.09%	9.37%
Crisis multiplier	**213.49%**	**331.97%**	**493.38%**

TRANSLATING DEFAULT RATES TO PD

For further estimations within the stress testing framework, we need segment-level stressed PD values. That means that we have to relate the forecasted default rates with the probability of default. This is done using migration matrices that tell us how clients' ratings change over time.

We took into account the recommendation of the European Banking Authority (EBA) about stress testing framework and migration matrices (EBA 2016). According to this suggestion, institutions need to calculate point-in-time transition matrices, and these matrices should meet at least the following two criteria:

- The PD for each grade should calculate in line with the scenarios, and
- The probabilities of moving between grades are adjusted according to the scenarios.

Considering the above, we created the one-year observed migration matrix. The matrix show how individual ratings changed between 2015 and 2016. As a next step, we stress these matrices using our estimated default rate multipliers. The methodology can be described as followed:

1. We define a common stress parameter φ that serves as a factor in which we shift the distribution of the ratings in the matrix (φ is different for every scenario).
2. We calculate the yearly default rates for the years 2016-2019 based on the observed migration matrix. These will be the "baseline" default rates.
3. We calibrate φ in a way that the ratio of the stressed and the baseline default rate at the end of the forecasting horizon (2019) would be the same as the corresponding default rate multiplier estimated by the linear regressions.
4. With these stressed migration matrices we calculate the total exposure in every rating category for all four years.
5. Using the fixed PD of the rating categories and the total stressed exposures we can calculate the segment-level average PD for all four years.

The observed one-year migration matrix ("base" migration matrix) is the following:

Table 6: Base migration matrix, number of clients

	C1	C2	C3	C4	C5	C6	C7	C8	D	Total
C1	24	6	1							**31**
C2	312	641	79	2	1	5	5	6	8	**1 059**
C3	89	1 121	601	49	30	33	9	17	39	**1 988**
C4	2	121	229	37	20	10	5	11	23	**458**
C5	1	53	143	39	230	15	3	14	11	**509**
C6		18	63	26	13	10	5	5	20	**160**
C7	2	10	38	14	23	30	8	6	19	**150**
C8		15	22	18	42	28	47	63	54	**289**
D					1			1	531	**533**
Total	**430**	**1 985**	**1 176**	**185**	**360**	**131**	**82**	**123**	**705**	**5 177**

Expressed in percent:

Table 7: Base migration matrix, percentage of clients

	C1	C2	C3	C4	C5	C6	C7	C8	D
C1	77.42%	19.35%	3.23%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
C2	29.46%	60.53%	7.46%	0.19%	0.09%	0.47%	0.47%	0.57%	0.76%
C3	4.48%	56.39%	30.23%	2.46%	1.51%	1.66%	0.45%	0.86%	1.96%
C4	0.44%	26.42%	50.00%	8.08%	4.37%	2.18%	1.09%	2.40%	5.02%
C5	0.20%	10.41%	28.09%	7.66%	45.19%	2.95%	0.59%	2.75%	2.16%
C6	0.00%	11.25%	39.38%	16.25%	8.13%	6.25%	3.13%	3.13%	12.50%
C7	1.33%	6.67%	25.33%	9.33%	15.33%	20.00%	5.33%	4.00%	12.67%
C8	0.00%	5.19%	7.61%	6.23%	14.53%	9.69%	16.26%	21.80%	18.69%
D	0.00%	0.00%	0.00%	0.00%	0.19%	0.00%	0.00%	0.19%	99.62%

Using the migration matrix and the sum of the clients in every rating category we can easily calculate the default rates for the forecasted years, that is the percentage of non-defaulted (C1-C8) clients that migrate to D (defaulted) rating category.

Then, we stress the matrix by the factor φ. "Stressing" means that φ% of the clients in a cell is shifted to the next cell to the right, that is they migrate to a worse category. For example, if φ = *10%* then the C1-C1 cell will be *90%·77,42%* = *69,68%*, while C1-C2 cell will be *90%·19,35%* + *10%·77,42%* = *25,16%*. With this methodology, we can stress the whole segment depending only on one factor that can easily be calibrated to be in line with our regression estimates.

As already mentioned, the stress factor φ is calibrated in a way to get the same ratio of the stressed and baseline default rates as estimated with the regressions (marked with blue in Table 5).

Following the same steps as before we calculated the average segment PD. The results are summarized in Table 8:

Table 8: Stressed PD results for micro segment

Scenario		2017	2018	2019	Stress factor (φ)
Baseline	Default rate based on observed migration matrix	2.35%	1.63%	1.24%	
	Average segment PD	**4.12%**	**3.17%**	**2.63%**	
Adverse	Default rate based on observed migration matrix	3.05%	2.35%	1.97%	**20.39%**
	Default rate multiplier	130.01%	144.42%	**159.14%**	
	Average segment PD	**5.13%**	**4.35%**	**3.91%**	
Severely adverse	Default rate based on observed migration matrix	4.99%	4.58%	4.47%	**71.41%**
	Default rate multiplier	212.74%	281.93%	**361.04%**	
	Average segment PD	**8.02%**	**7.90%**	**7.86%**	
Crisis	Default rate based on observed migration matrix	6.03%	5.93%	6.11%	**96.16%**
	Default rate multiplier	256.91%	364.37%	**493.38%**	
	Average segment PD	**9.62%**	**9.95%**	**10.18%**	

As a summary, Figure 4 presents the stressed segment PDs along the 4 scenarios:

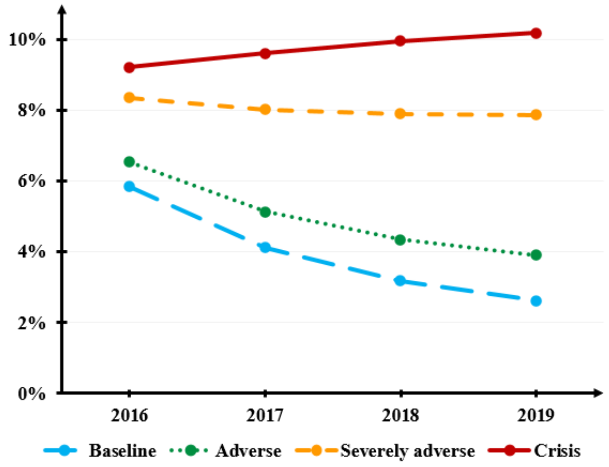


Figure 4: Stressed segment PDs

And finally, with the help of these PDs, we can estimate the effect of the different stress scenarios to the P&L and the capital adequacy of the given financial institution.

CONCLUSION

In this paper, we presented a possible methodology to calculate the stressed point-in-time PD parameter. This estimation is crucial for a sound stress testing, but the final methodology that a financial institution chooses has to be in line with the scope and complexity of the bank.
The regression method we applied requires long enough time series which is not always available. Therefore alternative methods and expert-based considerations should also be taken into account. And even if econometric modelling is possible, the results should be evaluated and sometimes modified by experts who have deep knowledge of the bank's portfolio and risk profile.

REFERENCES

Box G.E.P. and G.M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control.* Holden-Day. San Francisco.

Budapest Stock Exchange Statistics https://bet.hu/ (downloaded on the 27th of October 2016)

European Banking Authority (EBA). 2016 EU-Wide Stress Test Methodological Note. 24 February 2016.

European Central Bank (ECB) Statistics https://www.ecb.europa.eu (downloaded on the 27th of October 2016)

Eurostat Statistics http://ec.europa.eu/eurostat (downloaded on the 27th of October 2016)

Hamilton J.D. 1994. *Time Series Analysis.* Princeton University Press. New Jersey.

Hungarian Central Bank Statistics https://www.mnb.hu/ (downloaded on the 27th of October 2016)

Hungarian Central Statistical Office Statistics http://www.ksh.hu/ (downloaded on the 27th of October 2016)

Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms

Swiss National Bank Statistics https://www.snb.ch/en/ (downloaded on the 27th of October 2016)

YahooFinance Statistics https://finance.yahoo.com/ (downloaded on the 27th of October 2016)

AUTHOR BIOGRAPHIES

ZOLTÁN POLLÁK completed his MSc degree summa cum laude in Finance at Corvinus University of Budapest. He is currently doing a Ph.D. at the Department of Finances. He is lecturing financial courses such as Corporate Finance and Financial Calculations. Beside Ph.D. he works as a partner consultant for the International Training Center for Bankers (ITCB), where he also teaches on banking and investment courses.
His e-mail address is: ZPollak@bankarkepzo.hu

DÁVID POPPER graduated from Economics MA program of the Central European University. His main fields of interest are financial economics and economic growth (especially different aspects of economic convergence). He currently works as a junior consultant at ITCB where he is responsible for developing various credit risk models including stress testing methodologies.
His e-mail address is: DPopper@bankarkepzo.hu

COMBINATION OF TIME-FREQUENCY REPRESENTATIONS FOR BACKGROUND NOISE SUPPRESION

Eva Klejmová, Jitka Poměnková, and Jiri Blumenstein
Department of Radio Electronics
Brno University of Technology
Technicka 3082/12, Brno, Czech Republic
Email:klejmova@phd.feec.vutbr.cz, pomenkaj@feec.vutbr.cz, blumenstein@feec.vutbr.cz

KEYWORDS

continuous wavelet transform, time-varying autoregressive process, short time Fourier transform, time-frequency representation, noise suppression

ABSTRACT

The aim of the paper is to propose approach for enhancement of time-frequency representation leading to the background noise suppression. The approach is based on combination of continuous wavelet analysis, time-varying autoregressive process and short time Fourier transform. By such combination we make the identification of important areas in the time-frequency representation easier. The proposed method is an alternative approach to significance tests which can be problematic in some cases. The performance of methods is presented on the gross domestic product of the United Kingdom and Group of 7. The results show that in the UK, oil crisis has a bigger impact compared to financial crisis, while from the perspectives of G7 countries, the impact of financial crisis was stronger. The obtained results can be also used for consequent econometric analysis which identify dependencies, relations, bilateral causalities or other economic aspects.

INTRODUCTION

The need to analyse the data can be found across most disciplines. Despite the diversity of disciplines, it is a common goal to obtain the maximum information from data analysis that will help to solve the tasks set. With respect to the scientific area such data are given as observations in the form of time series or input signals. The common analytical instruments are given in time or frequency domain. The linking of both approaches giving us a more compact view can be done via time-frequency techniques (TF).

The literature includes many interesting papers from application in engineering (Stankovic et al. 2012; Liu et al. 2011), biology and medicine (Faust et al. 2015), or economics (Maršálek et al. 2014; Ftiti et al. 2014).

The estimation of TF representation of the data can be done via several approaches. The widely used is short time Fourier transformation (STFT) (Proakis et al. 2002), the time-frequency varying autoregressive process (TFAR) (Proakis et al. 2002), multiple window method (Cakrak and Loughlin 2001) and wavelet analysis (Rajmic 2014). There are also alternative approaches such as modified empirical mode decomposition (Sebesta et al. 2013), the usage of Wigner-Vile distribution (Orovic and Stankovic 2009) or methods for more complicated multicomponent signals (Stankovic et al. 2012).

Suitable methods for the analysis of non-stationary signals is STFT, continuous wavelet transform (CWT), multiple windows method or TFAR. The TFAR is a simplification of the general autoregressive moving average model. The comparison of these main methods can be found in Blumenstein et al. (2012) or Klejmová (2015). The results shows that while CWT has better time resolution, the TFAR has better frequency resolution.

In most economic application the wavelet analyses predominates. The reason can be found in its simple usage for non-stationary signal and better time resolution (Jiang and Mahadevan 2011). The use of CWT for estimation of co/cross-spectra, the co-movement analysis is very popular. Such an approach put in evidence the existence of both long run and short-run co-movement. Aguiar-Conraria and Soares (2014) generalize the concept of simple coherency to partial wavelet coherency and multiple wavelet coherency akin to partial and multiple correlations. Berdiev and Chang (2015) took TF framework to examine the strength of business cycle synchronization. Fidrmuc et al. (2014) apply wavelet spectrum analysis to study globalization and business cycles in China and G7 countries. And Maršálek et al. (2014) proposed an original method based on CWT for filtering-out the global shocks from the time series.

In order to have better predictive power it is suitable to support and validate the obtained results via some testing. The basic work is given by Torrence and Compo (1998). The paper presents statistical significance tests for wavelet power spectra are developed by deriving theoretical wavelet spectra for white and red noise processes and using these to establish significance levels and confidence intervals. Similar approach to Torrence and Campo can be found in Schulte et al. (2015) or Ge (2007). Ge derived the sampling distributions of the wavelet power and power spectrum of a Gaussian White Noise (GWN) in a rigorous statistical work. He proved that the results given by Torrence and Compo (1998) are

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

numerically accurate when adjusted by a factor of the sampling period. The different approach to model validation is proposed by Jiang and Mahadevan (2011). The author uses testing via Monte Carlo (MC) simulations to infer whether the model prediction and experimental observation represents two coherent processes.

Significance tests proposed by Torrence and Compo (1998) or Ge (2007) require a priori knowledge of the noise character. As shown in Pomenkova (2017), MC results for CWT differs from Ge (2007) especially in case of heteroskerasticity in input data. Therefore, we propose combination of several TF methods as an alternative to these tests. In each method the background noise is depicted with different characteristics. However significant spectral components should be captured in most cases. Based on such an assumption we should be able to suppress the noise and highlight required components by using their combination. While in engineering background noise can be considered as the rest after removal periodic components (GWN, red noise etc.), in case of economic data the situation is not the same. Economic data can be viewed as a composition of several cyclical components which can occur in different time sub-period (not in whole time). The nature of an economic indicator plays an important role and can influence the character of nested cycles. In such way the background noise character is usually taken as a weakly stationary series and is obtained in dependence on analytical approaches.

The objective of the paper is propose approach for enhancement of TF representation leading to the background noise suppression. Thus, on the basis of the proposed method, we make the identification of important areas in the TF representation easier. the application of the proposed methodology on economic data allows easier interpretations from time and frequency perspectives. It can be also used for consequent macro/micro-econometric analysis of dependencies, relations with other economic aspects or analysis of bilateral causalities of all series or its cyclical components. The performance of methods is presented on the gross domestic product data of the United Kingdom and G7. These representatives were chosen because of available sample size and because they represent leading economies.

METHODICAL BACKGROUND

Continuous Wavelet transform (CWT)

In order to describe the parameters of a signal not only in time but also in frequency domain, wavelet transform and its modifications can be used. One of these modifications, which is commonly used for assessing cyclical movements in different types of macroeconomics time series is continuous wavelet transform (CWT). It can be described as the integral of analysed signal with the base function (mother wavelet) (Walnut 2013):

$$\begin{array}{c} S_{CWT}(a,\tau) = \int_{-\infty}^{\infty} s(t)\frac{1}{\sqrt{a}}\psi\left(\frac{t}{a}-\tau\right)dt, \\ a>0, \tau \in R, \end{array} \tag{1}$$

where $s(t)$ is the time series, $\psi\left(\frac{t}{a}-\tau\right)$ is a scaled version of the mother wavelet, τ denotes the time shift, and a denotes the scale (or frequency) (Walnut 2013).

To be the invertible transform, basis (mother wavelets) functions must be mutually orthogonal, have zero mean value and limited to finite time interval. That is

$$\begin{array}{ll} i) & \int_{-\infty}^{\infty} \psi\left(\frac{t}{a}-\tau\right)dt = 0, \\ ii) & \int_{-\infty}^{\infty} \psi^2\left(\frac{t}{a}-\tau\right)dt = 1, \\ iii) & 0 < C_\psi = \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega}; \\ & \Psi(\omega) = \int_{-\infty}^{\infty} \psi\left(\frac{t}{a}-\tau\right)e^{-i\omega t}dt, \end{array} \tag{2}$$

where $\Psi(\omega)$ is the Fourier transform of $\psi(\omega)$. To satisfy the assumptions for the time-frequency analysis, waves must be compact in time as well as in the frequency representation. There are several types of mother wavelets which can be used (e.g. Gaussian, Haar, Daubechies, Morlet etc.) In this paper, we use the complex Morlet wavelet (Walnut 2013):

$$\psi(t) = exp\left(\frac{-t^2}{2\sigma^2}\right)exp(i\omega_0 t), \tag{3}$$

where σ is a Gaussian window width in time and ω_0 is the central frequency of the wavelet. The complex Morlet wavelet is a substantially complex exponential modulated by a Gaussian envelope. In order to recalculate the local frequency, corresponding to the scale a, the following equation can be used (Walnut 2013).

$$\omega(\tau) = \frac{\omega_0}{a(\tau)}, \tag{4}$$

where $\omega = 2\pi f$ and τ is the time shift.

Time-frequency varying AR process (TFAR)

This method uses a parametric approach and creates a model generating an input signal. The analysed signal $s(n)$ is then regarded as the output of a linear filter influenced by white noise w with variance σ_w^2. The autoregressive process can be described by the AR(p) model given by the equation

$$s(n) = c + \sum_{i=1}^{p} a_i s_{n-i} + w_n, \tag{5}$$

where $a_i,\ i = 1,\dots p$ are the parameters of the autoregressive model of the order p, c is a constant and w_n is white noise. The output spectrum can be described

$$S(f) = \left|H\left(e^{j2\pi fT}\right)\right|^2 \sigma_w^2, \tag{6}$$

where $H\left(e^{j2\pi fT}\right)$ is a linear time variant filter. Thus the spectrum estimation, when we use the AR(p) process, is done according to the formula (Proakis 2002)

$$S_{AR}(f) = \frac{\widehat{\sigma}_w^2}{\left|1+\sum_{i=1}^{p}\widehat{a}_i e^{-j2\pi fi}\right|^2}, \tag{7}$$

where $\widehat{a_i}$ are estimates of the AR(p) parameters and p is the lag order. Several methods for estimating AR(p)

model parameters can be used. The most common are the Burg method, Yule-Walker method, unconstrained least-squares method or sequential estimation methods. For the AR process, an appropriate selection of lag order p plays an important role. Selecting a low level order leads to an excessive smoothing of the spectrum. Furthermore, if the level of p is too high, a non-significant spectral coefficient can appear to be a high peak. For the optimal selection, several criteria can be used (Proakis 2002).

Fourier transform

One of the most common methods used for spectrum estimation is the Fourier transform (FT) and its modifications. If an input signal $s(n)$ is a discrete time series, then the Discrete Fourier transform (DFT) is used. It can be defined as (Proakis 2002)

$$S_{FT}(f) = \sum_{n=0}^{N-1} s(n) e^{-j2\pi fn}. \qquad (8)$$

A slight modification of this method is called Short Time Fourier Transform (STFT), when the Fourier transform is calculated using a sliding observation window. The individual spectrum estimations are then sorted in time and can be plotted in a 2D graph.

APPLICATION

Data

We use seasonally adjusted quarterly data of the gross domestic product (GDP), volume index in OECD reference year 2010 (OECD 2017) of the United Kingdom (UK) in 1956/01-2016/03 and Group of 7 (G7) in 1961/02-2016/03. All variables are in first differences of logarithms (Fig. 1), further they will be denoted as GDP. The motivation for the data selection was appropriate data sample size for testing the proposed method. We needed sufficient data range to have detailed time resolution. Both the UK and G7 data meet these requirements. Also, the data sets were chosen to be overlaping, and thus supporting the validation of proposed method. Additionally, because of the Brexit we were interested in analysing the data before this even which can be considered a structural break affecting the data.

Setting of TF methods

Our analysis consists of several steps. First, we analyse the data using CWT. We set scales to correspond to the range of 1 year to 10 years, with 257 individual scales. We selected the complex Morlet with center frequency $f_b = 1.5$ as mother wavelet (Poměnková and Klejmová 2015). The complex Morlet wavelet is based on the standard Morlet with the advantage of providing complex results making it possible to obtain a phase part (quadrature) of spectrum. In case of TF estimation via TFAR process we used Burg approach for coefficient estimates on 30 samples with 29 samples overlaping and with the Hann window. The optimal value of lag order was based on the AIC criteria (Klejmová 2015). The parameters of STFT were set to correspond to the TFAR settings (30 samples, 29 samples overlaping, Hann window) to simplify the process of combining the methods.

Combination of TF methods

Significance tests based on Ge (2007) require the knowledge of the noise character. This assumption can be broken when the data are heteroscedastic. Thereafter, Monte Carlo simulation for CWT can differs from Ge (2007) approach. To avoid this complication we suggest combination of several TF approaches as an alternative to significance tests. To obtain the best possible TF representation we combined results from the CWT, TFAR and STFT approach. Since the main focus was on the amplitude part of the spectra, we omitted the phase part of complex spectra S_{CWT} and S_{STFT}. When focusing on the amplitude and phase components whole signal can be used for subsequent processing.

Firstly we align the time axis (time resolution) of spectral representations S_{CWT}, S_{AR} and S_{STFT} so that each spectrum corresponds to one another. All three time vectors have linearly increasing trend, therefore for the time axis alignment the only requirement was to adjust the starting and ending point for each method. We omitted the first and last 15 columns of S_{CWT}, we denoted this remaining matrix as $S^{'}_{CWT}$. By doing this, we ensured corresponding the time axis for all three methods.

Secondly we needed to align the frequency/scale axis of $S^{'}_{CWT}$, S_{AR} and S_{STFT}. The frequency range of S_{AR} and S_{STFT} was cropped to correspond to the range of $S^{'}_{CWT}$ which was 1 year to 10 years cycles. He resulting frequency/business cycles vectors $\overline{f_{AR}}$ and $\overline{f_{STFT}}$ had a linearly increasing trend, however, the trend of $\overline{f_{cwt}}$ was non linear. To obtain the corresponding vectors we matched each point of $\overline{f_{cwt}}$ with one value of $\overline{f_{AR}}$/$\overline{f_{STFT}}$ with 1.4% tolerance:

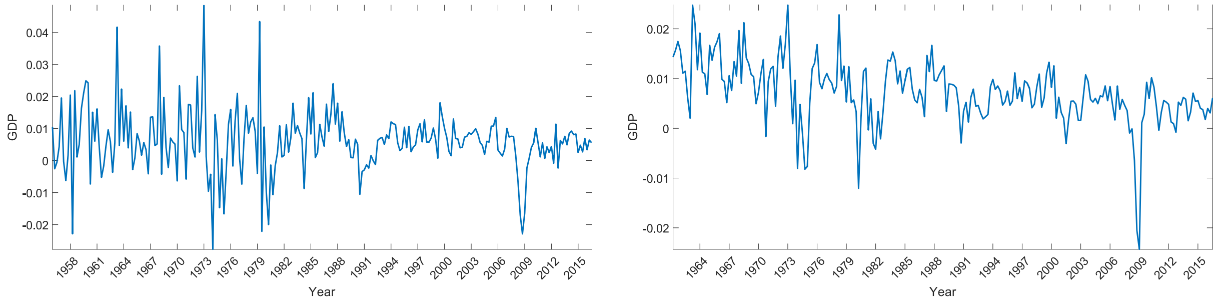
$$\begin{gathered} |f_{CWT} - f_{STFT}| \leq 0.014 \left| max(\overline{f_{CWT}}; \overline{f_{STFT}}) \right| \\ |f_{CWT} - f_{AR}| \leq 0.014 \left| max(\overline{f_{CWT}}; \overline{f_{AR}}) \right| . \end{gathered} \qquad (9)$$

With this step, we have gained the adjusted TF matrices $S^{'}_{AR}$ and $S^{'}_{STFT}$ making all three methods aligned. For the combination of methods we selected a simple multiplication (Klejmová and Poměnková 2017). We used combination of CWT and AR ($S_{CWT,AR}$) and the combination of CWT, AR and STFT ($S_{CWT,AR,STFT}$):

$$\begin{gathered} S_{CWT,AR} = S^{'}_{CWT} S^{'}_{AR} \\ S_{CWT,AR,STFT} = S^{'}_{CWT} S^{'}_{AR} S^{'}_{STFT}. \end{gathered} \qquad (10)$$

RESULTS

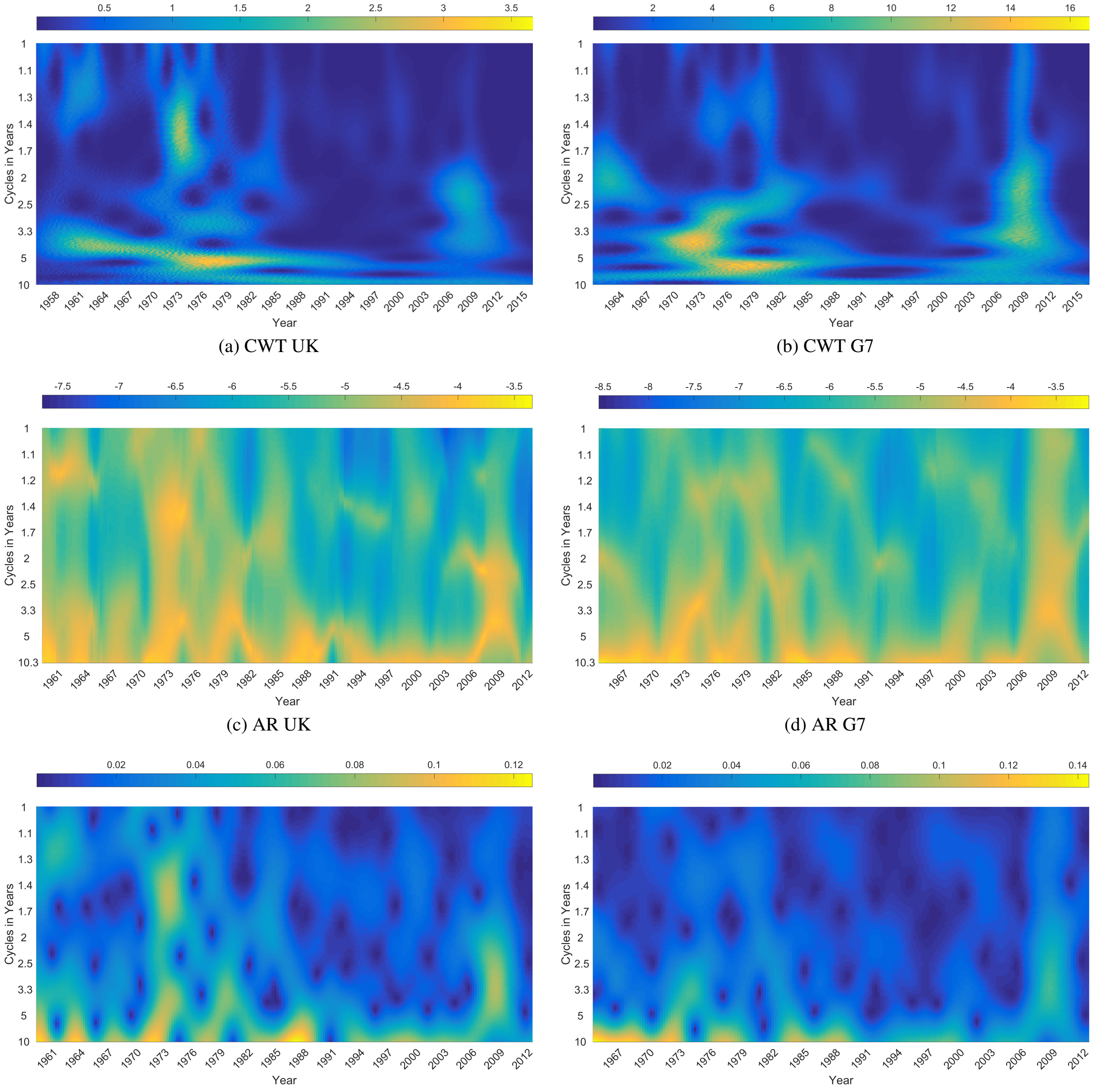
The data and results for the UK and G7 are presented graphically in the Fig. 1a-b, in Fig. 2a-f and Fig. 3a-d.



(a) UK

(b) G7

Figure 1: GDP of UK and G7 in time domain



(a) CWT UK

(b) CWT G7

(c) AR UK

(d) AR G7

(e) STFT UK

(f) STFT G7

Figure 2: Spectrum of GDP of UK and G7 in time domain

There are four types of figures. Namely the time representation of GDP for the UK and G7 (Fig. 1a-b), TF transform via CWT (Fig. 2a-b), TF transform via AR (Fig. 2c-d), transformation via STFT (Fig. 2e-f) and the adjustment of CWT picture with the help of AR (Fig. 3a-b) and with the help of AR+STFT (Fig. 3c-d).

Focusing on the time representation given in Fig. 1a-b, we can conclude the following. In the time representation of the United Kingdom data, we can see two sub-periods with different volatility, between 1956-1988 and 1989-2015. There are also visible several moments with higher/lower levels of the data, i.e. structural breaks (1958, 1964, 1968, 1973, 1979 and 2008) given by events in the UK economy such as oil crises, financial crisis. In the case of G7, there is a similar problem with volatility, but it is not visible as in the UK. In contrast with the UK, the G7 data has a slowly decreasing trend with a higher volatility between 1961-1988. Afterwards (1989-2015) the data character looks similar to the UK. In G7 we can see similar structural breaks (1973, 1979 and 2008).

After a pilot analysis of time representation of the data, we applied TF approaches. Firstly we modelled CWT (Fig. 2a-b), consequently TFAR (Fig. 2c-d) and STFT (Fig. 2e-f). As expected, CWT provides results with a very good time resolution. We can see several important areas across time and frequency. Focusing on TFAR representation the results are not so clear from time perspective as CWT, but they give us better information from frequency perspective in a similar way to STFT. Therefore, we decided to adjust the CWT picture with the help of TFAR and TFAR+STFT according to the calculation (eq. 9,10) described in Combination of TF methods.

The results of the adjustments can be seen in Fig. 3a-d. Focusing on the UK situation and comparing the enhanced figure (Fig. 3a) with the CWT figure (Fig. 2a), we can see a sharper picture with suppressed noise. Therefore, we can easily identify the most important events in the UK data from the time and frequency perspectives. We can find three most important events in the UK. The first is between 1960-1988 and can be divided into two sub-periods; 1960-1972 and 1973-1988. These results correspond to the time domain description. In addition to the time domain, we find that such an event results in a reaction in approximately 4 years (in the first sub-period 1960-1972) and around 5 years (in the second sub-period 1972-1988) cycles. The second important can be identified between years 1973-1976 and results in an economic reaction in short cycles of proximately 1.5 year. The last important area is between 2007-2010. It cover business cycle frequencies (from 4 to 2 year frequencies) and if compared to the previous one, it does not seem to have such impact in the UK as the previous events. To confirm this conclusion and to validate the results, we add an additional adjustment of three TF approaches leading to Fig. 3c. The result confirms the conclusion results from the adjustment of CWT and TFAR and the fact that the events between 1970-1976 have a stronger impact on the UK economy.

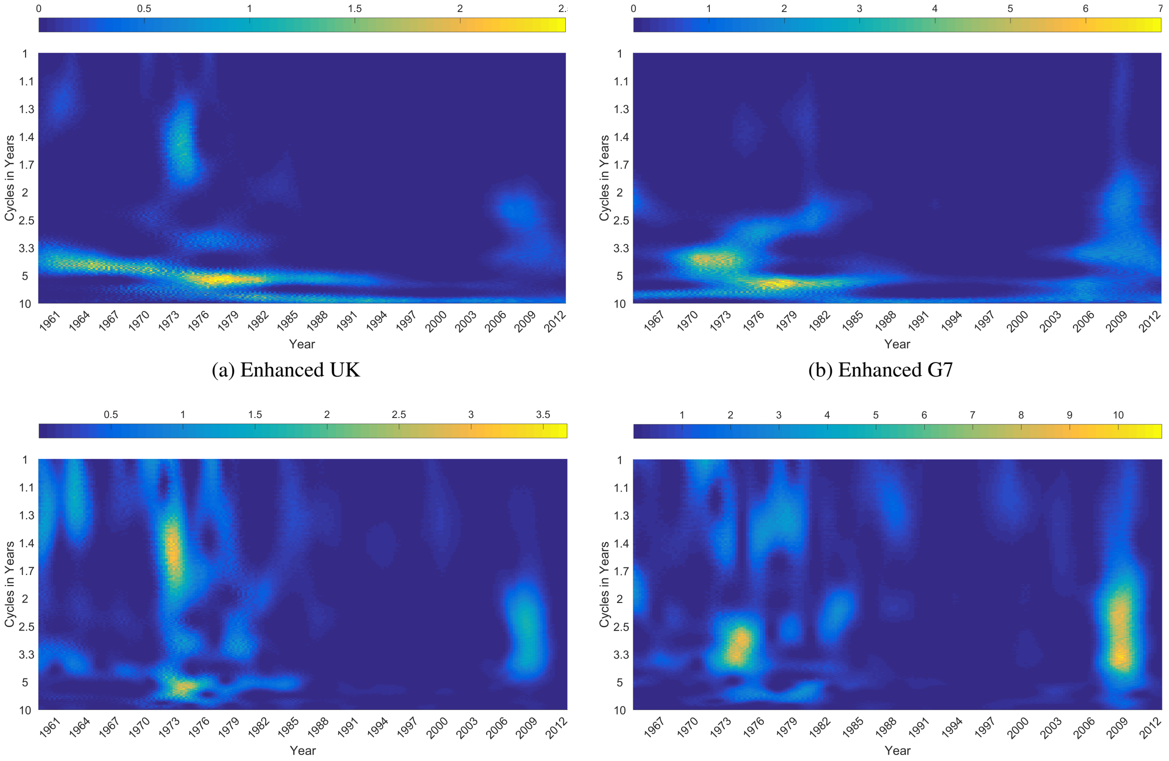
If we focus on the results for G7, we can find some similarities as well as dissimilarities. Again, we can see (Fig. 3h) that the most important area is between 1970-1974 in 4 years cycles and 1974-1981 in 5 years cycles consisting one period 1970-1982. The second important area is between 2007-2010 (financial crisis) which covers the range of business cycles, i.e. 1.5-5 years cycles. After adding the second adjustment for the cross validation of the obtained results presented in Fig. 3d, we see conformity in identification of important area. Also in this case, to confirm and validate the results, we added adjustment of TF approaches leading to Fig. 3d. The result confirms conclusion from the adjustment of CWT and TFAR and the fact that the events between 1970-1976 have a stronger impact on G7 economy than the financial crisis in 2007-2009.

Comparing the results from the economy point of view, we can see, that in the UK, the oil crisis has a bigger impact than the financial crisis, while from the perspectives of G7 countries, the impact of financial crisis was stronger. The obtained results can be used for consequent macro or micro-econometric analysis searching for dependencies or relations with other economic aspects. Moreover, they can motivate researcher to investigate further steps, e.g. decomposition analysis on specific component of the corresponding frequency which can be used for analysing bilateral causalities.

If we review the results from a methodological point of view, we successfully adjusted CWT with others TF approaches and suppressed the background noise. Consequently, certain events occurred to be much more visible and, further, the time as well as frequency identification was easier. Other possible research is significance testing according to Torenc and Compo (1998) or Ge (2007) with the investigation of Monte Carlo simulation or with the investigation of background noise description.

COMPARISON WITH SIGNIFICANCE TEST

To assess the performance of the proposed approach (combinations of TF methods), we carried out a significance test described in Ge (2007). To do this we assumed the background noise character to be GWN. In the case of different noise character or heteroskedasticity data character, the Monte Carlo simulation may be carried out (for details see Pomenkova (2017)). Significant parts of CWT spectrograms based on Ge (2007) can be seen in Fig. 4a-b. When applying the proposed method for the UK case, we identify the most important area of approximately 4-5 years cycles during 1960-1988and also the second important area between 2007-2010, covering business cycles. In the case of G7, the Ge approach was able to identify the most important area between 1970-1974 in 4 years cycles, 1974-1981 in 5 years cycles and also area between 2007-2010 covering the range of business cycles. In both (UK and G7) cases, the Ge approach tends to omit spectral peaks with shorter fluctuations.



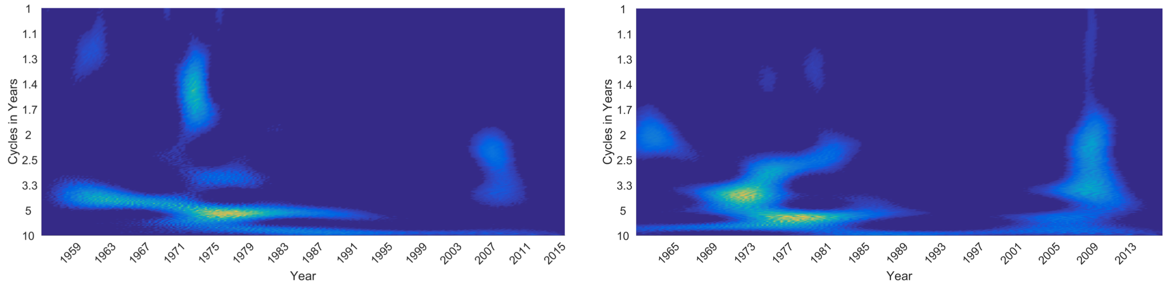
(a) Enhanced UK

(b) Enhanced G7

(c) Enhanced UK

(d) Enhanced G7

Figure 3: Adjustment of TF methods



(a) UK

(b) G7

Figure 4: Significant components based on Ge (2007)

CONCLUSION

The presented paper deals with enhancement of time-frequency representation leading to the background noise suppression. The new approach is based on the combination of CWT, TFAR and STFT making it easier to identify important areas in the TF representation.

The methods were performed on the GDP of the United Kingdom and Group of 7. The results show that in the UK, the oil crisis has a bigger impact compare to the financial crisis, while from the perspectives of G7 countries, the impact of the financial crisis was stronger. From a methodological point of view, we successfully adjusted CWT with others TF approaches and enhanced the time-frequency picture of the UK and G7 GDP with the suppressed background noise.

ACKNOWLEDGEMENT

The research described in the paper was supported by the Czech Science Foundation via Grant n. 17-24309S and by the Czech Ministry of Education in the frame of National Sustainability Program under grant LO1401. For research, the infrastructure of the SIX Center was used.

References

Aguiar-Conraria, L. and M. J. Soares. 2014. ”The continuous wavelet transform: Moving beyond uni-and bivariate analysis.” *Journal of Economic Surveys* , Vol.28, 344-375.

Berdiev, A. N. and Ch-P. Chang. 2015. ”Business cycle synchronisation in Asia-Pacific: New evidence from wavelet analysis.” *Journal of Asia Economics* , Vol.37, 20-33.

Blumenstein, J., J. Poměnková and R. Maršálek. 2012. ”Comparative Study Of Time- Frequency Analysis Approaches With Application To Economic Indicators.” In: *Proceedings of the 26th European Conference on Modelling and Simulation ECMS 2012* (Troitzsch, K., Mhring, M., Lotzmann, U., eds.). University Koblenz Landau, Koblenz, 291-297.

Cakrak F. and P. J. Loughlin. 2001. ”Multiple window time-varying spectral analysis.” *IEEE Transactions on Signal Processing*, Vol. 49, No. 2, 448-453.

Crowley, P. and A. Hughes Hallett. 2014. ”Volatility transfers between cycles: A theory of why the ”great moderation” was more mirage than moderation.” *Bank of Finland Research Discussion Papers*, Vol.23, 1-19.

Faust, O., U. R. Acharya, H. Adeli and A. Adeli. 2015. ”Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis.” *Seizure*, Vol. 26, 56-64.

Fidrmuc, J., I. Korhonen and J. Poměnková. 2014. ”Wavelet spectrum analysis of business cycles of China and G7 countries.” *Applied Economic Letters* , Vol.21, 1309-1313.

Ftiti, Z., A. Tiwari and A. Belanés. 2014. ”Tests of Financial Market Contagion: Evolutionary Cospectral Analysis V.S. Wavelet Analysis.” *Computational Economics* , Vol.46, 575-611.

Ge, Z. 2007. ”Significance tests for the wavelet power and the wavelet power spectrum.” *Annales Geophysicae*, Vol.25, No.11, 2259-2269.

Jiang,W. and S. Mahadevan. 2011. ”Wavelet spectrum analysis approach to model validation of dynamic systems.” *Mechanical Systems and Signal Processing*, Vol.25, 575-590.

Klejmová, E. 2015. ”Wavelet Significance Testing with Respect to GWN Background: Monte Carlo Simulation Usage.” In: *Proceedings of the 26th International Conference Radioelektronika 2017*, In print.

Klejmová, E. and J. Poměnková. 2017. ”Identification of a Time-Varying Curve in Spectrogram.” *Radioengineering*, In print.

Liu, S., D. Wang, T. Li, G. Chen, Z. Li and Q. Peng. 2011. ”Analysis of photonic Doppler velocimetry data based on the continuous wavelet transform.” *Review of Scientific Instruments*, Vol. 82, No. 2.

Maršálek, R., J. Poměnková and S. Kapounek. 2014. ”A Wavelet-Based Approach to Filter Out Symmetric Macroeconomic Shocks.” *Computational Economics*, Vol. 44, No. 4, 477-488.

Orovic, I. and S. Stankovic. 2009. ”A Class of Highly Concentrated Time-Frequency Distributions Based on the Ambiguity Domain Representation and Complex-Lag Moment.” *EURASIP Journal on Advances in Signal Processing*, Vol. 2009, 35:1–35:9.

Poměnková, J. and E. Klejmová. 2015. ”Identification of time-varying model using wavelet approach and AR process.” In: *33rd International Conference Mathematical Methods in Economics MME 2015*, University of West Bohemia, Plze, 665-670.

Poměnková, J., Klejmová, E. and T. Malach. 2017. ”Evaluation of Background Noise for Significance Level Identification” In: *Proceedings of the 24th International Conference on Systems, Signals and Image Processing*, Forthcoming.

Proakis, J. G., CH. M. Rader, F. L. Ling, CH. L. Nikias, M. Moonen and J. K. Proudler. 2002. *Algorithms for Statistical Signal Processing.* Prentice Hall. ISBN 0-13-062219-2.

Rajmic, P. and Z. Prusa. 2014. ”Discrete wavelet transform of finite signals: detailed study of the algorithm.” *International Journal of Wavelets, Multiresolution and Information Processing*, Vol.12, No.1, 1-38.

Schulte J. A., C. Duffy and R. G. Najjar. 2015. ”Geometric and topological approaches to significance testing in wavelet analysis.” *Nonlinear Processes in Geophysics*, Vol. 22 , No. 2, 139-156.

Sebesta, V., R. Marsalek, R. and J. Pomenkova. 2013. ”The Modified Empirical Mode Decomposition Method For Analysing The Cyclical Behavior Of Time Series.” In: *27th European Conference on Modelling and Simulation ECMS 2013.* Aalesund University College (AAUC), Norway, 288-292.

Stankovic, S., I. Orovic, and V. Sucic. 2012. ”Averaged multiple L-spectrogram for analysis of noisy nonstationary signals.” *Signal Processing*, Vol.92, 3068-3074.

Torrence, Ch. and G.P. Compo. 1998. ”A practical guide to wavelet analysis.” *Bulletin of the American Meteorological society*, Vol. 79, No. 1, 61-78.

Walnut, D.F. 2013. *An introduction to wavelet analysis.* Springer Science & Business Media.

Organisation for Economic Co-operation and Development: National Accounts [online database]. (2017) [cit. 2017-01-12]. Available at: http:\\stats.oecd.org \Index.aspx?DatasetCode=SNA_TABLE1.

AUTHOR BIOGRAPHIES

EVA KLEJMOVÁ received her Masters degree in Electrical Engineering from Brno University of Technology in 2014. At present she is a Ph.D. student at the Department of Radio Electronics, Brno University of Technology.

JITKA POMĚNKOVÁ received a Ph.D. degree in applied mathematics at Ostrava University in 2005, a habilitation degree in Econometric and operational research at Mendelu Brno in 2010. Since 2011 she has been the Senior researcher at the Department of Radio electronics, Brno University of Technology.

JIŘÍ BLUMENSTEIN received his Ph.D. degree in electrical engineering at Brno University of Technology in 2013. In 2011 he worked as a researcher at the Institute of Telecommunications, Vienna University of Technology, Austria. Nowadays, he is a researcher at the Department of Radio Electronics, Brno University of Technology.

A MARGIN CALCULATION METHOD FOR ILLIQUID PRODUCTS

Marcell Béli
E-mail: beli.marcell@gmail.com

Csilla Szanyi
KELER CCP
Rákóczi street 70-72. Budapest,
1074, Hungary
E-mail:
szanyi.csilla@kelerkszf.hu

Kata Váradi
Department of Finance
Corvinus University of Budapest
Fővám square 8. Budapest, 1093,
Hungary
E-mail: kata.varadi@uni-corvinus.hu

KEYWORDS

Margin, central counterparty, illiquidity, IPO, counterparty risk, Value-at-Risk.

ABSTRACT

The role of the central counterparties (CCPs) on the market is to take over the counterparty risk during the trading on stock exchanges. CCPs use a multilevel guarantee system to manage this risk. The margin has a key role in this guarantee system, and the paper will focus only on this level. The main motivation of this paper is to introduce a potential margin calculation method which is compliant with the EMIR regulation and also does not put unnecessary burden on the market participants. We will introduce this method for two special type of products: (1) the illiquid products and (2) for the case of initial public offerings (IPOs). The specialty of these two product types, that there is no available historical time series of the securities' prices, so no risk management models can be used by the CCPs to calculate the margin.

REQUIREMENTS OF THE REGULATOR AND MARKET PARTICIPANTS

The role of the central counterparties is crucial on the financial markets since all trades on stock exchanges are being settled through CCPs. In case of a trader's default, the CCP ensures that the trade will be fulfilled for the other party. In order to guarantee this settlement, a CCP must have a waterfall system of guarantees, in which margin has a notable weight. Since CCPs' effect on the market stability is important from risk point of view, the regulators turned towards them lately. According to this the European Parliament and Council has launched in 2012 the EMIR (European Market Infrastructure – 648/2012/EU) regulation. EMIR and its supplementation, the Technical Standard (TS – 153/2013/EU) containing the following requirements regarding the margin calculation method of the CCPs (EMIR Article 41, TS Chapter VI):

- General assumptions: 'Margin shall ensure that a CCP fully collateralises its exposures with all its clearing members, … at least on a daily basis. A CCP shall adopt models and parameters in setting its margin requirements that capture the risk characteristics of the products cleared and take into account the interval between margin collections, market liquidity and the possibility of changes over the duration of the transaction.' (EMIR, Article 41., 2012)
- Liquidation period: at least 'two business days for financial instruments other than OTC derivatives.' (TS, Article 26, 2013)
- Confidence interval: 'for financial instruments other than OTC derivatives, 99%.' (TS, Article 24, 2013)
- Portfolio margining: 'a CCP may calculate margins with respect to a portfolio of financial instruments provided that the methodology used is prudent and robust.' (EMIR, Article 41., 2012)
- Look-back period: 'Initial margins should cover …exposures resulting from historical volatility calculated based on the data covering at least the latest 12 months … including periods of stress. Margin parameters for financial instruments without a historical observation period shall be based on conservative assumptions.' (TS, Article 25, 2013)
- Procyclicality: 'Applying a margin buffer at least equal to 25% of the calculated margin which allows to be temporarily exhausted in periods, where calculated margin requirements are rising significantly.' (TS, Article 28, 2013)

In the literature several risk measures exist that quantify the risk, and could be applied by CCPs in order to fulfil the regulatory requirements. The two most common models are the Value-at-Risk (VaR) (Jorion, 2007) and Expected Shortfall (ES) (Acerbi – Tasche, 2002, Acerbi et al. 2001) models. Both of the models have their advantages and disadvantages, for example the advantage of the VaR model, that it is easy to interpret; less data is enough to calibrate the model; it is not sensitive to the outlier data; easy to backtest (Acerbi – Székely 2014, Yamai – Yoshiba 2005), and it is elicitable (Ziegel 2014, Gneiting 2011). While the advantage of the ES is that it is coherent (Artzner et al. 1997, 1999, Pflug 2000, Frey – McNeil 2002, Acerbi – Tasche 2002), and can handle the fat tail risk (Yamai – Yoshiba 2005). Most of the CCPs uses these measures in their margin calculation models. For example KELER CCP applies the VaR model in their risk management system, with the following VaR parameters in order to fulfil the requirements of the

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

regulators: minimum holding period is 2 days, confidence level is 99%, the look back period is at least one year, and the procyclicality buffer is 25% (KELER CCP, 2017).

Besides the regulatory requirements, there are needs of the market participants as well. These needs were identified by Béli – Váradi (2016), and they also provided a solution for the margin calculation, which calculation will be introduced in more details in the next chapter, now only the market needs and the solution are shown briefly:

- stable margin: using a margin band.
- easy to reproduce the margin, so only a few expert decision should be in the calculation: (1) creating margin groups, in which the assets have the same parameters; (2) using liquidity and expert buffers besides the procyclicality buffer, based on the result of the backtest; (3) defining stress in order to be able to calculate the look back period objectively.
- the margin should follow market trends efficiently: using exponentially weighted moving average (EWMA) standard deviation besides the equally weighted standard deviation during the calculation of the VaR model.
- automatic and objective procyclicality buffer management: procyclicality buffer exhaustion and build back based on the relative relationship of the two standard deviations (EWMA and equally weighted standard deviations).

This paper will be built on the model of Béli – Váradi (2016), so in the next chapter we will show the model in more details. The new findings in our paper are, that we will show how that model can be easily used for illiquid products and in the case of IPOs. We have chosen this model, because it fulfils every regulatory requirements and every need of the market participants.

MARGIN CALCULATION METHODOLOGY

In the model of Béli – Váradi (2016) the risk measure is the VaR, calculated with a delta-normal method (Jorion, 2007), where the assumption is that the logreturn of an asset is normally distributed. The parameters that are needed for the VaR model is the mean and the standard deviation. Since they use daily returns in their calculation, they assume that the mean is 0, while the standard deviation is being estimated from the one year look back period (assuming that it contained a stress event) in two ways, once an equally weighted standard deviation and once an EWMA weighted standard deviation. They always use the one, which gives the smaller VaR value based on Equation 1. Then they calculate the VaR for prices as well, according to Equation 2.

$$VaR_t^{yield} = min(\sigma^{equal} \cdot N^{-1}(99\%); \sigma^{EWMA} \cdot N^{-1}(99\%)), \quad (1)$$

where σ^{equal} is the equally weighted standard deviation, σ^{EWMA} is the EWMA weighted standard deviation, N^{-1} is the inverse of the normal distribution's cumulative distribution function, and VaR_t^{yield} is the Value-at-Risk at day t, calculated on the logreturn basis, which means that this is the maximum loss one can have on a daily basis on a 99% significance level, expressed in logreturn.

$$VaR_t^{price} = -P_t + P_t \cdot e^{\sqrt{T} \cdot VaR_t^{yield}}, \quad (2)$$

where T is the liquidation period, while P_t is the price of the asset at time t, and VaR_t^{price} is the Value-at-Risk at the day t, calculated on the price level. It means that this is the maximum loss one can have on a 2 days basis (requirement of the regulator) on a 99% significance level, expressed in HUF.

After this, they increase the value of the VaR with liquidity- and expert buffers according to Equation 3, which buffers change between every margin groups – the assets are being grouped into different margin groups in order to have as unified buffers as possible. The more risky an asset, the higher these buffers will be.

$$KSzFmargin_t = VaR_t^{price} \cdot (1 + \varphi) \cdot (1 + \theta), \quad (3)$$

where φ is the liquidity buffer, while θ is the expert buffer.

The next step is, that the procyclicality buffer is being taken into account as well, based on Equation 4, where π is the procyclicality buffer.

$$PROmargin_t = KSzFmargin_t \cdot (1 + \pi), \quad (4)$$

Based on the regulation the procyclicality buffer can be exhausted if the margin would change notable due to market conditions. Béli – Váradi (2016) worked out a method, in which they exhaust and build back the procyclicality buffer in an objective way, and by keeping the margin stable (more details in Béli – Váradi (2016)). The following Equations 5, 6 and 7 show the method, which is based on the relative relationship between the equally and EWMA weighted standard deviation. If the EWMA standard deviation is higher than the equally weighted, the buffer can be exhausted, while the equally weighted is higher with 25% than the EWMA, then the buffer should the built back fully into the value of the margin, which will be called MINmargin.

$$margin_t^{pro-exhaustion} = max(margin_{t-1}; KSzFmargin_t), \quad (5)$$

$$margin_t^{pro-build\ back} = min\left(margin_t^{pro-exhaustion}; PROmargin_t\right) \quad (6)$$

$$MINmargin_t = if\left(\begin{array}{c}\left(\sigma_{EWMA} \cdot max\left(\frac{margin_{t-1}}{KSzFmargin_t}; 1\right) > \sigma\right); \\ margin_t^{pro-build\ back}; PROmargin_t\end{array}\right) \quad (7)$$

The last step is, that the margin should be stabilized as much as possible, so they use a margin band – which is a certain percent above the MINmargin, and it is called MAXmargin – and till the margin do not reach one of the bands, that margin, which is effective on the market, will not be changed. In their model they have shown how it is working in case of liquid Hungarian stocks. We will introduce how it works in case of illiquid stocks, and IPOs. Their result can be seen in Figure 1 and 2 for OTP stock. On Figure 1 the relation of the two standard deviations can be seen, while on Figure 2 the margin calculation can be seen on different levels as it was introduced in the equations above.

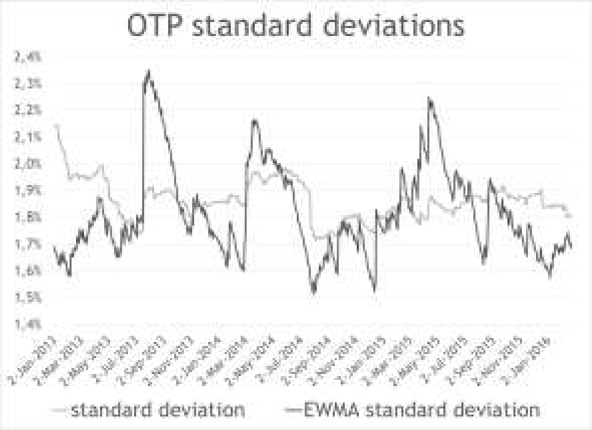


Figure 1: equally weighted and EWMA weighted standard deviation of the OTP

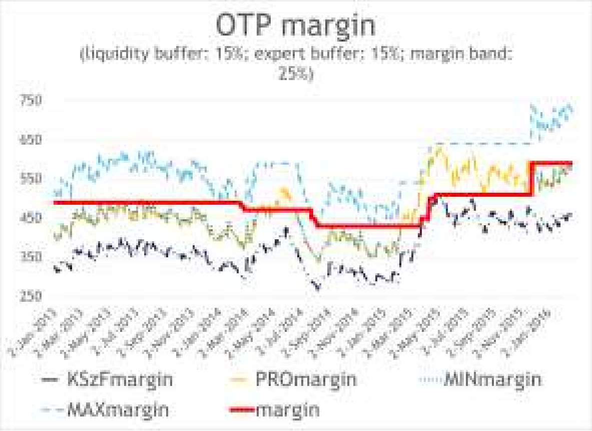


Figure 2: OTP margin

From Figure 1 and 2 it can be seen, that the procyclicality buffer was exhausted in the case when the EWMA standard deviation was greater than the equally weighted standard deviation. For example in the period from the beginning of May 2015, the PROmargin is 'visible', which means, that the MINmargin is lower than the PROmargin, so, the procyclicality buffer is exhausted. Also it can be seen, that the margin was not changed as often as the KSzFmargin, PROmargin, MINmargin and MAXmargin are changing. It only changes when the margin reaches one side of the margin band (the MINmargin or the MAXmargin).

MARGIN FOR ILLIQUID PRODUCTS AND IPOS

Under illiquid products we mean those products, that didn't have Instrument Liquidity Measure (ILM) – a weighted spread measure, calculated by KELER CCP on an intraday basis. It is the same measure as the Budapest Liquidity Measure (BLM) (more information about BLM can be found in Kutas – Végh (2005) – for more than 200 days in the last 250 trading days. Based on this, these assets do not have a historical time series of closing prices, since there was no trading at all.

Under IPO we mean the Initial Public Offering of a product, which is the first time in a company's life, when they sell their shares for investors through a stock exchange for institutional investors and also for private investors. So based on this, there is no price history for these assets, since it was not publicly traded before the IPO.

In case of illiquid securities and IPOs, in order to deal with the lack of price history, we base margin computation on the past time series of a market product that can be considered average and represents the market, to replace the insufficient time series of these products. We have chosen the index's time series for this purpose. In the case of the Budapest Stock Exchange's (BSE) products the BUX index's time series will be used, since the illiquid products and IPOs which we are analysing are traded on the BSE, and the index on that market is the BUX index. As margin determination is based on standard deviation parameters, as this is required to determine VaR, the BUX's historical logreturns provide this standard deviation parameter. However, as the BUX shows the movement of the entire market, where liquid, less risky products are overrepresented, using such data to determine the margin would result in significant underestimation of risk. To solve this we will increase the margin parameters of the model, the liquidity-, expert buffers and margin band to 100%, but calculating VaR based on the data of the BUX index. For comparison, the parameters Béli – Váradi (2016) have used in case of the OTP are: liquidity buffer: 15%; expert buffer: 15%, margin band: 25%.

The specialty of the IPOs compared to the illiquid products, that one year after the IPO – if it is not an illiquid stock of course – there will be enough data to calculate margin based on the VaR model with those parameters that belong to the margin group, in which the asset is going to be grouped into. We decided to do it this way, because it is the most prudent approach by a CCP to assume that an IPO asset is in the lowest risk category, and handle it, as a risky illiquid product.

A key element of the methodology is that the starting margin value can be objective in the case of IPOs, if on the first day the margin is determined as the arithmetic mean of MINmargin$_t$ and MAXmargin$_t$ computed in line with the above methodology (Equation 8).

$$margin_1 = \frac{MINmargin_1 + MAXmargin_1}{2} , \qquad (8)$$

However, the computation of MINmargin$_t$ is different than in the basic methodology, due to the lack of data, thus on the first day the PROmargin$_t$ value will be the MINmargin$_t$ value, according to Equation 9.

$$MINmargin_1 = PROmargin_1 , \qquad (9)$$

It is important to note that yield based VaR is computed from BUX values, but the price of the security is used to calculate the price based VaR, this is the reason why the margin will be different product-by-product.

There are no other changes in the methodology. Figure 3 and 4 illustrate the margin used in the case of the Update stock's initial public offering under the new methodology. Figure 3 shows the value of the standard deviations that is the basis of the margin, while on Figure 4 the margin is presented.

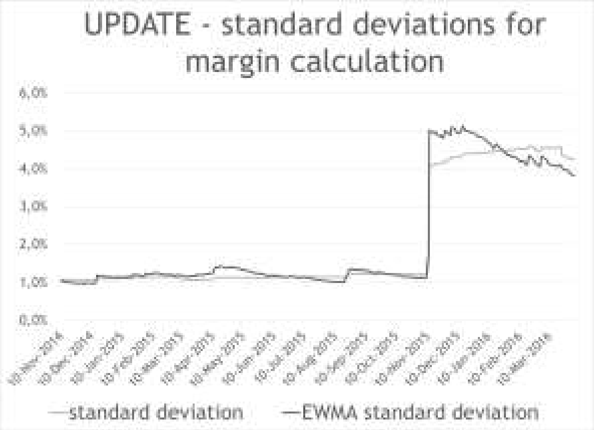


Figure 3: equally weighted and EWMA weighted standard deviation for UPDATE margin calculation

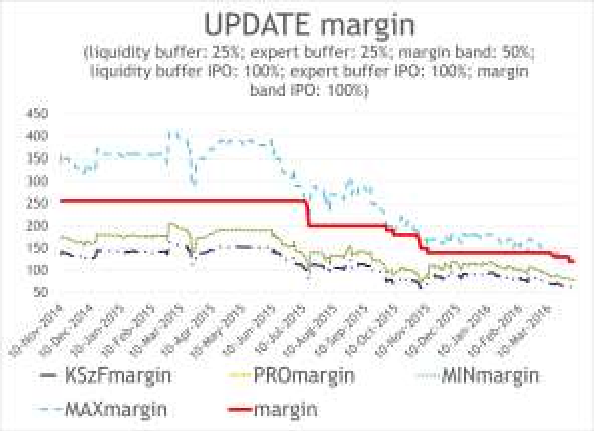


Figure 4: Update margin

The time series contain not only the IPO period, and the following one year, where the margin is based on the BUX index's parameter, but the period after the first year is over, and the parameters are being estimated from the own price history of the Update.

As one year after the IPO the Update equities were listed in the Standard category on the BSE (based on the categorization of BSE (2016)), the drop in the margin after one year is due to the major decrease in buffer values. The Figure 3 shows that there was a big jump in the standard deviation data (November 2015) when standard deviation was computed based on the own past time series of the security, not the BUX index's time series anymore. This increase in the value of the standard deviation should have been reflected in the increase of the margin in Figure 4, but it was offset by the decrease of buffers. Based on this we can conclude that it may be justified to use 100% buffer values in the case of all IPOs, as BUX presumably have lower risk than a newly listed product.

In Figure 5 we show the standard deviations of the BUX index for the whole analysed period. These standard deviations are needed for the margin determination for the illiquid products.

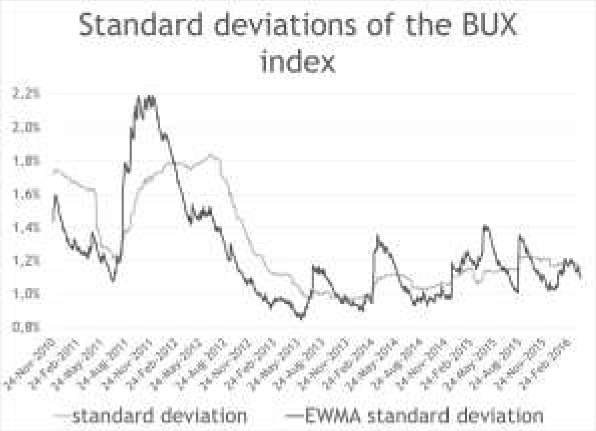


Figure 5: equally weighted and EWMA weighted standard deviation of the BUX index

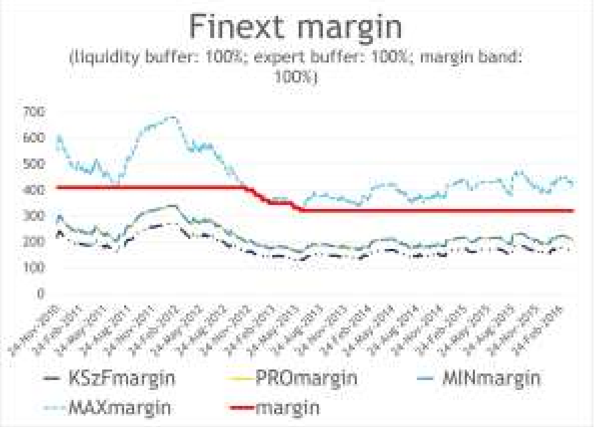


Figure 6: Finext margin

Figures 6 and 7 illustrate the margin determined in the case of illiquid products, with 100% buffers, for Finext and Őrmester. In addition to being illiquid, the Finext IPO took place in the analysed period shown in the figure.

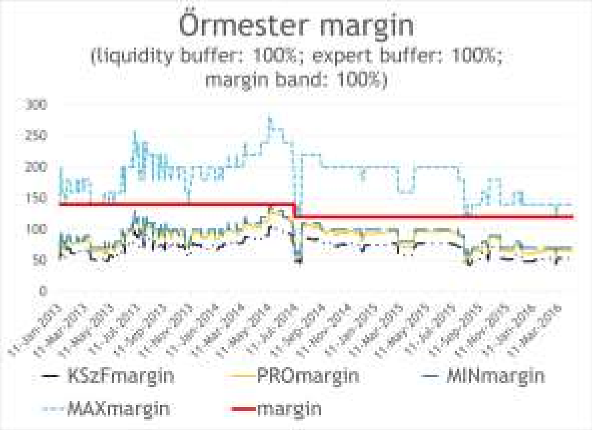


Figure 7: Őrmester margin

BACKTEST

To analyse whether the margin methodology we are using is appropriate for margin determination or not we are using backtests. A model can be considered good, if the backtest gives back the same result as the level of the significance level, so the 99% in our model. We have to check whether the actual price change have exceeded the value of the margin or not for each day in the last 250 days. If the price change exceeds the value of the margin in more than 99% of the days, then the model can be considered as not adequate.

We perform the backtest in two different ways for each product. On one hand we check how many times the actual daily price change exceeded the margin applied in the past 250 trading days, and on the other hand we check how many times the actual daily price change exceeded the VaR value. In the case of VaR, the VaR computed with equally weighted standard deviation and the VaR with EWMA weighting are not checked separately, but always the lower value is used in the back testing, as this is the one used for margin determination also.

However, a high knockout number was expected in the case of illiquid securities and IPO during the backtesting of the VaR values, as VaR parameters are determined from the index time series that have a lower risk than in the case of illiquid securities and IPO, which needs to be reflected in the back testing. The use of higher buffers is designed to manage this risk.

The result of the margin backtest was 100% in the case of all the three analysed asset, as it can be seen in Figures 8, 9 and 10. In these figures the columns show the actual price changes, the two lines in the bottom of the figures are the VaR values, while the uppermost line is the margin. The columns never exceeds the line of the margin.

In case of Update's backtest in Figure 8 it can be seen that the actual price change exceeded the value of the VaR for several days. Altogether it happened 14% of the cases, so the result of the backtest is only 86%, which we have expected before the backtest, namely to not to reach 99%. So this confirms in the case of these products, that 100% buffers and the margin band were needed, and was sufficient both for illiquid contracts and IPO.

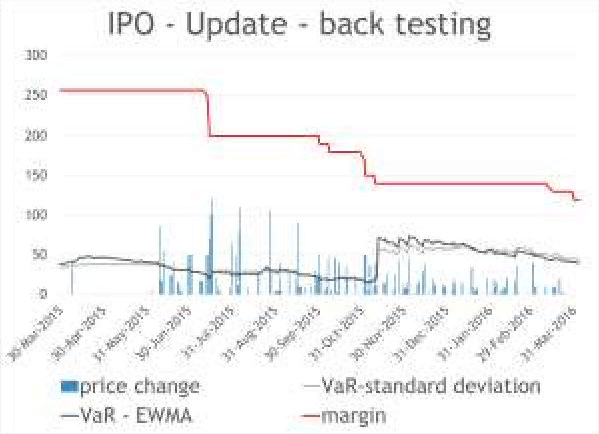


Figure 8: Update backtest

In Figure 9 we see the same as in the case of Update. On a VaR level the model was good only in 92.8% of the cases, but on margin level (uppermost line in Figure 9) the model was always good, the price change never exceeded the margin. So the model we have built is adequate for margin calculation purposes.

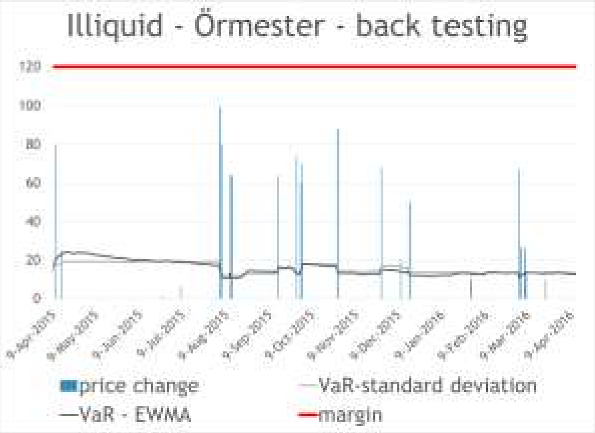


Figure 9: Őrmester back testing

The backtesting of Finext is not possible according to Figure 10, since there was no trading activity in the security at all during the whole analyzed period.

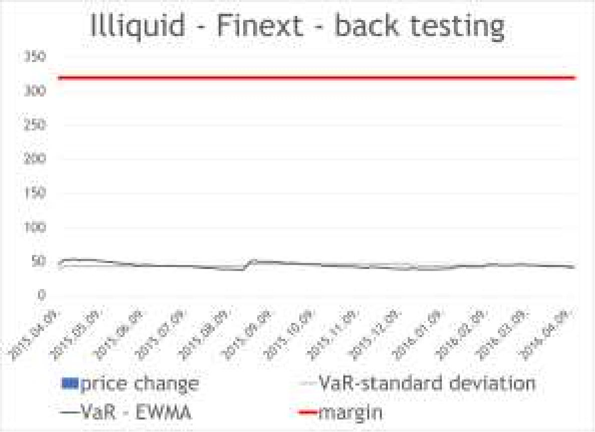


Figure 10: Finext back testing

CONCLUSION

We have built a margin calculation model for illiquid products and IPOs based on the model of Béli – Váradi (2016). Our model is easy to use, and understand, moreover the same methodology can be used as for liquid products. The new result of our model was, that we have estimated the parameters of the risk measure from an index's time series, since for illiquid products and IPOs we do not have adequate data for parameter estimation.

Also we have shown, that in case of IPOs it is necessary to handle the products the same way as we do in the case of illiquid products for risk reduction reasons.

To handle the high risk level of these products, we have used 100% buffers, and margin band values to handle risk, which was proven by the backtest.

REFERENCES

Acerbi, C., Nordio, C. and Sirtori, C. (2001): Expected shortfall as a tool for financial risk management. *arXiv preprint cond-mat/0102304*.

Acerbi, C. and Székely, B. (2014): Backtesting Expected Shortfall. MSCI working paper, 2014.

Acerbi, C. and Tasche, D. (2002): On the coherence of expected shortfall. *Journal of Banking & Finance*, Vol.26. No.7., pp.1487-1503.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1997): Thinking coherently. *Risk* 10, pp. 68–71.

Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999): Coherent measures of risk. *Mathematical finance*, Vol. *9*. No.3., pp.203-228.

Béli, M. and Váradi, K. (2016): Alapletét meghatározásának lehetséges módszertana [A possible margin determination methodology] Conference presentation, Győr, Hungary, at PRMIA conference, 21st October, 2016.

EMIR – European Market Infrastructure Regulation: Regulation (EU) No 648/2012 of the European Parliament and of the council of 4th July 2012 on the OTC derivatives, central counterparties and trade repositories (EMIR - European Market Infrastructure Regulation) Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32012R0648&from=EN downloaded: 8th April 2016.

Frey, R. and McNeil, A.J. (2002): VaR and expected shortfall in portfolios of dependent credit risks: conceptual and practical insights. *Journal of Banking & Finance*, Vol. 26. No.7., pp.1317-1334.

Gneiting, T. (2011): Making and evaluating point forecasts. *Journal of the American Statistical Association*, Vol. 106. No. 494., pp.746-762.

Jorion, P. (2007): *Value at risk: the new benchmark for managing financial risk.* Vol. 3. New York: McGraw-Hill.

KELER CCP homepage (2017): https://english.kelerkszf.hu/Risk%20Management/Multinet/Initial%20Margin/ downloaded: 7th February, 2017.

Kutas, G. and Végh, R. (2005): A Budapest Likviditási Mérték bevezetéséről. A magyar részvények likviditásának összehasonlító elemzése a budapesti, a varsói és a londoni értéktőzsdéken [Introduction of the Budapest Liquidity Measure]. *Közgazdasági Szemle (Economic Review-monthly of the Hungarian Academy of Sciences)* Vol. 52. No. 7. pp. 686-711.

Pflug, G.C. (2000): Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization.* pp. 272-281. Springer US.

TS – Technical Standard: Commission delegated regulation (EU) 153/2013 of 19th December 2012 supplementing Regulation (EU) No 648/2012 of the European Parliament and of the Council with regard to regulatory technical standards on requirements for central counterparties. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:052:0041:0074:EN:PDF downloaded: 8th April 2016.

Yamai, Y. and Yoshiba, T. (2005): Value-at-risk versus expected shortfall: A practical perspective. *Journal of Banking & Finance* Vol.29. No.4. pp. 997-1015.

Ziegel, J.F. (2014): Coherence and elicitability. *Mathematical Finance*.

AUTHOR BIOGRAPHIES

MARCELL BÉLI was a Market Risk Manager at KELER CCP. He graduated at Corvinus University of Budapest in 2014 and obtained the PRM™ designation also in 2014. In his earlier research he investigated the Equity Premium Puzzle and connected areas.

CSILLA SZANYI is a senior risk controller at KELER CCP. She majored in finance at Corvinus University of Budapest, at the Department of Finance in 2010. Her main responsibilities at KELER CCP are market risk management on the cleared capital and energy markets, and the evolvement of the risk management framework considering the compliance with EU and Hungarian regulations.

KATA VÁRADI is an Associate Professor at the Corvinus University of Budapest (CUB), at the Department of Finance. She graduated also at the CUB in 2009, and after it obtained a PhD in 2012. Her main research area is market liquidity, bonds markets, capital structure of companies and risk management.

MODELLING CIVIL SOCIETY’S TRANSFORMATIONAL DYNAMISM AND ITS POTENTIAL EFFECTS

Jozsef Veress
Institute of Informatics
Corvinus University of Budapest
H-1093, Budapest, Hungary
E-mail: veress.jozsef@yahoo.com

KEYWORDS
civil society self-organizing, sharing, on demand economy, regulation, hybrid modelling,

ABSTRACT

The paper proposes three ‘non-conventional’ patterns to model using System Dynamics: (i) transformational dynamism of civil society organizations; (ii) regulations effectively limiting disruptive effects of platform firms; (iii) policies facilitating new patterns of value creation through genuine sharing. The combination of SD and Agent Base Modelling may improve the effectiveness of analysing agency as complex interplay among “high-order, nonlinear, feedback systems” and actors who’s interactions co-create them. Such dynamic hybrids may enhance the effectiveness of modelling feedbacks among the civil society entities’ transformational dynamism and legislative and policy processes. Models may contribute to regulations limiting disruptive effects of platform firms and to policies enhancing alternative patterns of value creation in genuninely sharing economy similar to platform cooperativism and Commons Based Peer Production.

INTRODUCTIONS

“We do not live in a unidirectional world in which a problem leads to an action that leads to a solution. Instead, we live in an on-going circular environment. Each action is based on current conditions, such actions affect future conditions, and changed conditions become the basis for later action. There is no beginning or end to the process. Feedback loops interconnect people. Each person reacts to the echo of his past actions, as well as to the past actions of others”(Forrester, 1998:2-3). “…interactions in a social system correspond to those purposeful decision making processes that convert information into action - such is the definition of “decision making” of Forrester (1961) - through the exchange of resources, materials, information, meanings, communications, etc.”(Olaya and Gomez-Quintero, 2016:2). Since complex non-linear feedback systems unfold as continuous aggregation of the actors’ interactions System Dynamics facilitate to explore agency as complex interplay among “high-order, nonlinear, feedback systems” (Forrester, 1991) and the actors who participate in and also co-create them (Olaya and Gomez-Quintero, 2016). Goals of modelling process are inherently social (Vriens and Achterbergh, 2006) so System Dynamics can serve as effective instrument to elaborate and shape policies and legislative acts. Hybrid variants combining System Dynamics with Agent Base Modelling enable to model policy and legislation processes (Misuraca and Kucsera, 2016), their interplay with transformational dynamism of civil society entities.

MODELLING AND SIMULATION OF THE TRANSFORMATIONAL DYNAMISM OF CIVIL SOCIETY ORGANIZATIONS

The System Dynamics’ focus on underlying non-linearity makes it a proper analytic tool to explore, model and simulate sources, mechanisms, and outcomes of the civil society entities’ transformational dynamism. The volunteers’ (self-) empowering non-wage work unfolds as passionate and sharing co-creation improving life quality as the analysis of self-organizing communities exemplifying civil society entities (Veress, 2016) indicates. The interactions -, which presuppose and simultaneously regenerate motivation and trust - aggregate into the civil society entities’ and networks’ continuous emergence. The volunteers’ interactions generate multidimensional feed backing change processes. Their mutually catalytic character turns them into self-re-enforcing feedback loops and facilitates their aggregation into continuous emergence of civil society entities. Consequently, the civil society entities are “high-order, nonlinear, feedback systems” (Forrester, 1991) - subject to System Dynamics and the deployment of SD facilitates more ‘fine grained’ elaboration on diverse aspects of the transformational dynamism of civil society entities.

Motivation and self-communication in community

The community members’ voluntary cooperative interactions improve their perceived life quality in multiple ways. Their readiness to volunteer feeds back

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

with self-communication, which "…multiplies and diversifies the entry points in the communication process. This gives rise to unprecedented autonomy for communicative subjects to communicate at large"(Castells, 2009:135). The self-communication facilitates to recognize mutual benefits which collaboration provides. It generates awareness of an associational - rather than competitive - advantage. The awareness (R1) enhances motivation to volunteer (Figure 1). Growing motivation catalyzing more intense contributions to voluntary activities enables to increase the rate of cooperative interactions (R2) creating life quality improvements. Growing rates of cooperative interactions presuppose self-communication, catalyze its growing intensity (R3). Since these interplaying phenomena are mutually catalytic their feedbacks may aggregate into self-reinforcing loops (Figure 1).

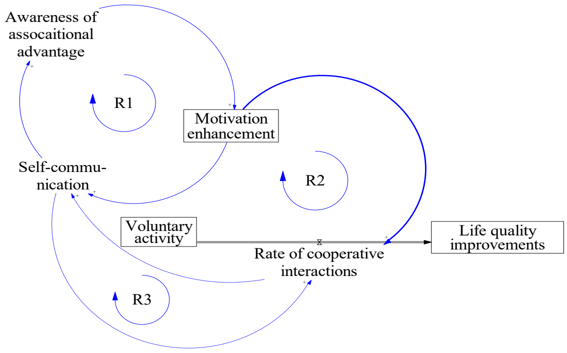


Figure 1: Enhancement of the Motivation to Volunteer

Social capital and trust (re-)creation

The community members' intertwined intra- and interpersonal dialogues carry out sense and decision making (Stacey, 2000). These dialogues enact various institutional settings and aggregate into self-communication. The association-prone institutional settings in turn may amplify motivation to join and contribute to cooperative interactions (R4) (Figure 2).

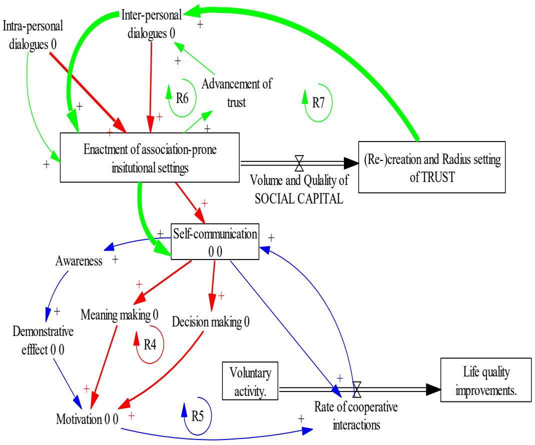


Figure 2: Social Capital and Trust (Re-)Creation

The intertwined dialogues aggregating into self-communication enable cooperative interactions catalyzing life quality improvements. The self-comunication re-creates the volunteers' awareness of associational advantage their cooperation brings about. Consequently, the awareness creates (also serves as) demonstrative effect which (re-) generates and amplifies motivation to volunteer. The enhanced motivation catalyzes participation in collective efforts and increases the rate of cooperative interactions (R5). The individuals who voluntarily join to a community serving as domain of cooperative efforts obviously have an inclination to collaborate. Due to such positive disposition toward cooperation their intra-personal dialogues enact (primarily) association-prone institutional settings. This constellation brings about the community members' readiness to (mutually) advance trust which is imperative to start to communicate (Luhmann, 1995). The (mutual) advancement of trust enables to launch inter-personal dialogue, which is generative and constitutive of their self-communication (R6).

Association-prone institutional settings which inter-twined intra- and inter-personal dialogues enact play multiple important roles. They serve as social capital, as "…informal norm that promotes cooperation between two or more individuals… [is] instantiated in an actual human relationship… [generates and sets the radius of] trust …epiphenomenal, arising as a result of social capital but not constituting social capital itself"(Fukuyama, 1999:1) (R7).

Consequently, the assocaition-prone institutional settings (i) serve as social capital which re-generates trust and sets its radius; (ii) 'catalyze and calibrate' self-communication; (iii) facilitate the volunteers' communicative interactions and those aggregation into - continuous emergence of - their community, i.e. (iv.) serve as institutional-type, soft organizing platforms. Actively catalyze multi-dimensional process feedbacks which carry out, aggregate into continuous (re-) emergence of the self-organizing community. This constellation provides the first instant dynamic character of communities as professor Gábor pointed out commenting the research.

Enhanced effectiveness of resource enactment

The volunteers' interactions usually are of small scale as empirical data and literature indicate. Benkler (2011) coins as "modularity of contributions" such limitation of the particular contributions' size. Since such modularity allows minimizing the particular tasks' resource intensity the volunteers are ready to take care also about resourcing. I.e. their interactions carry out also identification, accession, mobilization and sharing of resources. The interactions and the resourcing are in a sense identical voluntary activities. Consequently, the cooperative logic aims to decrease both individual tasks and their resource intensity. The decrease of burden related to particular tasks in turn may increase the number of contributors, i.e. increases both the frequency and overall number of contributions. Due to low resource intensity of tasks more people volunteer and mobilize limited volumes of required resources. Paradoxically by decreasing the tasks' resource requirements the modularity of contributions improves the effectiveness of resourcing and extends the mobilized resources' overall volume (B1) (Figure 3).

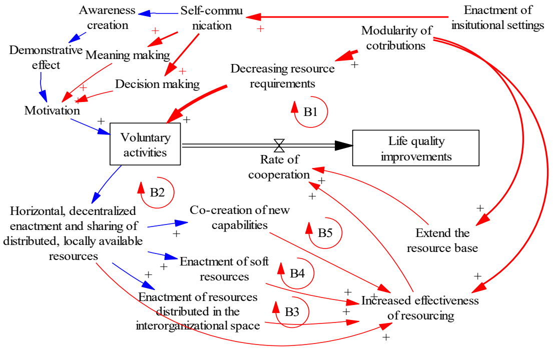

Awareness creation
Self-commu nication
Enactment of insitutional settings
Modularity of cotributions
Demonstrative effect
Meaning making
Decision making
Decreasing resource requirements
Motivation
B1
Voluntary activities
Rate of cooperation
Life quality improvements
B2
Horizontal, decentralized enactment and sharing of distributed, locally available resources
Co-creation of new capabilities
B5
Extend the resource base
Enactment of soft resources
B4
Enactment of resources distributed in the interorganizational space
B3
Increased effectiveness of resourcing
+

Figure 3: Enhanced Effectiveness of Resourcing

The extensive sharing makes obsolete to establish and maintain resource ownership. The horizontal and decentralized patterns of resourcing where interactions enact and share resources simultaneously allow avoiding accumulation and redistribution through vertical hierarchies (B2). The horizontal and decentralized enactment and sharing of distributed and locally available resources provide multiple ways to expand and upgrade collective resource base and improve the effectiveness of resourcing. Cooperating volunteers through vivid networking may mobilize also resources located in the 'inter-organizational space'. Although these frequently are dispersed into small quantities their overall volumes may be significant (B3). The community members frequently capitalize on various 'soft resources' similar to knowledge, information, creativity, and psychological energies (B4). Since these are non-depletable and non-rivalrous (Bollier, 2007:28) they are also multipliable or even self-multiplying as knowledge plausibly demonstrates. It is freely sharable; its pieces could be recombined into new knowledge. The knowledge through its implementation becomes even more 'voluminous' and of higher quality instead of decreasing and becoming 'worn out'. Additionally, the knowledge often may serve as 'ultimate substitute' enabling to decrease the required volume of or fully replace other resources.

The collaborating community members may co-create new capabilities improving collective resource access (B5). The growing awareness of feed backs among improved effectiveness of resourcing and increasing associational advantage may strengthen the motivation to cooperate, i.e. unleashes "cooperation trap" (Csányi, 1989). Self-enforcing feedback loops of innovative capability co-creation, improved (effectiveness of) resourcing, and enhanced functional (rather than organizational) complexity are characteristic also for broader evolutionary tendencies (Nowak, 2006). They may operate as driver(s) of emerging, self-organizing evolutionary tendencies and developments. In explored communities such feedback loops are important drivers of continuous self-organizing emergence providing robust transformational dynamism (Veress, 2016).

Recursive, multi-staged deployment of System Dynamics enables to explore sources of the dynamism of communities representing broad array of civil society organizations. Causal and stock and flow diagrams may facilitate to shed new light on the civil society organizations' dynamism and broader social transformational effects. The SD facilitates to analyse (i) the civil society players' ability to effectively fulfil various and altering needs in context of rapidly changing social dynamics; as wells as (ii) the role of civil entities in affecting and shaping social dynamics. Recursive mapping of causal loops and feedbacks among levels and rates may help to identify ways and tools of quantifying and measuring variables, construct and run quantitative model(s). The 'inverse logic' of modelling efforts may facilitate recursive fine-tuning of analytic tasks and tools, identify relevant data sources. Such attemtps should consider also effects of time distribution and relevance of metamodeling.

CAN SYSTEM DYNAMICS CONTRIBUTE TO POLICIES AIMING TO ENHANCE GENUINELY SHARING ECONOMY?

Sharing economy became the "…corporatization of the sharing movement ...sharing evolves from a peer-to-peer enterprise to a place where established market participants seek to assert themselves in the sharing economy's new domains"(Miller, 2016:149). The platform firms' business models robust disruptions among other on labour market. In the US during the last decade the "…share of workers in alternative work arrangements in their main job increased by 5.7 percentage points (or by over 50 percent) from 2005 to 2015. *A striking implication …is that all of the net employment growth in the U.S. economy from 2005 to 2015 appears to have occurred in alternative work arrangements* [italics in original]" - conclude Katz and Krueger (2016:7). These developments amplify robust disruptive tendencies created by "large corporations [which] have shed their role as direct employers of the people responsible for their products, in favor of outsourcing work to small companies that compete fiercely... The result has been declining wages, eroding benefits, inadequate health and safety conditions, and ever-widening income inequality"(Weil, 2014). "Low labour costs discourage investments in labour-saving technology, potentially reducing productivity growth… the rapid technological progress can coincide with lousy growth in pay and productivity"(Avent, 2016:2).

The ballyhooed 'sharing economy' consists of 'hollow companies' "redesigning…corporate responsibility and accountability… [these start-up] companies are dramatically claiming a new corporate »right«: set up operations first…and figure out the laws and tax requirements late …Uber does pay federal corporate income tax on the considerable business earnings generated from its cut of each fare (about 25 to 30 per cent of the bill). But just like Apple, Google and other companies, Uber (as well as Airbnb) has constructed a complex web of 30 foreign subsidiaries and tax havens, many of them no more than mailboxes in the Caribbean, as a way to greatly reduce its US tax obligations…."(Hill, 2016:3).

Negative socio-economic consequences are even more explicit in long term. As Galbraith (1987:290-291) points out: "In the modern industrial economy production is of first importance not for the goods it produces but for the employment and income it provides. ...In the industrial countries most people, *when employed,* are not primarily preoccupied with the size of their income. …Their principal worry is the danger of losing all or most of their income - of losing employment and the consequent loss all or most of the means of their livelihood …factors affecting the security of employment are now socially far more important than those determining the level of reward. This being so in the present, so it will be in the future. …All suffering was identified with the interruption in …income - with unemployment… Time and increasing well-being will… overtake the concern about how their proceeds are distributed." As Galbraith (1987:285) emphasizes: "The great dialectic in our time is …between economic enterprise and the state. Labor and labor unions are no longer the primary enemies of the business enterprise and of those who direct its operations. The enemy, the wonderfully and dangerously rewarding role of military production apart, is government …that reflects the concerns of a constituency that goes far beyond the workers - a constituency of the old, the urban and rural poor, minorities, consumers, farmers, those who seek the protection of environment, advocates of public action in such areas of private default as housing, as transportation and health care, those pressing the case for education and public services in general. Some of the activities thus urged impair the authority or autonomy of the private enterprise, others replace private with public operation; all…are at cost either to the private enterprise or to its participants. Thus the modern conflict between business and government". This broader context enables proper interpretation of recent assertions of the White House chief strategist, Stephen K. Bannon "…declaring that the new administration is in an unending battle for "deconstruction of the administrative state"…"(Rucker and Costa; 2017).

No regulation! – key factor of the on demand economy's business model

The loud marketing of 'sharing economy' transform - rather deform - the concept and practice of sharing, as well as employment, and its regulation. "…Uber and its ilk offer …a nearly magical user experience, but their innovation lies just as much in evading regulations as in developing new technology. Behind the apps lies an army of contract workers without the protections offered to ordinary employees, much less the backing of a union. This…on-demand service economy …spreading market relations deeper into our lives. …new middlemen [are] sucking profits out of previously un-monetized interactions, creating new forms of hyper-exploitation spreading precarity…" (Ehmsen and Scharenberg; 2016).

The platform firms' business models create "platform monopolies in the absence of a physical infrastructure of their own… [by] running off *your* car, *your* apartment, *your* labour, *your* emotions, and importantly *your* time" - emphasizes Scholz (2016:2-3). These logistics companies decouple productivity from income by "[u]sing the language of entrepreneurship, flexibility, autonomy, and choice …[they] shift …the burden of the biggest risks of life: unemployment, illness, and old age from the employee…to a more contingent worker, the freelancer…independent contractor…or gig worker"(Scholz 2016:5-7). Its crucial to re-make "…labour protections, including a portable safety-net, for the digital age. …New forms of work, such as »crowd work«, require new regulations. Minimum standards for wages, health and safety, working hours and social security must be established to prevent this form of work from becoming exploitative and the jobs precarious. …all workers must be included in a new kind of portable and universal safety net, including solo self-employed persons and crowd workers, ensuring co-financing of their social security contributions by the businesses that hire them" - points out Hill (2016:13).

These companies are very consistent and successful in preventing and escaping regulation (Hill, 2016). By "…using their apps as political platforms ...activate their clients to oppose any regulatory efforts against them"(Scholz 2016:7). "…within the academic literature… has been almost no discussion of how the sharing economy businesses relate to existing local government regulatory structures …which is a surprise given that many sharing economy businesses have violated state or local government laws…" - points out Miller (2016:149). By analyzing "how, or if, the mass scale of the sharing economies' non-compliance with local government laws can be rectified" Miller (2016) capitalizes on empirical data from San Francisco and Portland (Oregon). He indicates that legislative efforts aiming to provide effective regulation for platform economy provide rather controversial outcomes.

The platform companies and their lobbists stress that "outmoded labor laws are imposing costs on the gig economy …introduce a great deal of uncertainty …discourage the creation of the flexible and varied job opportunities that Americans increasingly need …do a poor job of benefiting the workers they are intended to protect"(Kennedy, 2016:2). The expert of Washington based Information Technology and Innovation Foundation proposes three practical variants to handle the problem of regulation: "1. … to create a new category of worker, between full employee and independent contractor… 2. … Congress to revisit each of the country's major labor laws and carefully tailor them to achieve their specific goals. This would be ideal, but it would involve a long and difficult political process… 3. … to draft a carve-out for workers who depend on Internet platforms to find gig work …None of these would entirely solve the problem, because each state also has its own labor laws. Without similar changes at the state level, many of the benefits of reform would remain out of reach. But any of the three paths would jumpstart the process of updating U.S. labor law…"(Kennedy, 2016:2). From the platform firms' point of view a long process with no genuine outcome is the second best option, while the very best one remains: No regulation! Indeed, various US and EU legislative attempts reached at best partial and disputed solutions if any. However, as the Taiwan experience justifies effective regulations exist.

Platform democracy, crowd legislation in Taiwan

Taiwan in 2015 elaborated a comprehensive new legislation on activity of Uber and Airbnb on local market. "With this regulation, other Uber-like apps, some created by the civil society, are entering the market" - sums up Tang (2016) the outcome of mass deliberation exercises. "The hardest hack is to hack into the society, especially things concerning the government. Still, "g0v.tw", an open source community, has significantly empowered the civil society in Taiwan since 2012" - points out Yun-Chen Chien (2015) analysing "How Open Government Movement Has Made Civil Society in Taiwan". The g0v.tw, a self-organizing network facilitated the civil society efforts of self-empowerment. "Following the model established by the Free Software community over the past two decades, we transformed social media into a platform for social production, with a fully open and decentralized cultural & technological framework…" - sum up major milestones the members of g0v.tw (http://g0v.asia/tw/) by adding 2013 was the year of "Dismantling our Government and Building It Anew".

Describing what made "Uber respond to vTaiwan's coherent blended volition" Tang, (2016) the Digital Minister @ Taiwan points out: "Modeled after Cornell's RegulationRoom, vTaiwan is a g0v (gov-zero) project run by volunteers that works with… administration on crowdsourcing internet-related regulations". G0v.tw developed a technology enabled platform providing tools for mass deliberation. Volunteering facilitators successfully catalysed the emergence of "a coherent set of reflections, expectations and suggestions". The facilitators using AI supported systems focused on identifying mutually acceptable points of emerging consensus. The process aiming for coherence (rough consensus) not convergence (coordinated consensus) aggregated mutually acceptable constructs into "blended volition". Various groups of citizens "fully explored all aspects of each stage, before moving on to the next".

Transparency and mutual trust were preconditions and outcomes of cooperation. Technology platforms provided full transparency of data and relevant knowledge by enabling substantive debate. The readiness of the government and legislation to formalize, endorse and approve the co-created legal constructs and mechanisms enhanced participative democracy in Taiwan. "The key to the vTaiwan model lies in its "symmetry of attention" - points out Tang (2016). Through live streaming and remote participation citizens can see how stakeholders present their views, how much effort invest in the process. Actual face-to-face meeting agenda itself are crowdsourced by online discussion. The participants may exert genuine influence by discussing policies which are in the stage of problem-identification and the ministries are committed to give an official response within seven days to any question during discussion.

Successful mass-deliberation created a wish to raise on higher level "crowd legislation" efforts. "As vTaiwan went on to deliberate transnational issues such as Uber and Airbnb, this model was proven to be feasible. …Is there a way to institutionalize this model? …Taiwan's efforts for open government and public participation are at frontiers of the world. An innovative democratic system - born among conflicts and oppositions - can become a gift that we share with all humanity. We must keep accelerating our efforts" - conclude Tang and Kao (2016) discussing how to go beyond crowdsourced laws and participatory budgets.

Platform cooperativism-return to genuine sharing

The Internet offers technology background to renew value creation driven by genuine sharing and participation. "There isn't just one, inevitable future of work. Let us apply the power of our technological imagination to practice forms of cooperation and collaboration. Worker-owned cooperative could design …apps-based platforms, fostering truly peer-to-peer ways of providing services and things" - emphasizes Scholz (2016:2). "We need to build an economy and an Internet that works for all… take lessons from the long and exciting history of cooperatives and bring them into the digital age. …Platform Cooperativism…is an emerging economy …cooperatives employ more people than all multinationals combined ...It's too hard to fix what you do not own …[but] cooperative ownership models of the Internet would address many of these issues…"- argues Scholz (2016:10).

Worker-owned cooperatives "…can offer an alternative model of social organization to address financial instability. They will need to be (i) collectively owned, (ii) democratically controlled businesses, (iii) with a mission to anchor jobs, and (iv.) offer health insurance and pension funds and, a degree of dignity"(Scholz, 2014:7). The platform cooperativism is a "core location for development of new ideas in pursuit of an open social economy" of communities and commons (Benkler, 2016). Network pragmatism enables massive experimentation, rapid iteration utilizing knowledge generated by applied inquires of volunteer cooperators. Local communities know best about their needs, use reflection through practical experience - trial and error. Flexible organizations continually adapt and innovate while engaging with investor capital. They may withstand pressures arising from the logic of "tyranny of margin", i.e. the need to compete in market and maximize profits. There is sufficient room in current market situation "…platform cooperatives will neither kill nor be killed by investors firms..."(Benkler, 2016:1). Platform cooperatives may allow overcoming inequality caused by extreme extraction of wealth by the top 10 %.

A Platform Cooperativism Consortium adopted nine priorities for 2017: elaborate legal templates to help communities to launch platform co-ops; establish a donation channel for systematic fundraising; catalyze active networking among members of emerging ecosystem; map platform cooperative initiatives and projects; launch a European sister organization and support events in various cities; prepare a Massive Open Online Course about platform cooperativism economy (Llewellyn, 2017). A major challenges is to determine how to contribute to commons acting in "soft enclosures" by insulating populations from economically hostile surroundings. Diversity of organizational forms, an "organizational bricolage" in cooperation with solidarity economy and pro-commons movement facilitate to capitalize on alternative ways of value creation (Benkler, 2016).

Modelling and simulation in policy making and legislation

System Dynamics may facilitate in multiple ways to explore and re-describe the civil society organizations' robust transformational dynamics as interplay among and aggregation of feedback loops of nonlinear changes. Findings provided by deploying SD confirm outcomes of previous qualitative analyses capitalizing on methodological pluralism combining process and variance approaches (Veress, 2016). The paper assumes that legislative efforts regulating platform companies and on demand economy, and policies aiming to facilitate sharing economy related initiatives similar to platform cooperativism or Commons Based Peer Production may capitalize on models enabling to better understand transformational dynamism of civil society entities. Modelling and simulation of similar dynamic process constellations require "a methodological pathway in dynamic model development that combine qualitative (CLD) and quantitative (stocks and flows and agent based models) approaches" - point out Misuraca and Kucsera (2016:9). "Dynamic Simulation Modelling …the combination of System Dynamics (SD) with Agent Based Modelling (ABM) would be the most appropriate approach…for modelling and simulation …the 'ecosystem' in which social protection 'organisms' operate…in the EU …characterized by …human behaviours and the unpredictable impacts they …have …on system"(Misuraca and Kucsera, 2016:46).

This approach of System Dynamics indicates the actor-driven nature of social systems 'producing' the problems to be modelled. It helps to improve conceptualization, which "…guides the model building process and leads to a shift from "variables" to "decisions rules" and "actions"…"(Olaya and Gomez-Quintero, 2016:1). Combined methodologies by analyzing agency should reflect the process character of (continuously emerging) social systems.

CONCLUSIONS

The civil society is 'uncharted territory' for modelling and simulation in general and for System Dynamics in particular. The SD may contribute to quantitative analyses deepening findings from qualitative analysis of the civil society organizations' transformational dynamism. Modelling and simulation of civil society entities create numerous challenges. The very concepts of civil society, volunteering and social capital are rather elusive not to mention challenges related to their measurement and quantification. The limited size and non-homogenous character of available data, lack of reliable sources may turn building models of civil society entities into 'non-standard exercise'. One can foresee the necessity of capitalizing also on "less traditional" solutions and sources, including open data, log analysis, accession of survey and poll results. More thorough study of the civil society organizations' transformational dynamism through concrete cases probably should consider also effects of time distribution and check relevance of metamodels.

Legislative efforts limiting and preventing disruptive effects of platform firms such as participative democracy related efforts in Taiwan and sharing economy related initiatives similar to platform cooperatives or Commons Based Peer Production providing alternative patterns of value creation may capitalize on the civil society entities' transformational dynamism. Each of them and their interplay could be more succesfully analysed by deploying innovative dynamic hybrids combining System Dynamics with Agent Based Modelling the paper assumes.

REFERENCES

Avent, R., 2016. The productivity paradox. https://medium.com/basic-income/what-if-you-got-1-000-a-month-just-for-being-alive-i-decided-to-find-out-9e8591976c37#.1zsfzafjr

Benkler, Y. 2011. The Penguin and the Leviathan. New York: Crown business

Benkler, Y. 2016. Network pragmatism - an alternative future. Summary of lecture on "Building the Cooperative Internet" Conference, November 2016, New York New School and Civic Hall http://platform.coop/stories/happy-new-year

Bollier, D. 2007. 'The growth of the Commons paradigm' In: Hess C. and Ostrom E. (reds.) (2007) Understanding knowledge as Commons From theory to practice. MIT Press. pp.27-41.

Castells, M. 2009. Communication power. Oxford University Press.

Csányi, V. 1989. Evolutionary systems and Society - A General Theory. Duke University Press.

Fukuyama, F. 1999. 'Social Capital and Civil Society'. In: IMF Conference on Second Generation Reforms http://www.imf.org/external/pubs/ft/seminar/1999/reforms/fukuyama.htm#I

Galbraith, J.K. 1987. A history of economics The past as the present. Hamish Hamilton London

Hill, S. 2016. The California Challenge How (not) to regulate disruptive business models http://library.fes.de/pdf-files/id-moe/12797-20160930.pdf

Katz, L.; Krueger A.B. 2016. The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015. https://krueger.princeton.edu/sites/default/files/akrueger/files/katz_krueger_cws_-_march_29_20165.pdf

Luhmann, N. 1995. *Social systems.* Stanford University Press.

Miller, S.R. 2016. First Principles for Regulating the Sharing Economy. *Harvard Journal on Legislation*, Volume 15, 2016, pp: 149-200.

Misuraca, G. (ed.); Kucsera Cs. (ed.) (2016) ICT-Enabled Social Innovation in support to the Implementation of the Social Investment Package. IESI. EU JRC Science Hub http://is.jrc.ec.europa.eu/pages/EAP/documents/20160226_IESI_D2.2_i-FRAME-V1.5_DRAFT_V1.0_000.pdf

Nowak, M. A. 2006 'Five rules for the evolution of cooperation'. *Science.* 8 December 2006 Vol 314: 1560-1563.

Olaya C.; Gomez-Quintero J. 2016. Conceptualization of Social Systems: Actors First. Proceedigns of the 34th International Conference of the System Dynamics Society Delft, Netherlands July 17-21, 2016. http://www.systemdynamics.org/conferences/2016/proceed/papers/P1360.pdf

Rucker, P., Costa R. 2017. Bannon vows a daily fight for 'deconstruction of the administrative state'. The Washington Post. https://www.washingtonpost.com/politics/top-wh-strategist-vows-a-daily-fight-for-deconstruction-of-the-administrative-state/2017/02/23/03f6b8da-f9ea-11e6-bf01-d47f8cf9b643_story.html?utm_term=.9fdce11b8ca9

Scholz, T. 2014. Platform Cooperativism vs. the Sharing Economy. https://medium.com/@trebors/platform-cooperativism-vs-the-sharing-economy-2ea737f1b5ad#.71imuwf10

Scholz, T. 2016 Platform Cooperativism: Challenging the Corporate Sharing Economy. http://www.rosalux-nyc.org/wp-content/files_mf/scholz_platformcooperativism_2016.pdf

Stacey, R.D. 2000. *Strategic management and organizational dynamics The challenge of complexity.* Financial Times, Prentice Hall.

Tang, A., 2016. Uber responds to vTaiwan's coherent blended volition.https://blog.pol.is/uber-responds-to-vtaiwans-coherent-blended-volition-3e9b75102b9b#.cn5ygcttm

Tang, A.; Kao, C-L. (2016) Challenges for Taiwan's Civic Hackers in 2016. https://medium.com/@audrey.tang/challenges-for-taiwan-s-civic-hackers-in-2016-385af61d6e79#.6jn89lpoh

Veress, J. 2016. Transformational Outcomes of Civil Society Organizations. Aalto University

Vriens, D.; Achterbergh, J. 2006. The Social Dimension of System Dynamics-Based Modelling. Systems Research and Behavioral Science, Volume 23, Issue 4 July/August 2006. http://onlinelibrary.wiley.com/doi/10.1002/sres.782/abstract

Weil, D. (2014) The Fissured Workplace Why Work Became So Bad for So Many and What Can Be Done to Improve It. Harvard University Press.

DETERMINANTS OF FX-RISK MANAGEMENT EVIDENCE OF HUNGARY

Barbara Dömötör
Erzsébet Kovács
Department of Finance
Department of Operations Research and Actuarial Sciences
Corvinus University of Budapest
1093, Budapest, Hungary
E-mail: barbara.domotor@uni-corvinus.hu

KEYWORDS

Corporate hedging, Forward hedge, Linear regression model

ABSTRACT

This paper investigates the motives of FX-risk management based on the changes of forward open positions of Hungarian corporations. We have found that Hungarian companies are significantly more exposed in short EUR forward position, than in EUR long one. Our linear regression model also showed that changing market conditions have an essentially higher impact on the EUR short positions. Our results confirmed that expectations are determining in the risk hedging decisions proving that financial risk management also has a speculative motive.

INTRODUCTION

The management of risks is one of the main tasks of corporate management not only in the theory but also in practice. It is a more and more common opinion (Merton, 2008) that despite the traditional theory that defines the task of risk management in reducing risks to a minimum level, the concept of enterprise risk management (ERM) (Casualty Actuarial Society, 2003) refers to the potential added value of taking certain types of risks.

Based on the classic model of mean-variance optimization, where utility is a positive function of the expected value and a negative one of the variance of the random income at maturity, if the hedging instrument is costless, the optimal hedging ratio (*w*) depends both on the expected value of the hedging instrument (*E*(*y*)) (speculative hedge) and the covariance between the underlying asset (x) and the hedging instrument (pure hedge) (Rolfo, 1980).

$$w = \frac{E(y)}{2a\,\mathrm{var}(y)} - \frac{\mathrm{cov}(x,y)}{\mathrm{var}(y)} \tag{1}$$

where the corporate specific risk aversion is denoted by *a,* while *var* stands for the variance.

For uninformed hedgers, the hedging ratio shall be equal to the second part of the equation, so that the variance of the whole portfolio is minimized (Duffie, 1989).

In practice, non-financial corporations are using derivatives not only for variance reducing purposes. Stulz (1996) explains this fact by showing that some companies may have a comparative advantage in forecasting price movements and bearing financial risk. Consequently, for them, risk management shall not necessarily aim to minimize variance, but it is worth to take risks in those areas where the company has comparative advantages, by ensuring the downside outcomes of significant costs at the required level.

According to Lessard (2008), the "hierarchy theory" of corporate risk management is as follows: the first and most important is to define those activities which the company has comparative advantages in, and risks shall be taken here. The first level of risk management is related to the operation of the company; business strategy and operative management are defined based on them. The management of financial risks means the management of the risks arising / remaining as a result of the strategic decisions.

Contrary to this, Hommel (2003) considers operating flexibility as an alternative to financial risk management, indicating the conditions under which it is worth to apply operative hedging instead of financial risk management at the company.

The available information about the practice of corporate risk management is limited, considering that companies have no recording or reporting obligation in this regard. The explanation of financial statements in the appendices of annual financial reports may contain guidance about the company's financial transactions relating to managing financial risk.

The international literature contains several empirical analyses in connection with corporate hedging, including Tufano (1996), Haushalter (2000), Mian (1996), Josepf and Hewins (1997), based on public databases and data from annual financial reports, or relying on the results of surveys. Dominguez and Tesar (2006) examined aggregate company exposure, Bodnar et al. (1998, 1999) have conducted a survey on the risk management practices of U.S. and German companies.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

In this paper, we aim to capture corporate hedging behavior based on the change of the foreign exchange forward positions.

The structure is the following: first, we introduce the data and the framework of the analysis, and then the results of the linear regression models are presented. Finally, we summarize our findings.

DATA

The most widespread assets of financial hedging are derivative products: options and forwards or futures. As FX market is a typical OTC (over-the-counter, non-exchange) market and the usage of options is marginal in risk management (Bodnar et al, 1998, 1999), we focus on the practice of forward FX-hedge.

The National Bank of Hungary (NBH) provided a database containing the daily aggregate *stocks* of forward transactions of domestic financial institutions with other resident partners (financial sector excluded) in the period between January 2003 and March 2016. The database is based on the reports of domestic banks and it also includes the positions' direction. Positions denoted by "EUR long" are those positions if the resident client of the bank buys the foreign currency for Hungarian Forint (HUF) on forward, and "EUR short" refers to the positions when the client sells the foreign currency. Foreign currencies are not restricted to the EUR, but as the major trading currency is the EUR, we will use it for the sake of simplicity.

"EUR short" forward position can derive either from the hedging activity of exporters or from speculation on the strengthening the forint. On the other hand "EUR long" forward position can serve the purpose of hedging import or speculation.

We could not differentiate the positions according to client type, so the data are based not only on corporate transactions, but retail forward deals as well.

Figure 1 shows the evolvement of the forward positions, it can be seen, that foreign currency short positions are higher than long positions; the currency sold on forward is about the double of forward currency purchases.

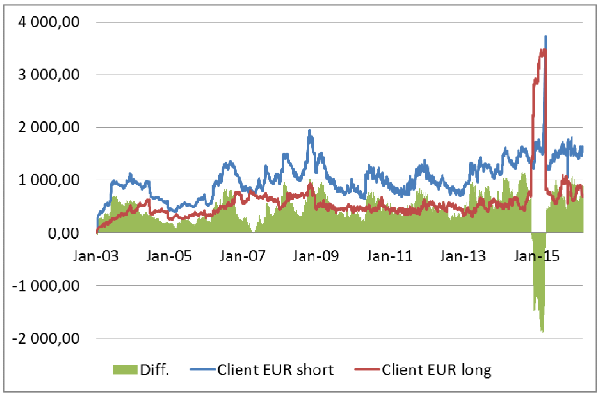


Figure 1: Aggregate open forward positions 2003-2016 (Bn HUF)

Source: the authors based on NBH data

The only exception is the period between November 2014 and March 2015, and it can be caused by the transactions relating to the compulsory conversion of the retail foreign currency loan portfolio.

The higher volume of the forward currency short positions can be explained either by the difference of the trading balance or by speculation.

The profit of a forward speculation is the difference between the forward rate and the spot price at maturity. According to the modern portfolio theory, the expected value of a forward position depends on the systemic risk of the currency. If the correlation between the spot FX-rate and the market prices of other assets is positive, the expected spot rate exceeds the forward price, so long FX-forward positions have positive expected value. If this correlation is negative, the systemic risk is also negative, resulting in a premium of the forward prices, and a positive expected value of forward currency sale (Hull, 2007).

In the studied period, the country risk premium of Hungary was positive, consequently, the expected value of the future spot price was lower than the forward price. The measures of the forward stocks prove the above fact, and that the expected value of the forward position has an important role in the risk management decisions.

In the followings, we investigate the forward positions and through that the corporate risk management behavior in details.

ANALYSIS

For analyzing the factors that influenced the changes of long and short forward positions[1], we built a multivariate linear regression model[2], in which the dependent variable is the monthly percentage change of the positions. Although the database allows for the analysis of daily data, we decided to examine monthly changes instead, due to the unsystematic effects and noises of the daily data. In doing so, we took the first available data of each month and assigned the percentage change to the expiration date. The study period consists of 145 months between January 2004 and March 2016. We decided to omit the data from the first year, 2003, because the data collection started then, and the scope of data collection was unstable in that period.

The open currency position to be hedged mainly derives from the foreign trade turnover, so the first set of explanatory variables is the monthly trading balance (export – import)[3].

As we aimed to capture the speculative motives of the hedge and the risk hedging behavior we included market data among the explanatory variables that can affect the expected value of the hedge and also of the speculative positions. According to market experience and our expectations, these factors could be the FX-rate, its

[1] SPSS 22 software package was used to the analysis.

[2] a description of the method is included in Kovács, 2009

[3] Data available from Bloomberg

volatility, the difference between the forward rate and the spot rate (swap-difference) and the foreign and domestic interest rates.

We used the data series of the spot exchange rate of two foreign exchange rates, EURHUF and EURUSD. Considering that a change in the spot exchange rate affects also the forward rate, changes in the spot rate are expected to determine the hedging decision, as in the case of an advantageous price movement, better hedging rates are available.

The next group of variables is the volatility of FX-rates. We downloaded the implied volatility of at-the-money option with 30, 90 days or 1-year maturity for both currency pairs[4]. Higher volatility of risk factors (foreign exchange rate) is expected to increase the utility available by hedging, and also the potential hedging needs.

The expected profit of a forward transaction is also affected by the difference of the forward exchange rate and the expected spot exchange rate, this difference, (swap-difference) is also quoted in the interbank market; its time-series is available on Bloomberg. The swap-difference is determined by the difference of the spot exchange rate and the interest rates of the two currencies. Therefore, in addition to the one-year development of swap-difference, 1-year BUBOR representing HUF interest level, and 1-year EURIBOR and 1-year USD-LIBOR data for the foreign interest were included among the explanatory variables.

In order to ensure the independence of the individual observations required by the methodology, both the dependent variable and the explanatory variables consist of the percent change in each factor.

By examining the time series of the derivative stocks (Figure 2), we found a periodicity in both the long and the short positions: they, in general, are strongly reduced between December and January.

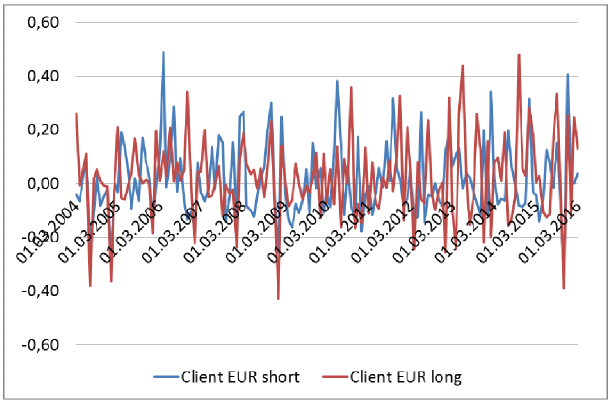


Figure 2: Monthly change of open forward positions
Source: the authors based on NBH data

This fall in the stocks may be explained by the positions expiring or closed out at the end of the year, since most corporations hedge to the reporting date, or in many cases, they do not want to report a substantial derivative portfolio in their annual statements.

Therefore, in addition to market factors, the "December effect" was included as a binary explanatory variable, with 1 standing for the period from the beginning of December to the beginning of January, in any other cases, it is 0.

The explanatory variables within a variable group are strongly correlated, but we included more variables to quantify the same factor, in order to find the variables with the highest explanatory power. However, variables are not uncorrelated between groups of variables either; therefore, we addressed the collinearity between the explanatory factors by choosing the stepwise method in the regression, thus the redundant variables are not included in the model.

Table 1 contains a summary of the explanatory variables of the regression model.

Variable clusters	Variable	Short name
Foreign trade	Monthly trade balance: export – import	Trade
Market prices	EURHUF spot rate	EURHUF
	EURUSD spot rate	EURUSD
	12-month BUBOR	BUBOR
	12-month EURIBOR	EURIBOR
	12-month USD-LIBOR	USDLIBOR
	12-month EURHUF swap difference	EURHUF_swap
	12-month EURUSD swap difference	EURUSD_swap
Market volatility	EURHUF 30-day implied volatility	EURHUF1MV
	EURHUF 90-day implied volatility	EURHUF3MV
	EURHUF 1-year implied volatility	EURHUF12MV
	EURUSD 30-day implied volatility	EURUSD1MV
	EURUSD 90-day implied volatility	EURUSD3MV
	EURUSD 1-year implied volatility	EURUSD12MV
December effect	The period is December	December

Table 1: Variables of the regression model
Source: the authors

We expect that EUR short positions (export hedge) are positively affected by the exchange rate, while EUR long positions (import hedge) correlate negatively with EURHUF.

[4] Data source: Bloomberg

RESULTS

In the first run the analysis of the residuals showed that one observation, 3rd November 2008, influenced critically the regression model, therefore, we excluded this month from the further analysis. Figure 3 shows Centered Leverage and Cook's Distance values. The horizontal and vertical lines represent acceptance limits based on the rule of thumb.

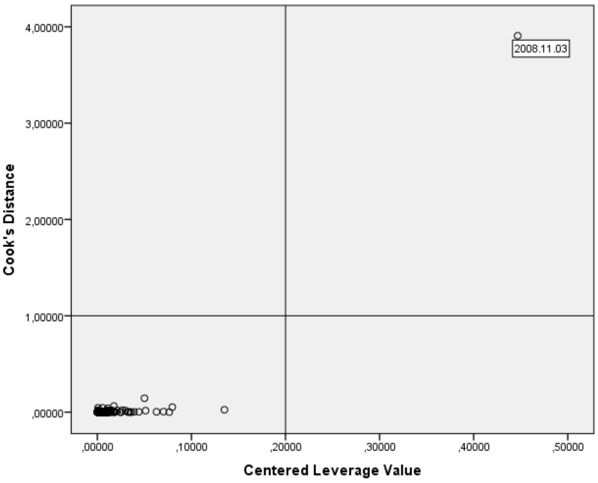


Figure 3: Exploring extreme data
Source: the authors based on NBH data

The extremity of this special date can be caused by the liquidity crisis of the Hungarian market in October 2008.

Modeling EUR short positions

Based on the *F-statistics* of the regression model, the model is significant at any conventional level (*p-value* 0.000), the adjusted R^2 coefficient is 0.58, i.e. nearly 60% of the total variance is explained. The essence of the stepwise method is that the independent variables are included gradually, on the basis of their explanatory power, as long as the new explanatory variable significantly improves the model. Table 2 shows that short forward positions were affected by three factors substantially: the change of the 30-day implied volatility of EURHUF; the EURHUF spot rate and the position closings in December.

It can be seen that the changes in foreign trade turnover had no effect on the derivative stock, which can be explained by the fact that products crossing the economic border of the country count to foreign trade, while hedging is carried out prior to the sale, in advance. In addition, we had data only about the net value of the foreign trade that proved to be almost uncorrelated to the forward positions.

Model Summary[d]

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	,678[a]	0,46	0,46	0,10	0,46	120,66	1,00	142,00	0,00	
2	,735[b]	0,54	0,53	0,09	0,08	25,05	1,00	141,00	0,00	
3	,768[c]	0,59	0,58	0,09	0,05	16,67	1,00	140,00	0,00	2,18

a. Predictors: (Constant), EURHUF3MV

b. Predictors: (Constant), EURHUF3MV, EURHUF

c. Predictors: (Constant), EURHUF3MV, EURHUF, December

d. Dependent Variable: EUR_short

Table 2: Summary of the linear regression model of short FX-forward positions
Source: the authors based on NBH data

Table 3 contains the details of the regression model. On the basis of the high t-statistics and low p-values of the table, each variable is significant at all conventional levels. From the above volatility data, the 90-day (3 months) volatility of the EURHUF exchange rate proved to be determining; because of the correlation of the volatility data, additional volatility time series do not provide further explanatory power. 1 percent increase in the EURHUF market volatility, the portfolio increases by 0.475 percent.

1 percent rise in EURHUF exchange rate, meaning weakening of the forint against the euro, as expected, increases the sold foreign currency portfolio by 1.896 percent in the period.

Short currency positions are reduced by 11% on average in December.

Although the expected value of the forward sale can be increased by the EURHUF swap difference, it is surprising that this factor was not included in the explanatory variables of the model and so the higher forward price arising from the difference in interests alone does not cause changes in the forward portfolio. The reason for this is possibly the strong relationship between the swap difference and the EURHUF spot rate.

The correlation between the explanatory variables is acceptable, the tolerance value is close to 1, as well as its inverse, the variance inflation factor (VIF), the value of which is also close to one.

	Unstandardized Coefficients		Standardized Coefficients			95,0% Confidence Interval for B		Correlations			Collinearity Statistics	
Model	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
3 (Constant)	,018	,008		2,340	,021	,003	,033					
EURHUF3MV	,475	,068	,471	7,021	,000	,341	,608	,678	,510	,380	,651	1,535
EURHUF	1,896	,350	,363	5,412	,000	1,203	2,588	,631	,416	,293	,651	1,537
December	-,110	,027	-,221	-4,083	,000	-,163	-,057	-,188	-,326	-,221	,998	1,002

Table 3: Coefficients of the linear regression model of short FX-forward positions
Source: the authors based on NBH data

The partial correlations cleaned from effects of other explanatory variables are moderate.

The distribution of the residuals slightly deviates from normal, which is confirmed by the result of the Kolmogorov-Smirnov test, p-value I s0.00.

All in all, the model is suitable for examining the relationship between the short forward position and the major market factors but about 40% of the variance is influenced by other factors like foreign trade and unsystematic effects.

Modeling EUR long positions

Similarly to the explanation of changes in short forward position, the changes of long forward FX-position is also analyzed in a linear regression model. The explanatory variables are the market-, foreign trade- and calendar factors as described above, while the dependent variable is the change in the stock of long forward stock. As mentioned above there was an extreme jump in the EUR long positions in October 2014 and a sudden fall in March 2015 due to a political regulation. That is why we substituted the values of the referring month by the average monthly change in the forward positions.

Model Summary[d]

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	,324[a]	0,10	0,10	0,15	0,10	16,63	1,00	142,00	0,00	
2	,405[b]	0,16	0,15	0,14	0,06	10,03	1,00	141,00	0,00	
3	,452[c]	0,20	0,19	0,14	0,04	7,03	1,00	140,00	0,01	2,17

a. Predictors: (Constant), December

b. Predictors: (Constant), December, EURHUF

c. Predictors: (Constant), December, EURHUF, EURHUF12MV

d. Dependent Variable: EUR_long

Table 4: Summary of the linear regression model of long FX-forward positions
Source: the authors based on NBH data

Although the model is significant at all conventional significance levels; the *p-value* of the *F statistics* is 0.001, the model explains only 20% of the total variance, according to the *adjusted* R^2 indicator, as shown in Table 4.

Table 5 summarizes the explanatory variables, the regression coefficients and associated other statistics of the model.

	Unstandardized Coefficients		Standardized Coefficients			95,0% Confidence Interval for B		Correlations			Collinearity Statistics	
Model	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
3 (Constant)	,036	,012		2,941	,004	,012	,061					
December	-,180	,043	-,319	-4,228	,000	-,264	-,096	-,324	-,336	-,319	,996	1,004
EURHUF	-2,334	,559	-,393	-4,176	,000	-3,439	-1,229	-,259	-,333	-,315	,642	1,558
EURHUF12MV	,475	,179	,250	2,651	,009	,121	,829	-,004	,219	,200	,641	1,560

Table 5: Coefficients of the linear regression model of long FX-forward positions
Source: the authors based on NBH data

In explaining the long forward positions, the constant proved to be significant; so if all other factors remain unchanged, the monthly portfolio increases by 3.6%.

For long forward position, the December effect is the most significant, this variable was involved first. The year-end position maturities and closings reduce the overall portfolio by almost 18% on average. The other two explanatory variables, similarly to the model in the previous subsection, are the change in the EURHUF exchange rate and the 12-month volatility of EURHUF. The weakening of the forint against the euro by 1 percent reduces the portfolio by 2.334%, in line with the expectations.

The growth of volatility is in positive correlation with the portfolio increase; however, in the case of long positions, the 1 year implied volatility had the highest explanatory power, so it was included in the model as an explanatory variable. All the three explanatory variables are significant at a level higher than 99%, tolerance is close to 1, and the VIF value supports that the explanatory variables are uncorrelated.

The distribution of residuals cannot be considered normal; the *p-value* of the Kolmogorov-Smirnov test is 0,004.

The foreign trade turnover and the EURHUF swap difference were not included as explanatory variables, the reasons can be similar to those referred to in the case of the short positions: the foreign trade turnover and the timing of the hedging decision are different, or may vary, while the swap difference does not have an additional significant effect on the development of the portfolio due to its co-movement with the EURHUF exchange rate.

The effect of the explanatory variables developed in line with the expectations, but the model does explain only one-fifth of the total variance, which indicates non-linear effects and further explanatory factors as well as the importance of individual factors also in the case of long positions.

CONCLUSION

We analyzed the aggregate long and short forward positions of the non-financial resident clients of Hungarian banks, in order to detect the motives of corporate FX-hedging and speculation. We found that market movements have an important impact on the change of the forward positions. As we expected, trading activity increased following favorable spot rate changes, long forward position correlated negatively, short forward positions correlated positively with the spot price. Higher volatility of the exchange rate caused an increase of the forward positions in both directions. We found that both positions fall significantly at the end of the year. We could explain the changes of the short foreign currency positions in 60 percent, but the explanatory power of the model was only 20% for long FX-positions.

REFERENCES

Bodnar, G. M., Hayt, G. S., & Marston, R. C. (1998). 1998 Wharton survey of financial risk management by US non-financial firms. *Financial management*, 70-91.

Bodnar, G. M., & Gebhardt, G. (1999). Derivatives Usage in Risk Management by US and German Non-Financial Firms: A Comparative Survey. *Journal of International Financial Management & Accounting*, *10*(3), 153-187.

Casualty Actuarial Society (2003). *Overview of Enterprise Risk Management*. Study of Enterprise Risk Management Committee

Dominguez, K. M., & Tesar, L. L. (2006). Exchange rate exposure. *Journal of international Economics*, *68*(1), 188-218.

Duffie, D. (1989). Futures Markets. *Englewood Cliffs, New Jersey*.

Haushalter, G. D. (2000). Financing policy, basis risk, and corporate hedging: Evidence from oil and gas producers. *The Journal of Finance*, *55*(1), 107-152.

Hommel, U. (2005). Value-based motives for corporate risk management. In *Risk management* (pp. 455-478). Springer Berlin Heidelberg.

Hull, J. C. (2007). *Options, futures, and other derivatives*. Pearson

Joseph, N. L., & Hewins, R. D. (1997). The motives for corporate hedging among UK multinationals. *International Journal of Finance & Economics*, *2*(2), 151-171.

Kovács, E. (2009) Statistical Analysis of Financial Data. Tanszek Kft, Budapest, 160.

Lessard, D. (2008). The Link Between Risk Management and Corporate Strategy. In MIT Roundtable on Corporate Risk Management. *Journal of Applied Corporate Finance* 20(4), 20-38.

Merton, R. C. (2008). MIT Roundtable on Corporate Risk Management. *Journal of Applied Corporate Finance,* 20(4), 20-38.

Mian, S. L. (1996). Evidence on corporate hedging policy. *Journal of Financial and quantitative Analysis*, *31*(03), 419-439.

Rolfo, J. (1980). Optimal Hedging under Price and Quantity Uncertainty: The Case of a Cocoa Producer. *Journal of Political Economy,* 88(1), 100–116.

Stulz, R. M. (1996). Rethinking risk management. *Journal of applied corporate finance*, *9*(3), 8-25.

Tufano, P. (1996). Who manages risk? An empirical examination of risk management practices in the gold mining industry. *the Journal of Finance*, *51*(4), 1097-1137.

AUTHOR BIOGRAPHIES

BARBARA DÖMÖTÖR is an Assistant Professor of the Department of Finance at Corvinus University of Budapest (CUB). She received her Ph.D. in 2014 for her thesis modeling corporate hedging behavior. Prior to her recent position, she worked for several multinational banks treasury. Her research interest focuses on financial markets, financial risk management, and financial regulation. Her e-mail address is: barbara.domotor@uni-corvinus.hu

ERZSÉBET KOVÁCS is Professor and Head of the Department of Operations Research and Actuarial Sciences at Corvinus University of Budapest (CUB). She received Candidate of Science degree in 1991 and became Dr Habil in 2003. Her fields of research are the followings: Pension Modelling, Mortality projection, Risk Analysis of Student Loan Scheme**,** Application of multivariate statistical methods in an international comparison of insurance markets, Statistical analysis of the economic transition and inventory characteristics in Central-Eastern Europe.
Contact: erzsebet.kovacs@uni-corvinus.hu

MODEL OF THE STATE AND EU INVOLVEMENT IN THE VENTURE CAPITAL MARKET

Erika Jáki, PhD. and Endre Mihály Molnár
Department of Enterprise Finances
Corvinus University of Budapest
Fővám tér 8.,Budapest,1093, Hungary
E-mail: jaki.erika@t-online.hu

KEYWORDS
Venture capital, state involvement, seed, EU funds

ABSTRACT

It is especially difficult for seed stage companies to find adequate financing. In the last decade venture capital (VC) has played significant role in funding seed and start-up stage companies. Our study focuses on the financing of seed stage companies via venture capital funds subsidized by the state and European Union. Seed stage companies are supported by incubator houses with infrastructure and expertise. Accelerators help them with their partner network, with intensive training and occasionally with capital. There is no sharp borderline between incubator houses and accelerators regarding the provided services. We give an overview about the history of the Hungarian VC market with its most important milestones. In our study, we pay extra attention to the appearance of the governmental and the EU funds, and focus on the model of the local VC market, presenting how funds operate and distribute state subsidies.

INTRODUCTION

Many authors have researched various aspects of state involvement on venture capital market. One of the most comprehensive books is written by Gompers and Lerner (2004), which presents systematically how venture capital industry works in the United States. They examined conditions and circumstances under governments can efficiently act as venture capitalists. They concluded that governments should help in the financing of small companies as it generates a positive social effect. This conclusion is supported also by other researchers (Harding, 2000; Sohl, 1999). Lerner (1999) examined the long-term effects of US venture capital programs called “Small Business Innovation Research” (SBIR). This program has run from 1983 to 2003 and had distributed 13 billion dollars to small high-technology firms.

The OECD survey (1997) categorized government programs as follows: 1) providing sources to invest in small companies, 2) providing financial incentives for investing in small companies and 3) regulations for venture capitalists. Government venture capital schemes intend to capture public benefits in terms of increased innovation, economic growth and job creation. According to EU financial market policies the role of venture capital finance is to facilitate employment and improvement in productivity (Schelter, 2006). Garbade (2011) did a comparative analysis of venture capital financing by U.S., British, German and French Information Technology Start-ups The EU also started it’s own venture capital program called Jeremie, which we will examine in the upcoming chapters regarding the Hungarian market.

Many researches were published concerning Hungarian venture capital market and financing. Most relevant publications are presented briefly later in that article. In the upcoming chapters, we present the model of the European Union and the state involvement in the Hungarian venture capital financing.

MODEL OF THE VENTURE CAPITAL MARKET IN HUNGARY

The most important actors of the Hungarian venture capital market are displayed on Figure 1. The *target companies* can be distinguished based on three stages of their business development: *seed*, start-up and expanding enterprises. We are going to focus on seed stage enterprises that can be supported by *incubator houses* or *accelerators.* There is no sharp borderline between the two types supporting entities. The *Vállalkozói Inkubátorok Szövetsége* (VISZ - Association of Entrepreneurial Incubators) gives all its support market participants to found more entrepreneur incubators.

Other groups of important actors are *the venture capital fund management* and *the capital funds.* The *Hungarian Private Equity and Venture Capital Association* (HVCA) represents the interests of the whole private equity and venture capital sector in Hungary. Its mission is to support its members and promote adherence to the highest possible professional and ethical standards. Another important actor is the *Central Bank of Hungary* (CBH – in Hung. *Magyar Nemzeti Bank*; MNB) who plays a supervisory role. Until the year 2013 the supervision was performed by the “*Pénzügyi Szervezetek Állami Felügyelete*” (Financial Supervisory Authority), which merged with the CBH.

On the one hand, governmental participation appears in the foundation of venture capital fund managements on the other hand in subsidizing funds. The EU involvement manifests by providing financial sources, distributed via tenders. The tenders of the EU funds are coordinated by the *Hungarian Development Bank Plc.* (HDB). Earlier this task was performed by the Magyar Vállalkozásfinanszírozási Zrt. (Hungarian Business

Financing Plc) which was merged into HDB in 2015. All the plans containing state subsidy in terms of EU regulation must be announced to the authority of *the State Aid Monitoring Office* (SAMO). It is responsible for examining competition regulatory aspects of state subsidies. As a rule, state subsidy is banned in the EU as it distorts market competition. The state can intervene only if there are market failures in a segment.

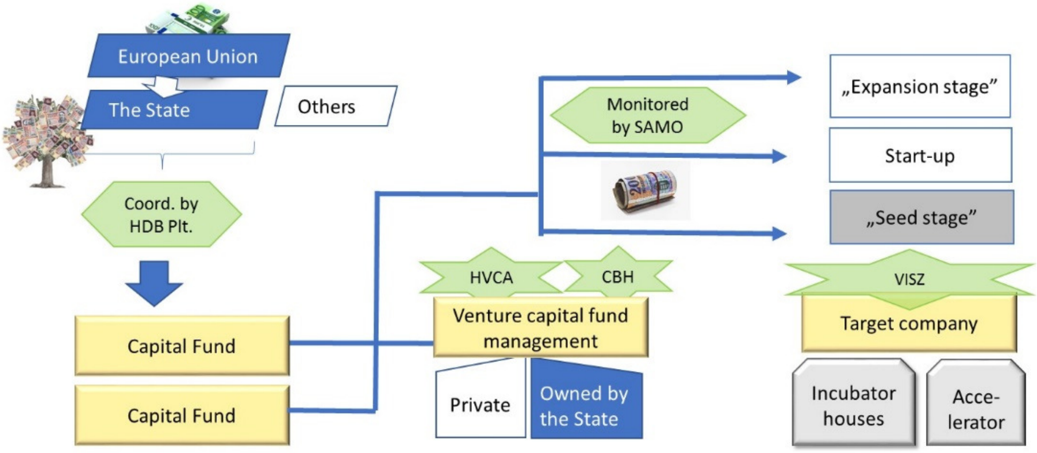


Figure 1: The most important actors of the Hungarian venture capital market in 2016

STAGES OF DEVELOPMENT OF STARTER COMPANIES AND THEIR TYPICAL FINANCING

Financial sources for starter enterprises according to different stages of their life-cycle are as follows:

- The "seed stage enterprises" often possess merely the product/service idea ("idea company"). Investors of these companies are usually business angels, seed funds, accelerators, or the 3F (Family, Friends, Fools).
- "Start-up enterprises" already developed an operating prototype and have some market response on the product or service. These enterprises are beloved targets of venture capital funds.
- The "expansion stage enterprises" have an established business but need additional financing to expand further on the market (marketing expenses, and to cover the initial losses). Venture capital funds and private equity investors are the typical investors of these companies.

In these early stages the enterprises cannot count on bank loans. But as we see venture capital funds are present in each of the three stages, emerging in several forms (Walter, 2014a).

One typical problem of starter enterprises is the lack of an economic/financial expertise to set up a business plan to present to potential investors. On the other hand, it is hard for starter companies to set up a feasible business plan, because the operation, business model has not evolved completely yet. Furthermore, the product/service creates in many cases a "blue ocean" in the sector. As there are no competition and benchmark in this case, so business plan should focus on how customers could be convinced. Good examples from such innovations from the near past are: "pick pack point" that is widespread package sending method by now; the smartphone; the smartwatch; virtual reality headsets. Not only is the forecasting of revenues difficult in such cases but the estimation of costs also.

Incubator houses and accelerators

On the Hungarian market there is always confusion in distinguishing incubator houses and accelerators in their name and also in their provided services.

Incubator houses do not invest, they just provide professional business support and infrastructure – e.g. office space, office and business related services – to start-ups on a favorable price in the growth stage of development. In Hungary, the Vállalkozói Inkubátorok Szövetsége coordinates the main traditional incubator houses:

- Közép-dunántúli Regionális Innovációs Ügynökség Nonprofit Kft.
- Primom Vállalkozói Inkubátorház és Innovációs Központ
- Dél-Dunántúli Regionális Innovációs Ügynökség Nonprofit Kft.
- Főnix Inkubátorház és Üzleti Központ
- Budapesti Politechnikum Alapítvány
- Bács-Kiskun Megyei Angol-Magyar Kisvállalkozási Alapítvány
- Vállalkozói Központ Közalapítvány, Székesfehérvár

Accelerators support starter enterprises in implementing their business idea. The starter entrepreneurs take part in a training to develop the given business activity. The program usually lasts for a couple of months. Entrepreneurs can meet business mentors via the business network of the accelerator, and establish their presence in the given industry sector. Certain accelerators – in return for a small equity share – occasionally even provide capital for the start-up company. It becomes common, that venture capital investors establish accelerators to finance the most promising enterprises from their own seed funds. An example for this is the SeedStar accelerator of DBH Investment Plc., which defines itself as an incubator house and as an accelerator. It organized the SeedStar Battle start-up competition in April, 2015, where the presenting starter enterprises were evaluated by a professional jury, moreover competitors had the opportunity to meet venture capital investors as well.

VENTURE CAPITAL FUND MANAGEMENT AND FUNDS

The VC fund management is legally separated from the fund it handles. The fund management collects liquid funds from different investors into the VC capital fund. The fund itself is without legal personality, and its role is to finance investments made by the capital fund management. One of the legal requirements for the establishment of a fund is the completion and authorization of the Management Guidelines. Among many conditions, it fixes the investment strategy, the target industry, the target development stage of the enterprises and also determines the expected return and investment tenor. Similarly, to the stages of development by start-up enterprises, we can differentiate "seed", "start-up" and the "expansion funds" based on their investment strategy.

The fund management charges a fund with management fee for their services. It mainly covers the cost of the operation of the fund management (infrastructure and employees). With the expiration of the fund's lifetime the accumulated amount in the fund must be paid back to the investors.

It is important to differentiate the *expected return from a single investment* and the *expected return of a total fund.* Table 1 shows the returns expected of a single investment by the US venture capital investors in different stages of development by the enterprises. Obviously, the expected return of a single investment is the highest at companies in the seed stage, where the chance of survival is the lowest. As the enterprise develops, the probability of survival becomes higher and the expected return of the investment decreases.

Table 1: Expected returns in the United States

Life cycle	**Expected returns of Venture Capital in the US**
Seed	>80%
Start-up	50-70%
Expansion 1st round	40-60%
Expansion 2nd round	30-50%

Source: Sahlman-Scherlis (2003)

While some of the investments produce great returns, several of them end up with a failure. That is why the expected return of one single investment is high, but realized return of the fund is much lower. (See Table 2 in case of US venture capital funds). It can be seen that in the seed stage expected return if 80% for a single investment, while the realized three-year-long annual return was 4,9% and the biggest return was 32,9% during a ten-year-long interval.

Table 2. Returns Earned by Venture Capitalists Looking Back from 2007

Investor / index type	**3 years**	**5 years**	**10 years**	**20 years**
Early/seed VC	4.90%	5.00%	32.90%	21.40%
Balanced VC	10.80%	11.90%	14.40%	14.70%
Later stage VC	12.40%	11.10%	8.50%	14.50%
All VC	8.50%	8.80%	16.60%	16.90%
Benchmarks:				
NASDAQ index	3.60%	7.00%	1.90%	9.20%
S&P index	2.40%	5.50%	1.20%	8.00%

Source: Damodaran (2009)

Based on the numbers of the two tables we can state that the realized returns achieved by the venture capital fund are deeply under the expected returns of a single investment, but over the returns of some stock market indexes like NASDAQ index and the S&P index.

On the Hungarian market Karsai (1997) estimates the expected return to 35-50% in 1997 and to 30-40% in 2002. Estimations were made by interviews with venture capital investors active on the local market. We must mention that the state owned "Széchenyi Tőkealap-kezelő Zrt." (Széchenyi Capital Fund Management Plc.) expects a return of 12% to 15% from its investments, which is much lower than the expected return of one single investment by other Hungarian or foreign fund managements. It must also be noted that this expected return range only applies to companies with at least 2 years of operating history at the Széchenyi Tőkealap-kezelő Zrt.

Table 3: Composition of the capital for Venture Capital and Private Equity Investments

Type of sponsors	**Proportion of total funds provided**
Government organizations	58,52%
Banks	18,00%
Private investors	10,48%
Corporate investors	4,41%
Other asset managers	3,86%
Superannuation funds	1,39%
Family asset managers	1,33%
Other	1,08%
Fund manager contributions	0,46%
Academic institutions	0,40%
Funds of funds	0,07%

Source: MNB (2015)

Table 3. shows that the biggest investors are governmental institutions (58,52%). A great amount of EU resources was distributed to venture capital funds through the "Jeremie program" which we will detail later.

Venture capital funds usually invest equity into starter or early phase enterprises as target companies. The maturity of investments is 5-7 years with exiting plan after the investment horizon. The main difference between venture capital and private equity is that private equity finances the enterprises that are already over the starting phase and enter the expansion stage. The target investments are typically very risky, the expected return of the investors is also high, and the realized return on the investments must compensate the losses produced by investments failed.

Historical overview

From the 1989 to 1992 the Hungarian market was dominated only by the so-called "country-funds", the funds that invest in a defined country. This time privatization played a central role in the investments.

The average size of these funds was around 50 million dollars. From 1992 the so-called "regional funds" entered the market too, who concentrated on the Central-Eastern-European region. The size of these funds reached the volume of 100 to 200 million dollars. Until about the year of 2000 the focus of investments was mainly on technological financing. In the early years of 2000 classic venture capital investors have also appeared who made their first investments into start-up companies (Karsai, 2011). From 2005 to 2008 the so-called global funds have also launched their activity in Hungary.

The Hungarian market was especially attractive for new investments after joining the EU. Working capital investments continuously increased until the crisis of 2007. Between 2007 and 2009 the market gained some attractiveness even from the fact that the pace of Western-European investments became slower. By 2010 the crisis heavily affected the Hungarian VC capital market as well, and the volume of investments decreased significantly (Karsai, 2012).

In 2009 the *Jeremie I.* program was launched in Hungary. Sources were distributed among the funds in more stages through a tender system. The Jeremie program was founded by the European Committee together with the European Investment Fund. The program supported micro-, small- and medium-sized enterprises. This capital infusion gave a new impulse to the Hungarian venture capital investments. These days – partly due to the Jeremie program –, culture of start-up enterprises have become well known and accepted.

The New Hungary Venture Capital Program (also called as Jeremie I) distributed funds between 8 winning venture capital funds and their management companies. The total volume of these funds was about 48 bn HUF with at least 30% private sector investment and with a maximum of 70% state involvement in each fund. The size of the smallest and the largest fund was 4 bn HUF and 7,36 bn HUF respectively (MV Zrt., 2013).

The Hungarian state provided 27 billion HUF equity to local venture capital investors in 2012. Main sources came from the European Regional Development Fund in the framework of the New Széchenyi Venture Capital Program - Economy Development Operative Program 4. This program was also called *Jeremie II.* program. In the program 6 billion HUF could be invested via the so called Common Seed Fund subprogram to finance micro- or SMEs established within 3 years with a maximum annual sales revenue of 200 million HUF. The funds could invest the remaining 22,5 billion HUF through the Common Growth Fund subprogram to micro companies, SMEs and to medium-sized companies established within 5 years with a maximum sales revenue of 5 billion HUF (Invitation to Tender, 2012). The total size of all the funds was 41 bn HUF, the size of the smallest fund was 2.14 bn HUF, and the largest fund received 6.5 bn HUF (MV Zrt. 2013).

The subprogram of the Széchenyi Capital Program Common Growth Fund expanded further in several steps. In the stage named as Jeremie III. eight venture capital fund managements received 3 billion HUF each. (Project Result Proclamation, 2013).

Several events and meeting opportunities are organized for the leaders of the start-up enterprises and potential investors. Such an event was the IVSZ Start-up Conference in the organization of the Informatikai, Távközlési és Elektronikai Vállalkozások Szövetsége (Association of IT, Communication and Electronical Enterprises) and the Start-up Underground Events, which was organized at the Corvinus University of Budapest in 2013.

Seed capital funds today in Hungary

Despite Sahlman-Scherlis (2003) who considers the seed funds primarily as investors who finance idea companies, investment guidelines of the seed companies in Hungary often ask for a prototype and market validation from the target companies.

On Figure 2 the distribution of the venture capital funds is presented according to the lifecycles of the target companies. Between 2010 and 2014 the seed stage companies received 13,43% of total funds distributed.

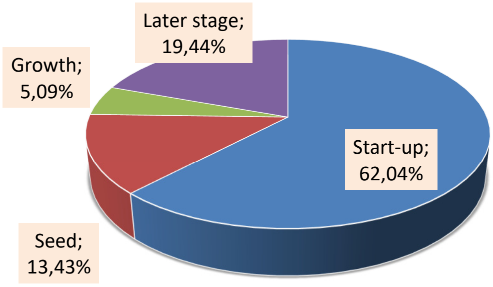


Figure 2: Distribution of Venture Capital Investments According to Lifecycle between 2010 and 2014

Source: MNB (2015)

The Hungarian accelerator called Traction Tribe set the objective to sell target companies to US venture capital investors, in other words to take them to the Silicon Valley, to one of the centers of starter companies. Between 2010 and 2015 39% of the US venture capital investments were executed in companies based in Silicon Valley (NVCA, 2016). Young start-up entrepreneurs come here from all over the world hoping that their idea will attract the investors' attention. As an example, iCatapult, a Hungarian Accredited Technological Incubator, also received a state subsidize. iCatapult also urges their supported start-up companies to build up US relations, to step onto the US market and contact US investors (Website of iCatapult). This is in strong contrast with the statement in the announcements of the National Research, Development and Innovation Office. Based on the statements, the goal of the program is to keep local start-up enterprises at home (NKFIH 2015).

STATE INVOLVEMENT ON THE VENTURE CAPITAL MARKET

The state is present on the market both as a venture capital investor and as a venture capital management. On the one hand the state appears as an equity holder in the capital funds, on the other hand there are venture capital fund management companies in state possession that invest state and EU funds. State intervention into the market mechanisms is necessary anyway if market failures occur. Market failures indicate that market mechanisms cannot create optimal market (Kovácsné 2011). These can appear in several forms and all can indicate market distortions: problems with public goods, presence of monopolistic and oligopolistic market participants, asymmetric information, transaction costs and externalities. (Lovas 2015). These latter three market failures are especially relevant from venture capital point of view. These failures are also responsible for the lack of financing for start-up companies, a typical market feature.

- The problem of asymmetrical information is very much in focus since the publication of Ackerlof (1970). Ackerlof demonstrates this effect on the lemons market, which we can interpret as example for financing start-up enterprises. Only the entrepreneur has any kind of information concerning the risk of the enterprise, the investor can just guess it (Leland–Pyle 1977). As the investor takes the average when it determines the conditions of financing, and he sets conditions that are adequate for the "bad" enterprises and not for the good ones. These conditions include also the investor's expected return. The investor's expected return for good start-up enterprises is too high, for the bad ones it is too favorable. In this case, the state can enter as the financer, as the expected return of one single investment of a state-owned capital fund manager is lower than the expected return of the private market capital funds.
- The second market failure is related to the topic of transaction costs. Start-up enterprises searching for financing occasionally require too negligible amounts. These amounts are not economic from private VC investors' point of view due to the relatively high transactional costs, administrative fees, and expert fees. For VC investors, it is not worth financing under a certain investment size (15 to 20 million). However, state actors may accomplish investments under this threshold.
- Finally, we also must mention externalities as a reason for market failures. State intervention on the venture capital market can stimulate local innovation, and social-economic development in a wider sense. It can follow more goals than pure profit goals, like the development of local, regional economy, job-creation, or increasing tax incomes as a fundamental base of social services. These can be translated as a positive externality that can justify state intervention.

State owned venture capital fund management and their funds

There were and there are many state-owned capital fund managers in Hungary during the last decade. The currently operating funds are as follows: Széchenyi Tőkealap-kezelő Zrt. (Széchenyi Capital Fund Management Plc.) manages the Széchenyi Tőkealap (Széchenyi Capital Fund). In 2002 the Informatikai Kockázati Tőkealap-kezelő Zrt. (IT Venture Capital Fund Management Plc.) started to manage a fund of 3 bn HUF. It was taken over by the Corvinus Tőkealap-kezelő Zrt. (Corvinus Venture Capital Fund Management Plc.) in 2015.

At the end of 2016 "Corvinus Tőkealap-kezelő Zrt." was renamed to HiVenture Venture Capital Fund Management Plc. It is planned to manage state fund of 50 bn HUF provided by the EU, in cooperation with the Hungarian Development Bank Plc. and the "Nemzeti Kutatási Fejlesztési és Innovációs Hivatal" (National Research, Development and Innovation Office).

State owned venture capital fund managements manage only state owned funds. Their expected returns are lower than those of the privately-owned capital fund managers, but investment decisions and processes are much more controlled. Therefore, the decision-making process is longer, which is also apparent during the cooperation phase with the target company after the investment.

Comparison of the characteristics of the private and the state venture capital

The investment structure of the state venture capital investors materially differs from those of the private venture capital investor companies. Venture capital fund management invests into equity. In the investment contracts, they define exit opportunities, the practice of ownership rights, voting rights, the decisional scopes of stakeholders, and the right to delegate members into distinct positions and boards (supervisory board members, board of directors, the person of the CEO, etc.) To identify the differences let us examine the characteristics of the private venture investment deals first.

The private venture fund managers concentrate on getting as big ownership stake as possible. If the target company becomes more valuable, then investors can realize substantial returns by the exit. The private investors focus on getting a majority share in the target companies to get control rights. They like to emphasize that they are strategic investors and partners with business network, market know-how. They also usually insist on including in the contract the so-called drag-along right, which obligates the founders to sell their shares together if the venture capitalist could set up an exit.

As opposed to that, state venture capital investors are typically financial investors: they do not wish to intervene into the everyday operation. They do not necessarily acquire a majority share in the target companies, their share usually remains under 49%. Thus, they leave the leadership in the hands of the original founders. Furthermore, state venture capital investors limit their profit potential on individual investments. The exit plan is that the target company will repurchase the fund’s share at the exit with a defined fixed expected rate of return. Capital investments are often combined with an ownership loan with continuous amortization to the exit. This can be considered as a risk mitigation step, which transforms state capital investments similar to hybrid financing. By these loans of course a lower interest is charged than the level of expected return on the equity.

State involvement creates the opportunity that the successfully developed enterprise could stay in the possession and control of the original founders and will not be sold necessarily to third (mainly foreign) investors.

CONCLUSION

Seed companies are in a difficult position in terms of financing. State intervention intends to solve this kind of market failure by providing state and EU funds. Distribution and utilization of EU funds are controlled by the state. On the Hungarian venture capital market two models of state intervention have been developed. The first form of state involvement is the indirect, when the state and the private sector cooperate. In that case, private venture capitalists manage funds containing state and EU sources, and the private venture capitalist attitude is dominating the investment process. The second form of state involvement is the direct intervention, when the state-owned venture fund managers directly control and monitor the investment process until the exit. Target companies can decide whether they need an active partner with higher return expectation or they would like to run their business alone beside a lower expected return from the financing partner.

REFERENCES

Ackerlof, G. A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism". *Quarterly Journal of Economics.* The MIT Press.

Damodaran, A. (2009): Valuing Young, Start-up and Growth Companies: Estimation Issues and Valuation Challenges. *Stern School of Business*, New York University.

Garbade, M. J. (2011): Differences in Venture Capital Financing of U.S., UK, German and French Information Technology Start-Ups - A Comparative Empirical Research of the Investment Process on the Venture Capital Firm Level, (June 15, 2010). Available at SSRN: https://ssrn.com/abstract=1819422 or http://dx.doi.org/10.2139/ssrn.1819422

Gompers, P. A. – Lerner, J. (2004): *The venture capital cycle.* The MIT press.

Harding, R. (2000): Venture Capital and Regional Developement: Towards a Venture Capital ’System’. Venture Capital, 4. pp. 287-311.

Karsai, J. (1997): A kockázati tõke lehetõségei a kis- és középvállalatok finanszírozásában. *Közgazdasági Szemle,* 44, February, pp. 165–174.

Karsai, J. (2002): Mit keres az állam a kockázatitõke-piacon? *Közgazdasági Szemle,* 49, 11, pp. 928–942.

Karsai, J. (2011): A kockázati tőke két évtizedes fejlődése Magyarországon.*Közgazdasági Szemle,* 58, 10, pp. 832–857.

Karsai, J. (2012): A kapitalizmus új királyai. Kockázati tőke Magyarországon és a közép-kelet-európai régióban. *Közgazdasági Szemle Alapítvány,* Budapest.

Kovácsné, A. A. (2011): Kockázatitőke-finanszírozás a hazai kis- és középvállalkozásokban. Doktori értekezés, *Kaposvári Egyetem*, Gazdálkodástudományi Kar.

Leland, H. E. – Pyle, D. H. (1977): Informational Asymmetries, Financial Structure, and Financial Intermediation. *Journal of Finance,* 1977, 32, 2, pp. 371-87.

Lerner J. (1999): The government as venture capitalist: The long-run effects of the SBIR programme. Journal of Business, 1999, 72, pp. 285–318.

Lovas, A. (2015): Innováció-finanszírozás aszimmetrikus információs helyzetben. Doktori Értekezés, *Budapesti Corvinus Egyetem,* Befektetések és Vállalati Pénzügy Tanszék.

MV Zrt. (2013): Új Magyarország Kockázati Tőke Program és Új Széchenyi Kockázati Tőke Program. Magyar Vállalkozásfinanszírozási Zrt., Budapest.

URL: http://www.mvzrt.hu/termekek/kockazati-toke/uj-magyarorszag-kockazati-toke-program-es-uj-szechenyi-kockazati-toke-programDate of download: 2016.10.30.

MNB (2015): Elemzés a hazai kockázati tőkealap-kezelők és alapok működéséről. Magyar Nemzeti Bank, Budapest.

NKFIH (2015): 2,1 milliárd forintos állami támogatás az ígéretes magyar vállalkozásoknak. Nemzeti Kutatási Fejlesztési és Innovációs Hivatal, Budapest.

NVCA (2016): *Yearbook 2016.* National Venture Capital Association, Washington.

OECD 1997: Government Venture Capital for Technology-Based Firms. Organization for Economic Co-operation and Developemet. Paris, 1997. OECD/GD(97)201.

Sahlman, W. – Scherlis, D. (2003): A Method for Valuing High-risk Longterm Investments. *Harvard Business School Press*, Boston.

Schertler, A. (2006): The venture capital industry in Europe; Palgrave macmillan, 2006, ISBN 978-0-230-50522-3

Sohl, J.E. (1999): The early-stage equity market in the USA. *Venture Capital: An international journal of entrepreneurial finance,* 1, 2, pp.101-120.

Project Result Proclamation (2013): *Az Új Széchenyi Kockázati Tőkeprogramok Közös Növekedési Alap Alprogram pályázati felhívásának eredményhirdetése.* URL: https://www.palyazat.gov.hu/az_uj_szechenyi_kockazati_tokeprogramok_kozos_novekedesi_alap_alprogram_palyazati_felhivasanak_eredmenyhirdeteseDate of download: 2016.10.30

Invitation to Tender (2012): Pályázati Felhívás a Gazdaságfejlesztési Operatív Program 4. Prioritás keretében finanszírozott ÚJ SZÉCHENYI KOCKÁZATI TŐKEPROGRAMOK Közös Magvető Alap Alprogramja közvetítőinek kiválasztására Kódszám: GOP-2012-4.3/A (2012)URL: https://www.palyazat.gov.hu/doc/3531Date of download: 2016.10.30

Website of Traction Tribe (URL): http://traction-tribe.com/ Letöltés dátuma: 2016.10.30

Walter, Gy. (2014a): Vállalatfinanszírozás a gyakorlatban - Lehetőségek és döntések a magyar piacon. Vállalatfinanszírozási lehetőségek. *Alinea Kiadó,* Budapest.

Walter, Gy. (2014b): Vállalatfinanszírozás a gyakorlatban - Lehetőségek és döntések a magyar piacon. Az állami támogatások. *Alinea Kiadó,* Budapest.

Websites of fund managers (URL): http://conorfund.com/http://www.krscapital.hu/http://www.coreventure.hu/http://www.dayonecapital.com/
Date of download: 2016.10.30.

Website of iCatapult (URL): http://icatapult.co/
Date of download: 2016.10.30

Authors

- Endre Mihály Molnár, PhD student, Corvinus University of Budapest (Budapest)
- Dr. Erika Jáki, PhD, senior lecturer, Corvinus University of Budapest (Budapest)

AUTHOR BIOGRAPHIES

Dr. Erika Jáki was born in Budapest, Hungary and went to the Corvinus University of Budapest (CUB), where she studied finance and marketingcommunication and obtained her degree in 2001. During her study, she was taking part in the CEMS program and obtained the CEMS degree in 2002. From 2004 she made her PhD study and she obtained her PhD degree in 2013. She worked for the CUB from 2008, and from 2013 as assistant professor and she is responsible for the Business Planning curse. She makes internal audit for MFB Invest Plc. from 2014 and Hiventures Venture Capital Funds Management Plc. from 2013. She does research in the field of early stage investments and behavioral finance. Her e-mail address is: jaki.erika@t-online.hu.

Endre Mihály Molnár was born in Budapest, Hungary and went to the Corvinus University of Budapest, where he studied finance and obtained his master's degree in 2016. He worked for a Hungarian venture capital fund management company for more than two years before starting his PhD at the Corvinus University of Budapest. He does research in the field of early stage investments. His e-mail address is : bandoolero@gmail.com

FACTORS ASSOCIATED WITH THAI EXPORTER'S INTEREST IN USING NEW DAWEI DEEP SEAPORT

Kanogkan Leerojanaprapa
Department of Statistics
King Mongkut's Institute of Technology Ladkrabang (KMITL)
Bangkok, Thailand, 10520
kanogkan.le@kmitl.ac.th

Kittiwat Sirikasemsuk
Department of Industrial Engineering
King Mongkut's Institute of Technology Ladkrabang (KMITL)
Bangkok, Thailand, 10520
kittiwat.sirikasemsuk@gmail.com

Komn Bhundarak
Thammasat Business School
Thammasat University
Bangkok, Thailand, 10200
komn@tbs.tu.ac.th

KEYWORDS

Port selection, Decision factor, Pearson Chi-Square test, t-test.

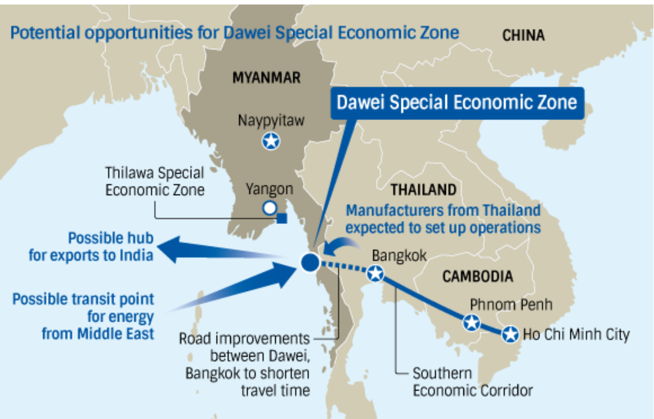
ABSTRACT

Dawei deep seaport in a part of the Dawei Special Economic Zone (Dawei SEZ) in Myanmar aims to support the new economics along the GMS Southern Corridor. The Dawei seaport can serve the potential new industries along the new industry zones. The new port will be the alternative route for Thai exporters in the future as it is under construction. The exploratory study by employing survey was selected and analyzed to identify the significant influencing factors. The results of hypothesis testing by Pearson Chi-Square test confirm the relation between the interest of using new Dawei deep seaport and the location of manufacturer (p-value = 0.027). In addition, the results of t-test confirm the significant six decision variables of Time for transportation (v7), Reliability of service (v10), Port size and capability (v15), Facility (v25), Professionals and skilled labors in port operation (v30), and Port accessibility (v31) are more important for exporters who are interested in Dawei seaport than the exporters who may not be intend to use the new seaport, p-value (1-tailed) < 0.05.

INTRODUCTION

The concept of transportation along a land bridge to a port can support the policy of Economic East-West Corridor sub-region GMS (Greater Mekong Sub-region) or GMS Southern Corridor. Therefore, the GMS Southern Economic Corridor links Dawei, Myanmar through Bangkok-Chonburi and Trat, Thailand to Phnom Penh and Sihanoukville, Cambodia to Nam Can, Vietnam. One of the big project is the development of the Dawei Special Economic Zone (Dawei SEZ), see Figure 1. In addition, the Dawei deep seaport will be a Western gateway through the Indian Ocean and the port can support to increase the competitiveness of countries in this region. The advantages of the Dawei seaport can support exporters to ship their products to the Indian Ocean instead of shipping via Malay Peninsula. The establishment of this project can also stimulate the development of transport routes and infrastructures across the corridor and then follow by the development of significant economic activity (Cabral & Ramos 2014) as the hinterland of the Dawei seaport.

Dawei city is in the southeast of Myanmar, about 360 km from Yangon and about 138 km to the Myanmar-Thailand border Baan Phu Nam Ron. The Dawei SEZ is officially demarcated by the Republic of the Union of Myanmar under the SEZ Law, enacted in January 2014. The Dawei SEZ is managed by Myandawei Industrial Estate Company Limited (MIE). The Dawei SEZ has also been supported by the joint cooperation between the governments, the Republic of Myanmar and Thailand (Myandawei Industrial Estate Company Limited, 2016).


Potential opportunities for Dawei Special Economic Zone
CHINA
MYANMAR
Dawei Special Economic Zone
Naypyitaw
Thilawa Special Economic Zone
Yangon
THAILAND
Manufacturers from Thailand expected to set up operations
Possible hub for exports to India
Bangkok
CAMBODIA
Possible transit point for energy from Middle East
Phnom Penh
Ho Chi Minh City
Road improvements between Dawei, Bangkok to shorten travel time
Southern Economic Corridor

Source: Motoka et. al. (2015)

Figures 1: GMS Southern Corridor and Dawei SEZ

The Dawei deep seaport can be an alternative route for exporters not only in the Dawei industrial estate but also in Thailand. The Dawei deep seaport is a competitive port of seaports in Thailand since Dawei is not far away from Bangkok. In order to understand the current Thai exporters' perceptions of using Dawei deep seaport, this research aims to explore their Thai exporters' opinion by survey. All relevant factors are investigated in order to identify significant relation between those decision factors and Thai exporters' interest of using Dawei seaport.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

The structure of the remaining of this paper is therefore organised as follows. The Section 2 shows critical literature review of possible factors that are relevant to the export route selection. The data collection and instrument construction are described in Section 3. Results analysed from obtained survey data are described in Section 4. Finally, in section 5, the interpretation of the results are defined.

RELEVANT FACTOR IDENTIFACTION

This section shows the relevant factors in decision making of exporting routes. According to the critical review, two main distinct sources were specified as factors related to internal factors and external factors.

Internal factors

The characteristics of particular companies can influence the port choice selection. This research defined the 4 characteristics as the internal factors as Size, Type of business, Product type, and Manufacturer location.

Size
Size of a company can influence the decision to remodel the exporting route via the new port (Pokharel 2005). First, the large company may not be flexible to make change since they have invested to facilities along the route and it is difficult to move or change. Second, some global companies may not be interested in changing to the new port because their headquarter branches may have the contracted freight company that won the global bidding for all branches in the region.

Type of business
Two main types of business are relevant to the export activities as Import-Export, Production for export or both activities. Different types of business can perceive criteria to select seaport for export differently (Kent & Stephen Parker 1999; Meixell & Norbis 2008).

Product type
Some products require the special needs for transportation especially via seaport. Different products can come with different sizes and weights so those are relevant to the availability of mode of transportation (Meixell & Norbis 2008).

Manufacturer location
The distance to the port is a critical factor for selecting the port and manufacturer location since it leads to overall cost reduction (Manic 2013). The location has been confirmed by recent research (Chang et al. 2008; Lee Lam & Song 2013; Lirn et al. 2004; Park & Min 2011).

External factors

The external factors were reviewed from different papers such as the competitiveness of ports, the efficiency of logistics performance, or the influential attributes in mode choice decision etc. There are 7 criteria and 31 factors were defined as below.

Cost
Transport costs are an important factor in supply chain costs and it is defined and needs to be reduced to increase competitiveness. Operators want to spend in the most cost-effective way (Foster 1978). The cost of particular processes during export route is defined into 6 factors as:

1. Transportation cost which includes inland transportation cost to pay for the vehicle charges and sea freight shipment cost.
2. Terminal handing charge which includes the dock charges, berth fees, electricity, etc.
3. Multimodal operation cost which includes cost of worker, cargo loading or discharging fees, cost transshipment, etc.
4. Customs regulation cost which includes customs fees, costs and expenses, port authority documents, cost for special permits, etc.
5. Insurance cost which is the cost to pay for the coverage of unexpected events for export shipment along either inland or sea transport.
6. Cargo storage fee or container storage fee which is employed when the waiting time is longer than the allowance.

Time
Time is one of the main criteria for a seaport selection. The individual activities require either different operating or waiting time (Kofjac et al. 2009; Kopytov & Abramov 2013).

7. Time for transportation refers to the time duration of the transportation to the destination.
8. Transferring time means time spent in transit both unloading and loading from one mode to another until products reach the destination countries.
9. Customs service time is the time to spend for conducting customs clearance. This may include the crossing border time.

Reliability
Reliability of the operations or services along the routes and there were explained in literature (Kopytov & Abramov 2013; Yeo et al. 2008; Manic 2013; Panayides & Song 2012; Tongzon 2009).

10. Reliability of service is the delivering accuracy.
11. Safety in the export route is the secure throughout transport routes. The product and package are not damaged or stolen during transportation.

12. Safety during handle transferring means the safety of the product from damage and lost in the process of loading and unloading at each point as well as the security of keeping products in a cargo.

13. Traffic condition considers the traffic along the route both inland and in the sea.

14. Capability to handle transferring from one mode to another is a critical competitiveness of the ports. The available facilities along the transport network should be in a good condition so it will take less time and reduce the likelihood of accidents.

Port Efficiency

The port performance can attract entrepreneurs to choose the port (Langen et al. 2007; Manic 2013; Tongzon 2009).

15. Port size and capability indicate the ability of the port to handle number of cargos, space of container yards to store containers during waiting to transport, etc.

16. Frequency of ship visits can cause a variety of the price competition and flexibility for the operators to have more schedules and can reduce waiting time.

17. Inter-modal link is the ability to link the port with inland. If the connection is inconvenient, it can lead to congestion and higher costs.

18. Port facility and infrastructure is necessary element to provide efficient services. For example; all necessary equipment e.g. cranes should be sufficient.

Existing Resources

Existing resources were defined in literature (AEC, ENRICH, PCBK, CMCL & PTL, 2000).

19. Infrastructure availability means facilities, resources, or devices for existing operators invested to support the current use of logistics. If they decide to change the route of transport, operators may be concerned about the possibility of using their available resources within the new route.

20. Familiarity of the routes is the confidence of along the routes that exporters can cope with unexpected problems during the transport process.

21. Balancing between inbound and outbound represents the utility of the vehicle transporting loads both inbound and outbound in the same route.

Legislations and Basic Factors

Legislation and basic factors are related to the convenience of the transport operations and logistics. There are factors defined in the literature (Park & Min 2011; Saeed & Aaby 2013; Chou 2010).

22. Customs regulation is crucial to the port selection, especially for exporting to destination country. When the customs regulation rules of particular countries along the route are different, it will be difficult to cope with too many rules.

23. Government policy on investment means the government policy that can encourage or discourage relevant projects to promote the establishment of new businesses or to encourage existing businesses.

24. Political condition means a factor that may affect the delay of the process and the safety of transportation along the route.

25. Facility is the necessary infrastructure such as electricity supply system, phone line or Internet including transportation links such as roads for transportation from one place to another.

Port Service

This characteristic arises after the port was operated (Panayides & Song 2012; Langen et al. 2007).

26. Port customer service quality is the service at the port to meet customers' needs in time.

27. Port flexibility is the port service that can cope with the special requirements from customers. The service can be adjusted to meet the customers' needs.

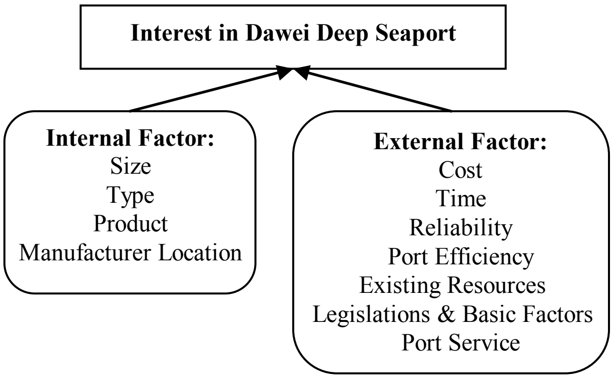
28. Efficiency of port management means the operation capability to manage port efficiency by concerning about speed and waiting time by providing the schedule appropriately.

29. Port information system is relevant to information sharing in the whole supply chain to enhance the logistics accuracy, reduce waiting time and increase speed of service.

30. Professionals and skilled labors in port operation are needed since labors contribute significantly to the efficiency of port services.

31. Port accessibility is channel to access to the port services such as customers service office hours, the working hours of customs at the port, etc.

According to the literature review above, the conceptual framework for this research is defined as Figure 2.



Figures 2: Conceptual framework

RESEARACH METHODOLOGY

Data collection

Data for this research were collected by mean of questionnaire that were sent via email with both a copy of questionnaire and a link to the online questionnaire to the members of Thai National Shippers' Council during October 2015 and May 2016. Therefore, Quota

Sampling was employed for this study with 150 samples and only one representative defined as a respondent for a company. The final returned questionnaires were 157 samples since extra 5% were prepared to prevent missing and incomplete questionnaire. The respondents are the staff who have main tasks relevant to the export process in the companies that have experience in exporting through major ports of Thailand.

Research instrument

Questionnaire is a structured questionnaire and the questionnaire is divided into three parts as:

Part 1: Overview and characteristics of the establishment;

Part 2: Factors influencing the choice of exporting ports which were evaluated by 5 levels of important factors (Table 1);

Part 3: The interest to the Dawei deep seaport.

Table 1: 5 Point Different Levels of Important Factors for Export Routing Decision Making

Level	Meaning
1	Insignificant
2	Slightly important
3	Fairly important
4	Very important
5	Vital

Statistical analysis

Data analysis was performed by using the Statistical Package for the Social Sciences (SPSS) Version 19.0 for Windows.

Test of dependence

Test of dependence was employed to investigate the factors related to the exporters' interest in the Dawei deep seaport as the main purpose of this research. Pearson Chi-Square (χ^2) was selected as the most suitable techniques for collected data of this research within nominal scales. The Pearson's chi-square test using χ^2 statistic plays the important role for testing of independence between two categorical variables. Consequently, the null hypothesis asserts the independence of variables under consideration. χ^2 statistic can be calculated as follows:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}} \qquad (1)$$

$$E_{ij} = \frac{(r_i)(c_j)}{n} \qquad (2)$$

Where:

O_{ij} = Observed frequency in ith row and jth column

E_{ij} = Expected frequency in ith row and jth column

r_i = Total frequency in ith row

c_j = Total frequency in jth column

n = Total number

r = Number of rows

c = Number of columns

The χ^2 can then be used to calculate a p-value by comparing the value of the statistic to a chi-squared distribution with (r-1)(n-1) degree of freedom. The limitations of the test should be ensured that the number of cells that $E_{ij} < 5$ should not be more than 20%.

Independent t-test

The independent-samples t-test is an inferential statistical test that determines whether there is a statistically significant difference between the population means in two unrelated groups. One tailed test of hypothesis can be employed for this study to compare the mean score of 'interested group' (group 1) will exceed the mean scores of 'not interested group' (group 2) to use the Dawei seaport in particular factors. The relevant decision factors to the interest of using Dawei seaport can be identified from the excess of the sample means which is large enough to be statistically significant evidence as defined by the hypothesis below.

H$_0$: $\mu_1 - \mu_2 \le 0$

H$_1$: $\mu_1 - \mu_2 > 0$

Two cases of test statistic are defined as:

Case 1: $\sigma_1^2 = \sigma_2^2$

$$t = \frac{\left(\bar{x}_1-\bar{x}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{\left(n_1-1\right)S_1^2+\left(n_2-1\right)S_2^2}{n_1+n_2-2}} * \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \qquad (3)$$

$$df = n_1 + n_2 - 2$$

Case 2: $\sigma_1^2 \ne \sigma_2^2$

$$t = \frac{\left(\bar{x}_1-\bar{x}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \qquad (4)$$

$$df = \left[\frac{\mathrm{S}_1^2}{\mathrm{n}_1}+\frac{\mathrm{S}_2^2}{\mathrm{n}_2}\right]^2 \Bigg/ \left(\frac{\left[\frac{\mathrm{S}_1^2}{\mathrm{n}_1}\right]^2}{\mathrm{n}_1-1}+\frac{\left[\frac{\mathrm{S}_2^2}{\mathrm{n}_2}\right]^2}{\mathrm{n}_2-1}\right) \qquad (5)$$

Where:

n_i = Number of sample for group i, i = 1 and 2

$\bar{x}_i$ = Sample mean of group i, i = 1 and 2

S_i = Sample standard deviation of group i, i = 1 and 2

RESULTS

Characteristics of respondents

Table 2 shows the characteristics of the respondents. The respondents are working in companies with different sizes. Most of the respondents work in the medium size company with 50-200 employees. Most of their companies mainly produce the products for export (and also domestic market) (68.2%). They work in different industries and the majority of product is Food/Beverages industry (17.8%). Furthermore; their company is located in different areas of Thailand and the majority of their manufacturers is located in the outskirt around Bangkok (39.3%).

Table 2: Respondent Characteristics

Characteristics	Frequency	%
Size		
< 50 employees	32	20.5
50-200 employees	47	30.1
201-500 employees	32	20.5
501-1,000 employees	21	13.5
> 1,000 employees	24	15.4
Total	156	100.0
Type of Business		
Import-Export	30	19.4
Production for export	105	68.2
Both	19	12.3
Total	154	100.0
Product Type		
Agriculture and agricultural products	13	8.9
Automotive and automotive parts	7	4.8
Chemical/Petrochemical	5	3.4
Construction materials	4	2.7
Electrical products /Electronic equipment and parts/Software	21	14.4
Food/Beverages	26	17.8
Furniture	11	7.5
Leather	2	1.4
Plastic/Plastic packaging	7	4.8
Rubber and rubber products	7	4.8
Textile and clothing	13	8.9
Other	30	20.5
Total	146	100.0
Manufacturer Location		
Bangkok	22	15.7
Outskirt	55	39.3
East	30	21.4
Middle	3	2.1
West	8	5.7
North	10	7.1
North East	4	2.9
South	8	5.7
Total	140	100.0

Table 3: Number of Samples, Standard Deviation (S.D.), Mean, and Rank of Export Routing Decision Factors

Factors	n	S.D.	Mean	Rank
Cost				
1. Transportation cost	156	0.807	4.519	**4**
2. Terminal handing charge	156	0.925	4.269	18
3. Multimodal operation cost	156	0.910	4.173	23
4. Customs regulation cost	156	0.980	4.019	27
5. Insurance	154	1.113	3.948	29
6. Cargo storage fee or container storage fee	155	1.308	3.813	31
Time				
7. Time for transportation	156	0.686	4.481	**5**
8. Transferring time	155	0.966	4.187	22
9. Customs service time	156	0.818	4.314	15
Reliability				
10. Reliability of service	156	0.540	4.724	**2**
11. Safety in the export route	156	0.528	4.750	**1**
12. Safety during handle transferring	154	0.593	4.708	**3**
13. Traffic condition	154	0.773	4.318	14
14. Capacity to handle transferring from one mode to another	152	0.742	4.421	9
Port Efficiency				
15. Port size and capability	154	0.786	4.273	17
16. Frequency of ship visit	154	0.727	4.344	13
17. Inter-modal link	154	0.831	4.364	12
18. Port facility and infrastructure	154	0.934	4.149	25
Existing Resources				
19. Infrastructure availability	152	0.887	3.967	28
20. Familiarity of the route	152	0.897	3.868	30
21. Balancing between inbound and outbound	152	0.848	4.059	26
Legislations and Basic Factors				
22. Customs regulation	154	0.746	4.409	10
23. Government policy on investment	154	0.893	4.266	19
24. Political condition	154	0.920	4.234	20
25. Facility	154	0.838	4.208	21
Port Service				
26 Port customer service quality	155	0.706	4.477	6
27. Port flexibility	156	0.687	4.308	16
28. Efficiency of port management	156	0.675	4.449	7
29. Port information system	156	0.685	4.378	11
30. Professionals and skilled labors in port operation	156	0.777	4.167	24
31. Port accessibility	156	0.701	4.429	8

The prioritization of factors influencing the export route choice

The respondents evaluated level of importance of decision factors in 31 factors. The number of valid answers by respondents in particular questions was counted. Furthermore; standard deviation, mean, and rank sorted by mean are shown in Table 3. Respondents evaluate the importance of each factor more than four score, except only 4 factors as Insurance (v5), Cargo storage fee or container storage fee (v6), Infrastructure availability (v19), and Familiarity of the route (v20). In addition, the respondents defined the top three of the most important factors in reliability criteria as Safety in the export route (v11), Reliability of service (v10), and Loss and damage during handle transferring (v12) as 4.750, 4.724, and 4.708 respectively. The forth important factor is Transportation cost (v1) and the average score is 4.519. The fifth rank is Time for transportation (v7) and the average score is 4.481.

The interest of using new Dawei deep seaport

Table 4 shows that 25.34% of the respondents are interested in using the Dawei seaport. However, the majority of respondents are still unsure since they need more concrete information of the seaport when the seaport can commercially be operated.

Table 4: Interests in Dawei Deep Seaport

Openions	Frequency	%
Interest	37	25.34
Not interest or Uncertain	109	74.66
Total	146	100.00

The dependent factors associated to the interest of using Dawei deep seaport

Table 5 describes the result of Pearson Chi-Square test of independence. Under the condition of Pearson Chi-Square test, number of cells which show the expected count (see Equation (2)) of particular cells should not be less than 5 more than 20% of all cell in the contingency table. Therefore, the category of variable are adjusted by reducing the levels of importance for decision factors into 3 groups as fairly to lower (≤ 3), very important (4), and vital (5). However, after some cells are merged, the remain classes are not relevant as some variables still show number of cells that $E_{ij} > 5$ is a bit greater than 20% as shown in the remarks of the Table 5.

Table 5: Pearson Chi-Square and p-value for Test of Independence

Factors	Pearson Chi-Square	p-value
Size	1.954	0.751
Type	1.344[a]	0.563
Product	5.956[b]	0.669
Location	12.453[c]	0.027*

[a] 1 cells (16.7%) has expected count less than 5.

[b] 4 cells (25.0%) have expected count less than 5.

[c] 3 cells (25.0%) have expected count less than 5.

* p-value < 0.05

Table 6: t-test and p-value for Test of Mean Difference

Factors	$\bar{x}_1$	$\bar{x}_2$	t	Sig (1-tailed)
Cost				
1. Transportation cost	4.622	4.481	0.899	0.185
2. Terminal handing charge	4.297	4.231	0.367	0.357
3. Multimodal operation cost	4.135	4.176	-0.232	0.409
4. Customs regulation cost	4.027	4.000	0.144	0.443
5. Insurance	4.000	3.925	0.35	0.364
6. Cargo storage fee or container storage fee	3.811	3.757	0.212	0.416
Time				
7. Time for transportation	4.703	4.389	2.85	0.003*
8. Transferring time	4.351	4.103	1.34	0.091
9. Customs service time	4.297	4.306	-0.055	0.478
Reliability				
10. Reliability of service	4.838	4.667	1.869	0.033*
11. Safety in the export route	4.811	4.713	0.95	0.172
12. Safety during handle transferring	4.838	4.682	1.7	0.465
13. Traffic condition	4.324	4.327	-0.019	0.493
14. Capacity to handle transferring from one mode to another	4.459	4.419	0.289	0.387
Port Efficiency				
15. Port size and capability	4.486	4.196	1.977	0.025*
16. Frequency of ship visit	4.432	4.308	0.901	0.185
17. Inter-modal link	4.514	4.336	1.159	0.124
18. Port facility and infrastructure	4.378	4.075	1.761	0.040
Existing Resources				
19. Infrastructure availability	4.000	3.981	0.111	0.456
20. Familiarity of the route	3.946	3.822	0.712	0.239
21. Balancing between inbound and outbound	4.054	4.084	0.186	0.427
Legislations and Basic Factors				
22. Customs regulation	4.486	4.346	0.974	0.166
23. Government policy on investment	4.351	4.215	0.804	0.212
24. Political condition	4.297	4.150	0.831	0.204
25. Facility	4.417	4.159	1.665	0.049*
Port Service				
26 Port customer service quality	4.568	4.495	0.574	0.284
27. Port flexibility	4.324	4.324	0.002	0.499
28. Efficiency of port management	4.514	4.444	0.542	0.294
29. Port information system	4.514	4.352	1.25	0.107
30. Professionals and skilled labors in port operation	4.378	4.120	1.766	0.040*
31. Port accessibility	4.595	4.370	1.706	0.045*

* p-value < 0.05

The hypothesis is determined by considering the p-value which is shown as Exact Sig. (2-sided) in the SPSS output. If p-value < 0.05, dependence between both variables is significant. For internal factors, only Location of the manufacturer is significantly dependent on the interest of using new Dawei deep seaport at a significance level of 0.05 (p-value = 0.027). On the other hand, p-value of Size, Type of business, and Product type is greater than 0.05, so there is no relationship between those internal factors and the interest of the Dawei seaport.

The results of statistical testing of means in Table 6 are also confirmed that the mean score of the importance of Time for transportation (v7), Reliability of service (v10), Port size and capability (v15), Facility (v25), Professionals and skilled labors in port operation (v30), and Port accessibility (v31) given by Thai exporters who are interested in using Dawei ($\overline{x}_1$) exceeds another group ($\overline{x}_2$). Those are significant external factors since t-value is positive and the p-value (1-tailed) < 0.05 at 5% level of significance. In other words, the data provide sufficient evidence that Thai exporters who are interested in Dawei seaport emphasize those 6 external decision factors more than Thai exporters who are not interested in Dawei seaport.

CONCLUSIONS AND FURTHER STUDIES

It is clear that Dawei deep seaport is under construction so it is unable to compare their performance indicators with the standards of general seaport. At the early point of Dawei seaport construction, if they can define the influencing decision factors of their potential customers, they can provide the proper policies to meet their future customers' expectation. Therefore, this study used survey to explore the relationships between Thai exporters' perceptions of decision factors and their interest in using new export route via the Dawei seaport.

There are only one internal factor and six external factors which are significantly relevant to the interest of using new export route via the Dawei seaport. It is found that location and time are highly significant because the location factor is a concordance factor to the time for transport. Any seaport would directly affect the industries in the hinterland of the port, manufacturer who are in the Lower Mekong sub-region (GMS Southern Corridor) would show their interest to the Dawei seaport. The respondents prioritise the time of transportation as the important factor for their company, they will be interested in the Dawei seaport since they can save time 2-3 days via the Malay Peninsula and they agreed to the important of time of transportation to the export process. Furthermore, if the Dawei seaport project developer would like to be successful and maintain their potential customers such as Thai exporters, they should be able to maintain their expectations of the six significant factors as the Time for transportation (v7), Reliability of service (v10), Port size and capability (v15), Facility (v25), Professionals and skilled labors in port operation (v30), and Port accessibility (v31).

Interestingly, some important factors relevant to cost are not significantly relevant to the interest of Dawei seaport. Although the exporters who are not interested in Dawei seaport think that those factors are very important (high rank), they may not prioritise the importance of the cost higher than the exporters who are not interested in Dawei seaport large enough to be statistically significant for many reasons. First, they may have to pay higher cost of inland transport so they may not ensure that the overall cost will be decreased. Second, the Dawei SEZ and the deep seaport project has been delayed so it can affect Thai exporter' confidence in this project. Finally, a little information has been promoted and passed to Thai exporters so it is unclear to the advantages of using Dawei seaport.

ACKNOWLEDGEMENT

This paper is a part of a project fully funded by the Thailand Research Fund (TRF) under Grant 2015. The authors would like to thank Thai National Shippers' Council and research assistants to support during data collection. In addition, the appreciation also goes to all representatives of particular export companies to be participants in this survey.

REFERENCES

AEC, ENRICH, PCBK, CMCL & PTL, 2000. *The feasibility study and preliminary design of freight rail link between the port of the Gulf of Thailand and Andaman*. Retrieved Jan 4, 2017, from http://webcache.googleusercontent.com /search?q=cache:http://www.pcbk-world.com/field-expertise/crw0126r.html&gws_rd=cr&ei=uVPTWInsBcX ZvATGoaHACw

Cabral, A.M.R. & Ramos, F. de S., 2014. Cluster analysis of the competitiveness of container ports in Brazil. *Transportation Research Part A: Policy and Practice*, 69, pp.423–431.

Chang, Y.-T., Lee, S.-Y. & Tongzon, J.L., 2008. Port Selection Factors by Shipping Lines: Different Perspectives between Trunk Liners and Feeder Service Providers. *Marine Policy*, 32(6), pp.877–885.

Chou, C.-C., 2010. AHP Model for The Container Port Choice in The Multiple-Ports Region. *Journal of Marine Science and Technology*, 18(2), pp.221–232.

Foster, T., 1978. Ports: What Shippers Should Look For. *Distribution Worldwide*, 6(3), pp.32–36.

Kent, J.L. & Stephen Parker, R., 1999. International containership carrier selection criteria. *International Journal of Physical Distribution & Logistics Management*, 29(6), pp.398–408.

Kofjac, D., Kljajic, M. & Rejec, V., 2009. The anticipative concept in warehouse optimization using simulation in an uncertain environment. *European Journal of Operational Research*, 193(3), pp.660–669.

Kopytov, E. & Abramov, D., 2013. Multiple-Criteria Choice of Transportation Alternatives in Freight Transport System for Different Types of Cargo. In *the 13th International Conference Reliability and statistics in Transportation and Communication*. pp. 180–187.

Langen, P. de, Nidjam, M. & Horst, M. van der, 2007. New

Indicators to Measure Port Performance. *Journal of Maritime Research*, 4(1), pp.23–36.

Lee Lam, J.S. & Song, D.-W., 2013. Seaport Network Performance Measurement in the Context of Global Freight Supply Chains. *Polish Maritime Research*, 20(Special Issue), pp.47–54.

Lirn, T.C. et al., 2004. An Application of AHP on Transhipment Port Selection: A Global Perspective. *Maritime Economics & Logistics*, 6(1), pp.70–91.

Manic, B., 2013. Benchmarking Analysis of Port Services from a Perspective of Freight Forwarders.

Meixell, M.J. & Norbis, M., 2008. A review of the transportation mode choice and carrier selection literature M. Waller, ed. *The International Journal of Logistics Management*, 19(2), pp.183–211.

Panayides, P. & Song, D.-W., 2012. Determinants of User's Port Choice. In W. K. Talley, ed. *The Blackwell Companion to Maritime Economics*. Oxford, UK: Wiley-Blackwell, pp. 599–622.

Park, B. & Min, H., 2011. The Selection of Transshipment Ports Using a Hybrid Data Envelopment Analysis/Analytic Hierarchy Process. *Journal of Transportation Management*.

Pokharel, S., 2005. Perception on information and communication technology perspectives in logistics. *Journal of Enterprise Information Management*, 18(2), pp.136–149.

Saeed, N. & Aaby, B., 2013. Analysis of Factors Contributing to Criteria for Selection of European Container Terminals. *Transportation Research Record: Journal of the Transportation Research Board*, 2330, pp.31–38.

Tongzon, J.L., 2009. Port Choice and Freight Forwarders. *Transportation Research Part E: Logistics and Transportation Review*, 45(1), pp.186–195.

Yeo, G.-T., Roe, M. & Dinwoodie, J., 2008. Evaluating the Competitiveness of Container Ports in Korea and China. *Transportation Research Part A: Policy and Practice*, 42(6), pp.910–921.

AUTHOR BIOGRAPHIES

Kanogkan Leerojanaprapa was born in Bangkok, Thailand and went to the Thammasat University, where she studied Applied Statistics and obtained her degree in 1999. She continued to study Master degree in Statistics in Chulalongkorn University and obtained her degree in 2002. She worked for four years for King Mongkut's Institute of Technology Ladkrabang (KMITL) before doing her PhD in 2008, University of Strathclyde, UK. After her graduation, she returned to KMITL where she is now a lecturer in Statistics department. Her research focus in the area of supply chain risk and statistical analysis. Her current e-mail address is : `kanogkan.le@kmitl.ac.th`

Kittiwat Sirikasemsuk is an Assistant Professor in Industrial Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Thailand. He received Doctor of Philosophy (Ph.D.) from Industrial Systems Engineering, Asian Institute of Technology, Thailand, in 2013. His teaching and research interests include design of experiments, supply chain design, measures of bullwhip effect in supply chains, quality engineering. He has published articles in various peer reviewed international journals. His email address is : `kittiwat.sirikasemsuk@gmail.com`

Komn Bhundarak was born in Bangkok, Thailand and went to the Thammasat University, where he studied Business Administration in Industrial Management and obtained his degree in 1985. He worked for 3M Thailand more than 10 years while pursuited MBA from Thammasat University in 1989 and MSIT from Kasetsart University in 2001. Then he became a lecturer in 2009 and earned his Doctoral in Business with Management since 2014 from University of Plymouth, UK. After his graduation, he returned to Thammasat Business School, Thammasat University where he is now a lecturer in Operations Management department. His research focus in the area of supply chain management and data analytics. His e-mail address is : `komn@tbs.tu.ac.th`

VALUATION OF THE PREPAYMENT OPTION IN THE BANKING BOOK

Petra Kalfmann, PhD
János Száz, CSc
Ágnes Vidovics-Dancs, PhD, CIIA

Corvinus University of Budapest
H-1093, Fővám tér 8, Budapest, Hungary
E-mail: kalfmannp@gmail.com

KEYWORDS

Mortgage loans, prepayment risk, affine term structure

ABSTRACT

One of the most important perspectives of interest rate risk in the banking book (IRRBB) is the valuation of so-called embedded options and the quantification of their impact on the value of bank portfolios. One unequivocal characteristic of mortgage portfolios is the option of prepayment, providing the borrower the possibility of redeeming their debt before maturity. We have created a theoretical model for valuing the prepayment option, based on which it can be demonstrated that, depending on the composition of the given bank portfolio (interest rate level, term to maturity), the prepayment option may have a significant effect on the sum of short-term interest income, as well as on the discount value of the bank portfolio via changing cash flows, and through this the value of economic capital. In this paper we analyse the impact of a possible model specification risk and the impact of changing the composition of banking portfolio under investigation.

INTRODUCTION

Mortgage based loans have become widely known for the public during the crisis of 2008. The importance of the issue and its widespread impact has reached even Hollywood – it is enough to mention only the film “Big short” which is excellent from professional point of view as well.

The paper we are about to present for the conference is related to the previously published analysis and results of Petra Kalfmann related to impact of loan prepayment on the value of the banking book (Kalfmann (2016)). The conclusions of the published model are based on simulation results of the stochastic Cox-Ingersoll-Ross (CIR) interest rate model, analysing the impact of prepayment on a mortgage portfolio in case of ascending and descending yield curves. We analyse the problem further in the following directions:

- we quantify the effect of a possible *model specification risk*;
- we analyse the impact of the changes in the *composition of the banking portfolio* on the original results, considering to change the composition in a way that the average interest rate level and maturity doesn’t change, but its deviation.

THE PROBLEM AND THE BASIC MODEL

The logical framework of the original calculations is the following.

a) *Yield curve modelling*. Based on the CIR model, 30-year time horizon, monthly intervals.

b) *Determining the par yield curve*, as current refinancing interest rates.

c) *Determining the refinancing incentive*, based on a comparison between the par interest rate for the given remaining term to maturity and the average interest level of the loan portfolio, until the point the simulated par interest rate drops below the coupon value, when prepayment occurs.

d) *Determining the interest income effect*, assuming refinancing at the current interest rate, i.e. at the par interest rate, calculating the difference between the original cash flow and the altered cash flow, which is calculated based on the new interest rate.

Stochastic interest rate dynamics

The Cox-Ingersoll-Ross model is a continuous, one factor stochastic interest rate model, first proposed by Cox et al. (1985). According to the model, the dynamics of the instantaneous interest rate (r) in the risk-neutral world is given by the following stochastic differential equation

$$dr = a(b-r)dt + \sigma\sqrt{r}dW \qquad (1)$$

where a, b and σ are constant, dW is normally distributed random variable with zero mean and dt variance. Mean

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

reversion prevails in the model, and volatility is proportional to $\sqrt{r}$. This means that if the short rate increases, then so does its deviation as well. Parameter *b* is the long run equilibrium level of the interest rate, and the *a* parameter represents the speed of the mean reversion. Denoting with *t** the length of time, until the half of difference of the actual and equilibrium value disappears:

$$at^* = ln2 \tag{2}$$

The stylized mortgage portfolio of the bank

For the sake of simplicity, the loan portfolio comprises five elements, each representing a sub-portfolio. These sub-portfolios vary in average interest rate levels and remaining term to maturity, and we summarise their characteristics in Table 1.

Table 1: Composition of the examined loan portfolio

Sub-portfolios	**1.**	**2.**	**3.**	**4.**	**5.**
Weight in the portfolio	20%	20%	20%	20%	20%
Avg. interest rate level	4%	5%	6%	7%	8%
Avg. remaining term to maturity (years)	10	5	6	7	4

The aim of the portfolio settings is to obtain diversity in both interest levels and maturities, similarly to the composition of real-life portfolios.

The cited paper compares two different parameter settings for the CIR model (see Table 2) and consequently two scenarios in terms of the yield curve. Either of them results a declining yield curve, while the other an ascending one.

Table 2: Parameters of the CIR model

	Declining YC	*Ascending YC*
r_0	**6%**	**5%**
a	0.5	0.5
b	**4%**	**7%**
σ	5%	5%

One new portfolio was added to the original calculations: we calculated the results with one item considering the average interest rate level and time to maturity of the original sub-portfolios, if we had substituted the 5 sub-portfolios with one item (results are called "one CF" in Table 3).

The interest levels of the examined sub-portfolios and the prevailing interest rate environment, as well as assumptions of changes in the latter, have a significant influence on the results. As a consequence of the assumed effect of declining interest rates, the prepayment option had a significant effect in the case of sub-portfolio 5, while around 5% of the interest income on the portfolio as a whole was endangered. With these sub-portfolios the effects were concentrated in the first 12 months, so that the interest income effect was substantial within the first year.

The effect within one year appears much more forcefully. The results obtained within one year are a potential maximum, since we assumed an optimal decision-making mechanism, while not reckoning with prepayment and transaction charges. Accordingly, assuming a declining interest rate environment for the hypothetical portfolio, at a 95% confidence level 6% of the planned one-year interest income is potentially endangered. The interest income effect is considerably smaller than this, since the declining interest rates are also apparent in declining funding costs, so that the net effect must be considerably more favourable than the theoretical maximum determined for interest income.

Table 3: Statistics of interest income effect – declining yield curve

Sub-portfolios	**1.**	**2.**	**3.**	**4.**	**5.**	**Overall effect**	**One CF**
Coupon rate	*4%*	*5%*	*6%*	*7%*	*8%*		*6%*
Remaining term to maturity	*10*	*5*	*6*	*7*	*4*		*6.4*
Average	-0.28%	-4.27%	-5.93%	-6.54%	-12.37%	-5.03%	-5.65%
Volatility	0.62%	0.46%	0.28%	0.18%	0.41%	0.16%	0.23%
95% confidence level	-1.68%	-5.01%	-6.38%	-6.81%	-13.02%	-5.35%	-6.03%
99% confidence level	-1.73%	-5.34%	-6.58%	-6.95%	-13.35%	-5.44%	-6.18%

In case of an increasing interest rate trajectory, the potentially endangered interest income is also considerable, meaning that the effect remains significant even when assumptions of the interest rate environment theoretically do not favour prepayment. The scale and nature of the effect are fundamentally influenced by the composition of the examined loan portfolio, since the effect appears in the case of sub-portfolios with a high coupon rate, where – as interest rates start from a low level compared to the coupon rate, and thus rising interest rates can be assumed – it makes sense to prepay. Naturally the result thus obtained can also be regarded as a potential maximum.

The Income-based approach vs Capital value-based approach

In the original paper the cash flow effect was also determined both without a discount and based on discounted cash flow. The cash flow effect is useful for examining the impact of the income-based approach, the goal of which is to estimate the interest income effect. The goal of the discounted cash flow effect is to estimate the change in asset value, and to calculate the impact of the economic capital-based approach accordingly. The impact on capital value was lower than on the cash flows only.

Inclusion of early prepayment fee

In the original paper the calculations were performed also with the incorporation of early repayment fee. In respect of the cost 2% fixed fee was considered to be payable in the case of early prepayment. The cost element affects cash flows through the refinancing incentive. Refinancing has taken place in the model if the par rate pertaining to the given residual maturity and the amount of the annualised value of the early prepayment fee distributed for the residual maturity were even lower than the coupon. In certain cases the incorporation of the early prepayment fee diverts the refinancing decision made merely on the basis of the par rate level since refinancing is not worth any more if the fee is taken into account.

The implementation of the early repayment fee further deteriorates the interest income impact. The reason for this is that due to the fee, early prepayment takes place fewer times but when the loan is refinanced according to the model, on average it takes place at an interest rate that is lower than that in the case when there was no early repayment fee in the model.

SUMMARY OF ORIGINAL CALCULATION RESULTS

The model examines the effect of the optimal prepayment option on the cash flow of bank portfolios and the value of economic capital. Based on the model's results, it can clearly be stated that the prepayment option can have a significant impact on both short-term, one-year total interest income and, via changing cash flows, on the bank portfolio's discounted value and through this the value of economic capital.

The results are largely influenced by the portfolio's interest rate structure (coupon rates) and how it relates to changes occurring in the interest rate environment (declining/ascending yield curve). Considering extending the original model with new yield curve calculation method and changing portfolio composition we expect that portfolio composition change shall have more adequate impact on the original results.

Based on the results of the model, it can be asserted that the interest income effect, depending on assumptions made about the interest rate environment, can be very considerable with respect to expected interest income both in the short term and throughout the loan duration. Change in capital value we attribute in the model to the result of change in the present value of cash flows. This approach also allows long-term effects to be quantified, since it determines a theoretical bond price, as well as any change occurring therein. Methodologically, this approach fits into the logic of determining the capital requirement, on which long-term capital management decisions can be based.

EXTENSION 1

In this section we quantify the effect of a possible *model specification risk*. Our question is the following. What happens if we can observe the yield curve, but we do not know the actual dynamics of the interest rate? How serious is the model specification error if we use the CIR model but the 'reality' is driven by a Hull-White (HW) interest rate model with constant volatility? This model enables us to fit an exactly matching yield curve to the one, which has been obtained by the CIR model. In a specific case of the HW model, namely the Vasicek model we can achieve a close fit to the CIR yield curve just changing the 3 parameters of the model: a, b and σ.

According to the Vasicek (1977) model, the dynamics of the instantaneous interest rate (r) in the risk-neutral world is given by the following stochastic differential equation

$$dr = a(b - r)dt + \sigma dW \tag{3}$$

where a, b and σ are constant, dW is normally distributed random variable with zero mean and dt variance. Mean reversion prevails in this model as well, but now the volatility is constant.

Since the formulas of the yield curves (both spot and forward) are known in both interest rate models, we can use numerical approximation to find various parameter settings that result very similar yield curves in the two models. (Of course, there are numerous 'matching' pairs of parameter settings.) We picked the declining yield curve of the original paper (see Table 2) and searched for corresponding Vasicek parameters. Four possible results are summarised in Table 4.

Table 4: Parameter settings resulting in similar yield curves

	CIR	*Vas1*	*Vas2*	*Vas3*	*Vas4*
r_0	**6%**	**6%**	**6%**	**6%**	**6%**
a	0.5000	0.5147	1.0356	0.5000	0.500
b	**4.00%**	**4.11%**	**4.00%**	**4.47%**	**4.00%**
σ	5.00%	2.66%	5.00%	5.00%	1.01%

Figure 1 shows the difference in the forward yield curves derived from the different models summarised in Table 4. On the same figure, we plotted also the forward yield curve arising from Vasicek model if we use exactly the same parameters as in the CIR model. In this case the resulting two yield curves are different on the long end (see the curves VasCIR and CIR on Figure 1).

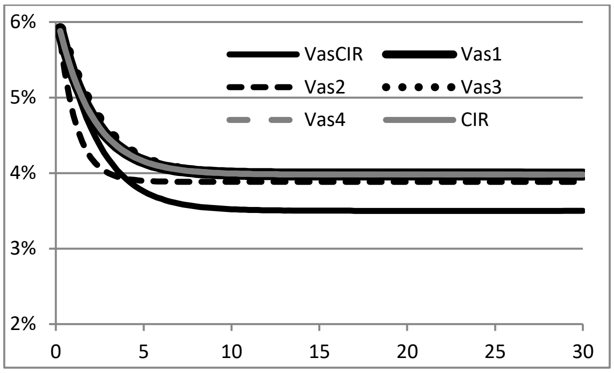


Figure 1: Forward yield curves in different parameter settings

We can achieve a nearly perfect fit to the forward yield curve from CIR by alternating the parameters in the Vasicek. The greatest improvement can be achieved by changing the sigma parameter. Table 5 shows the relative errors related to the different Vasicek parameter settings compared to CIR model.

Table 5: Relative errors related to different Vasicek parameter settings

	Relative error
VasCIR	11,13%
Vas1	5,04%
Vas2	0,87%
Vas3	0,39%
Vas4	0,03%

Doing so, the yield curve is the same, but the short term interest rate dynamics are quite different: unlike in the CIR model, the effect of the random component in the level of the interest rate is now independent from the interest rate itself.

Now we recalculate the interest income effect with our new Vasicek-parameters and compare it to the original results. The results are summarised in Table 6.

Table 6: Results on overall effect in different parameter settings

	CIR	*Vas1*	*Vas2*	*Vas3*	*Vas4*
Avr.	-5,03%	-5,00%	-5,61%	-4,73%	-4,99%
Vol.	0,16%	0,29%	0,28%	0,46%	0,13%
95%	-5,35%	-5,50%	-6,04%	-5,47%	-5,27%
99%	-5,44%	-5,67%	-6,23%	-5,75%	-5,40%

The impact of changing the model with comparable parameters with lowest relative error (*Vasicek1*) is twofold:

1. the average of the interest income effect is somewhat lower compared to the original results;
2. while its volatility is higher, that is true also for the values of the percentiles.

The other calculations compared the results with the original CIR results ceteris paribus:

1. *a*: with higher mean reversion the interest income effect is also higher, while the volatility is also higher;
2. *b*: with higher equilibrium interest rate level the average of the interest income effect is lower, while the percentiles are higher;
3. *volatility*: with lower volatility the effect is almost the same as in case of the original model.

Considering *Vasicek2* results, higher mean reversion means that prepayment happens more often, this is the reason for higher interest income effect. In case of

Vasicek3 the equilibrium interest rate level was increased, which means that the new interest rate at which prepayment happens is also higher, so the final result is also more favourable. In this model the volatility of the interest income effect was the highest from all the cases, also causing higher percentile values at the same time. The *Vasicek4* yield curve was approximating the original yield curve the most, with changing the σ parameter. In this case the result was the closest to the original results, just somewhat lower because of the lower σ.

Figure 2 shows the distribution of results of different parameter settings. As stated above the distribution of the results also show the difference in the general statistics of the results. Since *Vasicek4* yield curve was approximating the original yield curve the most the distribution of the results is almost the same as in case of CIR model.

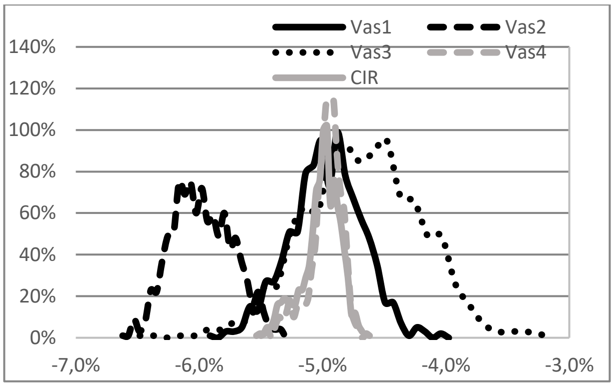


Figure 2: Distribution of results of different parameter settings (CIR and Vasicek)

EXTENSION 2

In this section we quantify the effect of *changes in the bank portfolio composition*. We examine portfolios that have the same *average* interest rate levels and the same *average* maturities, but the standard deviation of these parameters are different.

We consider 4 different compositions:

1. both interest rates and maturities are more diversified compared to original composition;
2. both interest rates and maturities are less diversified compared to original composition;
3. interest rates are unchanged, but maturities are more diversified;
4. maturities are unchanged, but interest rates are more diversified.

The compositions used for the calculations are summarised in Table 7.

Table 7: Composition of the examined loan portfolios

Sub-portfolios	**1.**	**2.**	**3.**	**4.**	**5.**
Weight in the portfolio	20%	20%	20%	20%	20%
Original					
Interest rate level	4%	5%	6%	7%	8%
Maturity (years)	10	5	6	7	4
Case1					
Interest rate level	3%	4%	6%	8%	9%
Maturity (years)	12	4	7	6	3
Case2					
Interest rate level	5%	5.5%	6%	6.5%	7%
Maturity (years)	9.5	5.5	6	6.5	4.5
Case3					
Interest rate level	4%	5%	6%	7%	8%
Maturity (years)	12	4	7	6	3
Case4					
Interest rate level	3%	4%	6%	8%	9%
Maturity (years)	10	5	6	7	4

As a result we can see that changing the interest rate level and the maturities have different impact on the interest income. The results show that if we consider a portfolio composition where interest rates can vary on a wide scale ceteris paribus (*Case4*) it causes higher possible interest rate risk. While changing the maturities in the same direction, i.e. considering a portfolio composition where maturities can vary on a wide scale ceteris paribus (*Case3*), the impact is opposite, it causes lower interest rate risk.

In that case when both interest rates and maturities were put on a less wide scale (*Case2*) the final result actually hasn't changed compared to the original one.

When analysing the results it must be stated that all the above results come from a hypothetical simulation, the portfolio compositions should be varied on a more wide scale to reach a more comprehensive result that could be comparable with real banking portfolios. The results are summarised in Table 8.

Table 8: Statistics of interest income effect of different compositions

Sub-portfolios	Original	Case1	Case2	Case3	Case4
Average	-5.03%	-6.67%	-5.15%	-4.92%	-5.47%
Volatility	0.16%	0.19%	0.10%	0.17%	0.15%
95% confidence level	-5.35%	-7.08%	-5.32%	-5.25%	-5.78%
99% confidence level	-5.44%	-7.22%	-5.38%	-5.35%	-5.88%

CONCLUSION

We analysed the impact of changing model specification and banking portfolio composition of a theoretical model aiming at analysing the effect of prepayment option on a stylized mortgage portfolio. As a result we can state that it makes sense to alternate the yield curve models used for modelling interest rate risk in the banking book. We analysed the results with CIR and a special HW model with constant equilibrium interest rate (Vasicek). In case of Vasicek we used parameter settings causing nearly perfectly fitting yield curve with the originally used CIR model. Even with almost perfectly fitting yield curves we received different results from the original CIR model. We changed parameters *a*, *b* and σ. We received the best fit with changing sigma parameter. In this case interest income effect was very close to the original results, but somewhat lower because of the lower sigma.

We also analysed the impact of different banking portfolio compositions. We diversified the portfolio changing the interest rates and maturities of the sub-portfolios, but considering the same average interest rate level and maturity of the original portfolio composition. An interesting outcome was that if we diversify the portfolio along the interest rates we received higher interest income effect, while diversifying the portfolio along the maturities resulted in opposite effect. It must be stated that we used a simple portfolio composition, more robust results can be achieved with analysing more comprehensive portfolio composition closer to real banking portfolios.

REFERENCES

Cox, J., Ingersoll, E., and Ross, A. 1985. “A Theory of the Term-Structure of Interest Rates”. *Econometrica*, No. 53, 385-407.

Kalfmann, P. 2016. "When is prepayment worthwhile?" *Economy and Finance*, Vol. 3, Issue 2, 129-158.

Vasicek, O. 1977. “An equilibrium characterization of the term structure”. *Journal of Financial Economics*, No. 5, 177-188.

AUTHOR BIOGRAPHIES

Petra KALFMANN, PhD is adjunct professor at the Department of Finance at Corvinus University of Budapest. Previously she worked at Erste Bank Hungary Zrt. as managing director being responsible for micro segment, CRM and campaign management; at Deloitte Zrt. in management consulting and International Training Center for Bankers as consultant and trainer. Her main expertise areas are risk management, Basel 2 implementation, business strategy preparation and implementation. Her e-mail address is: kalfmannp@gmail.com

János SZÁZ, CSc is full professor at the Department of Finance at Corvinus University of Budapest. He was the first academic director of the International Training Center for Bankers in Budapest and he has been its president for many years and until recently. Formerly he was the dean of the Faculty of Economics at Corvinus University of Budapest and President of the Budapest Stock Exchange. Currently he is president of EFFAS Hungary. His main field of research is financing corporate growth when interest rates are stochastic. His e-mail address is: janos.szaz@uni-corvinus.hu

Ágnes VIDOVICS-DANCS, PhD, CIIA is adjunct professor at the Department of Finance at Corvinus University of Budapest. Her main research areas are government debt management in general and especially sovereign crises and defaults. She worked as a junior risk manager in the Hungarian Government Debt Management Agency in 2005-2006. Currently she is chief risk manager of a Hungarian asset management company. Her e-mail address is: agnes.dancs@uni-corvinus.hu

EXPERIMENTS ON RISK PERCEPTION AND INVESTMENT DECISIONS OF ECONOMIC ACTORS

Nora Felfoldi-Szucs
Department Finance
Pallasz Athene University, Corvinus University of Budapest
Izsaki ut 10, Kecskemet 6000
Hungary
E-mail: szucs.nora@gamf.kefo.hu

Peter Juhasz
Department Finance
Corvinus University of Budapest
Fovam ter 8, Budapest 1093, Hungary
E-mail: peter.juhasz@uni-corvinus.hu

KEYWORDS
risk measures, experimental economics, risk aversion

ABSTRACT

In a simple simulated experiment we compare the risk perception and risk taking of participants to the concept of coherent risk measures. Using a sample of 50 participants the aim of our preliminary research is to test the defined experimental environment and define the further directions of its development. Using simulated financial positive homogeneity, subadditivity and monotonicity are perceived by the participants. Translation invariance is applied only by half of the sample. Capital allocation between risky and risk free asset highly correlates with price changes in the most risk averse group of participants. The most risk taking part of the sample followed a strategy to reallocate their gains from risky asset to the risk free one and reorganized their portfolio less frequently but by larger amounts. The behavior of medium risk taker/risk averse participants has to be tested in a more detailed research.

INTRODUCTION

Risk aversion of economic actors is a basic concept both in economics and finance when describing decision making under uncertain circumstances. Risk differs from uncertainty in a way that in risky situation all the possible future outcomes and also the related probability distribution is known. Risk is commonly seen as the deviation from the expected value of the outcome. Though the risk treats the case when the realization differs from the expected value.

There exist a widespread literature of risk measures. The well-known modern portfolio theory (MPT) of Markowitz (1952) applies a mean-variance analyses where the variance of returns stands as a risk measure. The beta in the Capital Asset Pricing Model (CAPM) (Sharpe 1964) is also a well-known risk measure. It measures only the non-diversifiable part of the total risk, the systematic risk thus it does not includes the whole variance of expected returns. Beta shows how the expected excess return of assets are related to the market risk premium trough the non-diversifiable risk of the asset compared to the market risk. Several decades later the Value-at-Risk concept is used among practitioners and regulators to measure and manage market risk of portfolios. (Jorion 2007) But in case of normally distributed returns also VaR is closely related to variance. Artzner et al (1999) define the properties of measures which are appropriate from a theoretical point of view to measure risk. Coherent risk measures are ρ functions which satisfy the following four characteristics:

Translation invariance: adding an amount to our initial X position and investing it into risk free investment reduces the risk measure by *a*:

$$\rho(X+a)=\rho(X)-a \qquad (1)$$

Subadditivity: there are X and Y investment opportunities, and ρ function measures their risk the following way:

$$\rho(X+Y)\leq\rho(X)+\rho(Y) \qquad (2)$$

Positive homogeneity: for all λ≥0 and X initial investment:

$$\rho(\lambda X)=\lambda\,\rho(X) \qquad (3)$$

Monotonicity: for all X≤Y initial investments:

$$\rho(x)\leq\rho(Y) \qquad (4)$$

(Artzner et al. 1999)

Risk aversion of investors is not only a basic element in the above cited models, but also in many of standard theories. (i. e. Bernoulli 1738; Pratt1964; Arrow 1965 as also Holt and Laury (2002) cite.) This can be derived from the expected utility theory of von Neumann and Morgenstern (1953) where assuming a concave utility function over wealth leads to risk aversion of consumers. Although risk aversion is a corner stone of many models, there is no evident experimental method available in the literature how it should be tested or modeled.

The results of Kahnemnann and Tversky (1973, 1984) provide an unevadable challenge to the rationality assumption of economics. According to them the

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

decision making of humans in risky situations is based on shortcuts instead of the classical rationality defined by von Neumann and Morgenstern. Thus the maximization of expected utility does not hold. From their work evolved the behavioral finance. This field provides already plenty of empirical studies on experiments and methods to test the risk aversion of actors or more generally to explorer the perception of risk and the attitude towards risk. Charness, Gneezy and Imas (2013) provide a summary on experimental methods on risk preferences. They describe several measurement tools like elicitation methods and multiple price list method. But most of their analytical tools are static. Eckel and Grossman (2008) tested the risk aversion differences of men and women. They defined three categories of methods: abstract gambling, contextual environment experiments and field studies. Simulated environments belong to their second category. Cohn et al (2015) published one of the most recent study in the topic, they focused on professionals' countercyclical risk aversion. They used a questionnaire with simulated asset prices.

Developing a measurement tool which belongs to the group of contextual environment experiments, we will focus on a new question in this field. The aim of our paper is to compare the professional risk concept to the perceived risk by non-professional participants.

In our paper we compare the risk perception of economic actors to the theoretical construction of coherent risk measures. We test whether the four assumptions of coherency is perceived the same way by the participants of an experiment when evaluating simulated financial data and comparing riskiness of different datasets. Simulating a single risky price process we also collect some evidence on capital allocation decisions of participants. The results can be the first step toward our further research where risk perception and decisions on risky investment are tested in a more complex experiment.

The paper continues as follows. First we describe the applied methodology for the simulation, then we provide details on the experiments and dataset. After the results on coherency axioms we summarize the capital allocation strategy of participants over 10 periods. Finally we conclude.

THE APPLIED METHODOLOGY

In our paper we test the perception of financial risk using a simulated financial data. We designed a questionnaire which was an Excel file where the Structured Monte Carlo simulation (SMC) was run by Visual Basic (VB) codes. We chose the Geometric Browninan Motion (GBM) to describe the process of the financial instruments representing different level of financial risk. The commonly used GBM is a stochastic process which is continuous over time. It assumes that the changes of a S_t stochastic process (i.e. the value of a stock) are characterized by the following equation:

$$dS_t=\mu S_t dt + \sigma S_t dW_t \qquad (5)$$

The W_t represents a stochastic process, the Wiener process and incorporates risk to the model. The parameters of the μ drift and the σ standard deviation are constant. The index t is standing for a given point of time t, while dt is time horizon of the price changes. (Hull 2009)

To analyse the results we use descriptive statistics.

THE EXPERIMENTS AND THE DATASET

Our research project is a pilot study to test the developed SMC and experimental environment. Thus the size of the sample refers to this goal: 50 participants were part of the experiment. Based on the collected initial results the setting and the framework of the experiments will be improved in a later research project.

Instead of real investment situations the experiments provided the participants simulated financial data where they had to make financial decisions. There were no monetary incentives, participants did not receive any monetary payoff depending on their profit/loss attained in the simulation.

The sample is a group of students of a Hungarian College in their first and second year of studies. The participation was voluntary and anonym. Who decided to participate at the experiment they run the SMC-file on their own computer. A short description of rules and the research was attached.

Using the randomly simulated data all the participants met one realization of the simulated distributions. In the first six answers, participants had to decide which one of the simulated price process is riskier. In all the cases the prices were illustrated on a chart, so participants had the possibility visually differentiate between the time series.

The participants had to compare the riskiness of the following assets where the time horizon was one year and price changes occurred on a weekly bases (dt=1/52):

Question (1): A_0=100, μ_A=10%, σ_A=15%, B_t=4A_t

Question (2): A_0=100, μ_A=10%, σ_A=20%, B_0=100, B_t=10+A_t

Question (3): A_0=100, μ_A=10%, σ_A=20%, B_0=100, μ_B=5%, σ_B=20%, C_0=100, C_t= 0.5A_t +0.5B_t

Question (4): A_0=80, μ_A=10%, σ_A=15%, B_0=100, μ_B=12%, σ_B=25%,

Question (5): A_0=80, μ_A=10%, σ_A=15%, B_0=60, μ_B=12%, σ_B=25%

Question (6): A_0=100, μ_A=10%, σ_A=15%, B_0=100, μ_B=12%, σ_B=25%

In the last question all the participants faced an investment decision in a SMC. They had to divide their initial capital to risk free and to risky investment assets. Then the VB code simulated the prices of the chosen asset for the first year of the investment where price changes occurred on a monthly basis. After the first period the investor had the opportunity to restructure the portfolio. In these question we simulated investment decisions over a horizon of 10 years where the investor could reallocate the portfolio at the beginning of each year. There were only one risky asset available thus the question represents the capital allocation problem between risky and risk free sub portfolio instead of security selection. (Thus we followed the logic of Capital Asset Pricing Model (CAPM) (Sharpe 1964) where security selection starts only after the capital allocation was made.)

Question (7): A_0=100, μ_A=10%, σ_A=20%,

The dataset contained all the individually simulated price changes and the investment decisions made by the participants. We also collected the self-defined risk aversion (on a range of 1 to 4) of the participants.

In this preliminary research we focused on the applicability of the simulated experiments. We apply only descriptive statistics to analyse the results because the sample size does not allow more sophisticated methods.

HOW PARTICIPANTS PERCIEVE COHERENCY AXIOMS

In the first part of the experiment we confront the risk perception of participants and the axioms of coherent risk measures. We tested whether the heuristical risk notion of non-professionals is similar to that of the well-designed risk measures if professionals. Table 1 contains the summary of answers in the first six questions (Q1-Q6) representing the axioms of coherency. (The first row of Table 1 represents that 12% of the 50 participants perceived simulated A series of financial data riskier than the B one. But 88% of the 50 participants find that simulated B series of financial data incorporates higher risk than A one.)

Table 1: Answers on Q1-Q6. Testing of coherency. Source: Own calculation

	A	B	C	sum
Q1	10%	90%		100%
Q2	52%	48%		100%
Q3	88%	8%	4%	100%
Q4	12%	88%		100%
Q5	30%	70%		100%
Q6	28%	72%		100%

Question (1) gave the most homogene result. Positive homogeneity was tested here according to $\rho(\lambda X)= \lambda \rho(X)$ axiom the following way: A_0=100, μ_A=10%, σ_A=15%, B_t=4A_t. 90% of participants knew that the higher the exposure the higher the risk in a certain investment.

Translation invariance was represented by Question (2). For the assumption of $\rho(X+a)= (\rho(X)-a)$, we defined A_0=100, μ_A=10%, σ_A=20%, B_0=100, B_t=10+A_t. As the results show, approximately half of the sample realized that process contains a risk free components thus B always dominates A. However the stochastic component in both processes is the same Wiener-process, so the perfect linear correlation may explain why participants of the experiment was unsure how to compare riskiness of A and B.

Subadditivity is tested by Question (3), where according to the subadditivity axiom C should be less risky, then A plus B. Only 4% of participants perceived the diversified C portfolio riskier than the other ones. Most of the answers (88%) was correct evaluating A as the most risky process. So subadditivity was perceived by most of the participants.

Monotonicity has been tested a certain way already in Question (1). An appropriate way for testing could be in a later research also a process setting as follows: e.g. A_0=80, μ_A=10%, σ_A=15%, B_0=100, μ_B=10%, σ_B=15%, where the generated random processes of A and B would not be perfectly correlated as in Question (4). We assume that this question would have been resulted in a slightly lower percent of right answers. (Compared to 90%.)

On the above presented preliminary results the risk perception of participants in our experiments shows similarities to the risk idea of researchers. Three from the four axioms of coherency, namely positive homogeneity, monotonicity and subadditivity are perceived also in the experiments. Translation invariance shows a less evident result. In further research it could be interesting to add a larger risk free part to the risky element of process B_t.

PERCIEVED RISK AND THE LEVEL OF RISKY PROCESS

The remaining questions (Question 4-6) analyze how the same dynamic with different starting value is perceived. A and B processes are defined by their different A_0 and B_0 initial value while their parameters are the same. (Q4-5: μ_A=10%, σ_A=15%, μ_B=12%, σ_B=25%)

The results somehow interfere with monotonicity or positive homogeneity. The fact, that in most of the scenarios process B in Question (4) dominates process A contributed to the perceived risk level. (Remember, Question (1) for positive homogeneity or monotonicity gave the most homogene answers.) We can also value our assumption on the cross tables of Questions (4)-5 or Questions (1) and (4). (See Table 2). 31 participants evaluated B riskier independent of B_0 or A_0. Only 17 participants ranked riskiness according the B_0 or A_0, and 13 of them perceived positive homogeneity or monotonicity in Question (1). Thus these 13 participants applied their initial idea on positive homogeneity as well in other questions. In Question (5) (A_0>B_0) and Question (6) (A_0=B_0) where ranking of A_0 and A_0 contradict the dynamic of the processes, answers are less homogene. (See Table 1.) Thus we assume that the higher the absolute values of the process the higher the perceived risk. This assumption could be tested in our further research using a sample of a larger size. (Having only 50 participants, the cells' values in a cross table are too low to calculate measures of independence. So Table 2 is only an illustration of a cross table.)

Table 2: Cross table of Q4-Q5.
Source: Own calculation

	Q5			
		A	B	Sum
	A	2	4	6
	B	13	31	44
Q4	Sum	15	35	50

Table 3: Cross table of Q1-Q4.
Source: Own calculation

	Q4			
		A	B	Sum
	A	0	6	6
	B	5	39	44
Q1	Sum	5	45	50

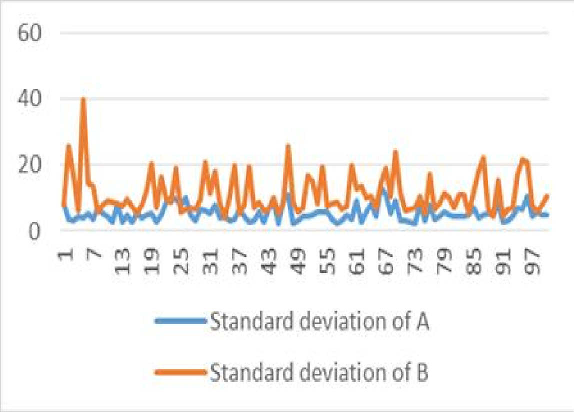
Standard deviation or variance which do not belong to coherent risk measures but are used to capture risk in several models, we can say that it has been an industry standard for decides. (I.e. MPT, CAPM – Capital Asset Pricing Modell) It is interesting that these two show the same pitfalls than the above describes answers. We run 1000 scenarios for processes A_t and B_t, and calculated the standard deviation of the simulated series, on all the processes of Q1-Q6. Table 4 and Figure 1-3. show that the simulation corresponds to the perceived risk of participants.

Table 4: Comparison of standard deviation of A_t and B_t for Q4-Q6.
Source: own simulation

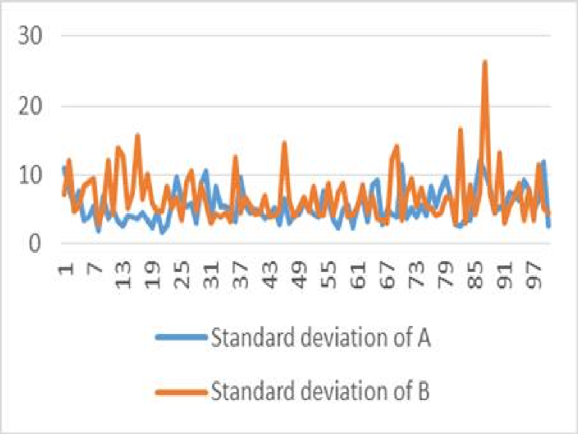
	σ_A	σ_B		% of B answers
Q4	126*	874		88%
Q5	408	592		70%
Q6	227	773		72%

*Number of cases from 1000 simulations where $\sigma_A > \sigma_B$ for Q4. Other cells' values are explained similarly.

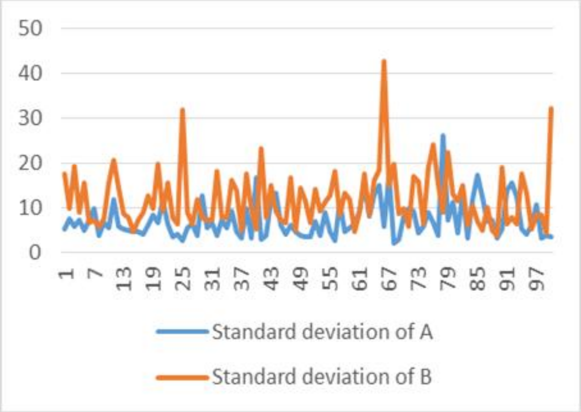
Figures 1: Standard deviations of A and B over 100 realizations in Q4
Source: own simulation



Figures 2: Standard deviations of A and B over 100 realizations in Q5
Source: own simulation



Figures 3: Standard deviations of A and B over 100 realizations in Q6
Source: own simulation

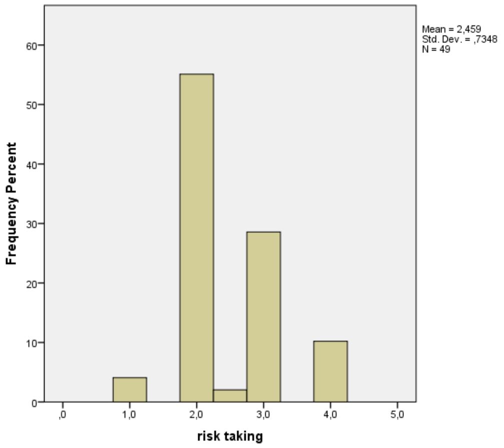


INVESTMENT DECISIONS IN THE CASE OF ONE SIMULATED RISKY ASSET

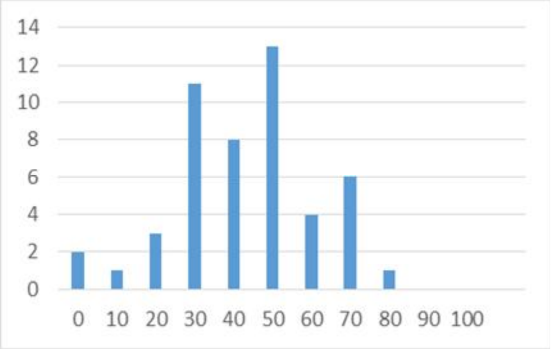
In Question 7 participants faced a 10 years long investment opportunity where they could allocate their initial amount of 100 capital to risky and risk free investment. In Figures 4 there is the histogram of the risk taking ability they defined at the beginning of the experiment (from 1 – totally risk averse to 4 – totally risk taking)

Figures 5 illustrate the first decision on risky sub-portfolio of participants. These two figures suggest that participants are sligthly risk averse at the beginning of the simulation.

Figures 4: Histogram of risk taking level of participants.
Source: SPSS



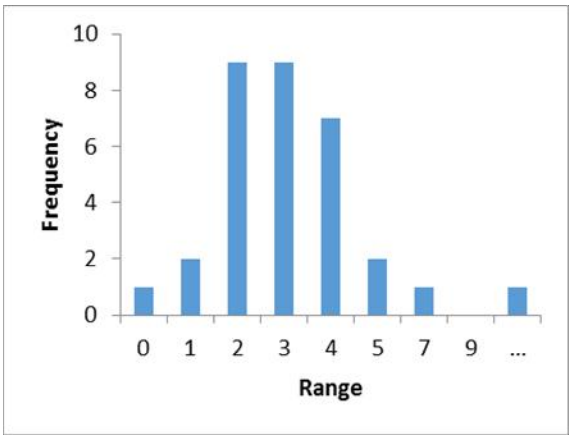
Figures 5: Histogram of allocation decision: The frequency of the allocated amount to risky investment**



** i.e.: 3 of the participants allocated 20 from the initial capital of 100 to the risky asset.

To evaluate the later decisions which are already based on the simulated price process of the chosen risky subportfolio Figures 6 show the distribution of total cumulated price changes (Approxomatly $\Pi_i(1+r_i)$ where i=0; 1…10) of the 49 participants. (One of the 50 participants failed to run the simulation.) The number of cases where a participant dicreased the risky part of the portfolio over the 10 years horizon correlates with the cumulated price changes of the risky investment. (Depending on the definition of cumulated change – absolute values or percentages – the linear correlation coefficient ranges from -0.613 to -0.632.) The correlation atteins 0.85 the highest value in the most risk averse group (risk taking level 1), but it has a weak negative value (-0.28) at risk taking level 4. Less frequent but high volume changes in risky asset allocation explain this correlation. The absolute value of linear correlation decreases over time at riks taking level of 4. These participants allocated the gains from the risky asset to the risk free one at the beginning of the simulation. The middle groups (risk taking level 2-3) show a positive correlation near to 0.5. But there is a periodicity in correlation over time at both of these risk taking levels however the price changes are unique and randomly generated for all participants in all the 10 periods. This autoregressive phenomenon needs further research and larger sample size to be explained.

Figures 6: Frequency of cumulated price changes in risky assets***
Source: Excel

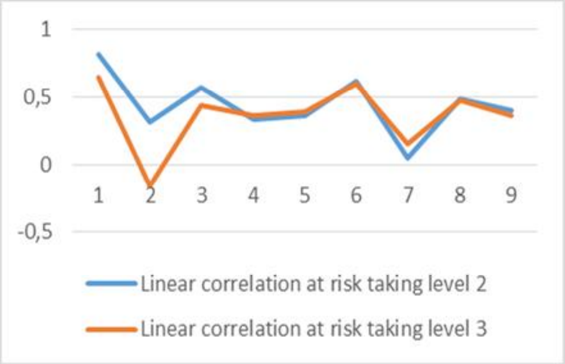


***i.e.: a cumulated change of 2 means that the value of the risky asset doubled over the 10 years.

Table 5: Linear correlation between price changes of risky asset and investment decision among different risk taking levels.
Source: own calculation

Risk taking level	Linear correlation of absolute changes	Linear correlation of % changes
1	0,85	0,60
2	0,49	0,30
3	0,40	0,26
4	-0,28	0,28

Figures 7: Linear correation of absolute value changes over 10 periodes
Source: Excel



CONCLUSIONS

Risk aversion of economic actors derived from the expected utility theory of von Neumann and Morgenstern (1953) is a basic concept both in economics and finance when describing decision making under uncertain circumstances. On the risk measurement for professional tools (i.e. capital allocation, regulatory capital) there exist a widespread literature. The most accepted concept which is the coherency of risk measures was first published by Artzner et al. (1999). But for the risk perception or risk aversion of non-professionals there is no evident experimental method available in the literature how it should be tested or modelled.

In this paper we developed a simulated investment environment and tested this questionnaire on a sample of 50 participants. First we explored the points where more detailed questions should be applied. Second we also present some preliminary results which have to be tested on a larger sample.

In our paper we compared the risk perception of economic actors to the theoretical construction of coherent risk measures. The experimental environment was appropriate to this goal. We need to define new questions only for the axiom of monotonicity. Our results show that the intuition on risk of the participants does not contradict the coherency axioms set by researchers. Positive homogeneity, monotonicity and subadditivity are perceived also in the experiments by non-professional participants. Translation invariance shows a less evident results.

There are further analysis on the relationship between the perceived risk and the level of risky process. This topic is related to coherency axioms like monotonicity or positive homogeneity. We analyzed in three questions how the same dynamic with different starting value is perceived. Thus we assume that the higher the absolute values of the process the higher the perceived risk. The perceived riskiness in these questions is similar to the risk measured by standard deviation or variance which are the most widely used to describe risk. All these three questions helped to deeper understand the results concerning coherency axioms thus we can include them in a further research.

In a simple investment decision game participants are sligthly risk risk averse at the beginning of the simulation. They risk only a smaller part of their capital. Later, the more risk averse the participant the higher the correlation between price changes and capital allocation. The most risk taking part of the sample changed less frequently capital allocation but their changes were larger than those of participants with other risk taking levels. Risk taker participants allocated the gains from the risky asset to the risk free one at the beginning of the simulation which effect in later periods disappeared. The middle groups (risk taking level 2-3) show a positive correlation near to 0.5 between price changes and capital allocation. The autoregressive charahcteristic of the reallocating decisions will be part of our further resaerch.

Thus the experiment had only 50 participants, the results on Questions (1-6) are to be tested in the future at a larger sample. The capital allocation decisions of 49 participants over 10 periods provide sufficent data on reallocations but the autoregression should be tested before treating investment decisions independent of the point of time they were made.

There are several limits to our pilot study not only the size of the sample. The most important one is that the participants are not professionals yet and there were not any financial concecvences of their investment decisions. Thus their behavior could be more close to gampling than to decisions in real investment situations. More robust results can be achieved with more experienced participants.

REFERENCES

Arrow, K. J. 1956: „Aspects of the Theory of Risk Bearing” Helsinki *Academic Bookstores* cited in Holt and Laury 2002.

Artzner, P., Delbaen, F. and Eber, J.M. 1999: „Coherent Measures of Risk”. *Mathematical Finance*, Vol. 9, No. 3. .203-228.

Bernoulli, D. 1738: "Specimen Theoriae Novae de Mensura Sortis." *Comentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, pp. 175-92, translated by L. Sommer 1954 in *Econometrica*, Vol. 22 .23-36. cited in Holt and Laury 2002

Charness, G., Gneezy, U. and Imas, A 2013: „Experimental methods: Eliciting risk preferences” *Journal of Economic Behavior & Organization* Vol. 87. .43-51.

Cohn, A., Engelmann, J., Fehr, E. and Marechal, M. A. 2015: „Evidence for Countercyclical Risk Aversion: An Experiment with Financial Professionals” *American Economic Review* Vol. 105. No, 2. .860-85.

Eckel, C. C. and Grossman, P. J. 2008: „Men, Women and Risk Aversion: Experimental Evidence” (Chapter 113) in *Handbook of Experimental Economics Results*, Vol. 1, .1061–1073.

Holt, C. A. and Laury, S. K. 2002: "Risk Aversion and Incentive Effects" *American Economic Review,*Vol. 92. No. 5. .1644-1655.

Hull, J. C. 2009: „Options, Futures, and other Derivatives” Pearson (9 ed.).

Jorion, P. 2007: „Value-at-Risk”. *McGraw-Hill* ISBN 0-07-146495-6

Khaneman, D. and Tversky, A. 1973: On the psychology of prediction. *Psychological Review* Vol. 80. No.4. .237–25l.

Khaneman, D. and Tversky, A. 1984: Choices, Values, and Frames. *American Psychologist* Vol. 39. No. 4. .341–350.

Markowitz, H.M. 1952: "Portfolio Selection". *The Journal of Finance.* Vol. 7 No. 1. . 77–91.

Neumann, J.von and Morgenstern, O. 1953: „Theory of Games and Economic Behavior”. *Princeton*, NJ. Princeton University Press, 1953

Pratt, J. W. 1964: “Risk Aversion in the Small and in the Large.” *Econometrica* Vol. 32. No. 1-2. . 122-36. cited in Holt and Laury 2002

Sharpe, W. F. 1964: “Capital asset prices: A theory of market equilibrium under conditions of risk.” *Journal of Finance*, Vol. 19 No. 3. .425–442.

ACKNOWLEDGEMENT

We would like to thank the students of Pallasz Athene University, especially Izabella Kutasi, Mate Sarai-Szabo and David Timari for their participation at the experiments.

AUTHOR BIOGRAPHY

NÓRA FELFÖLDI-SZŰCS has attended Corvinus University of Budapest, where she has studied financial investment analysis and risk management and obtained her MA degree in 2006. She started her career at Corvinus University as a lecturer in 2006. She obtained her PhD in 2013. Since 2015 she is the coordinator of the Business Administration BA Program at Pallasz Athene University in Kecskemet. Her e-mail address is: nora.szucs@uni-corvinus.hu .

PÉTER JUHÁSZ is an Associate Professor of the Department of Finance at Corvinus University of Budapest (CUB). He is a CFA charterholder. He holds a PhD from CUB and his research topics include business valuation, financial modelling, and performance analysis. His e-mail address is: peter.juhasz@uni-corvinus.hu.

VOLATILITY SURFACE CALIBRATION IN ILLIQUID MARKET ENVIRONMENT

László Nagy
Mihály Ormos
Department of Finance
Budapest University of Technology and Economics
Magyar tudósok körútja 2, Budapest H-1117, Hungary
E-mail: nagyl@finance.bme.hu, ormos@finance.bme.hu

KEYWORDS
SVI, SSVI, gSVI, stochastic volatility, arbitrage free pricing

ABSTRACT

In this paper, we show the fragility of widely-used Stochastic Volatility Inspired (SVI) methodology. Especially, we highlight the sensitivity of SVI to the fitting penalty function. We compare different weight functions and propose to use a novel methodology, the implied vega weights. Moreover, we unveil the relationship between vega weights and the minimization task of observed and fitted price differences. Besides, we show that implied vega weights can stabilize SVI surfaces in illiquid market conditions.

INTRODUCTION

Vanilla options are traded with finite number of strikes and maturities. Thus, we can observe only some points of the implied volatility surface. It is known that vanilla prices are arbitrage free hence exotic option traders would like to calibrate their prices to vanillas (Dupire 1994). The main difficulty is that calibration methods need the implied volatility surface. To overcome this problem we have to construct an arbitrage free surface from the observed points (Schönbucher 1998, Gatheral 2013). In this paper we provide a robust arbitrage free surface fitting methodology.
Chapters are structured as follows: Section 2. is a brief overview of SVI. In Section 3. we compare the different weight functions and present our implied vega weight L^1 methodology. In Section 4. we summarize the findings.

SVI

After the Black Monday in 1987, traders behavior changed. Implied volatility skew became more pronounced. Risk aversion incorporated in the volatility. Hence, risk transfers between tenors and strikes get more sophisticated.
The changes were in line with human nature, because people have different risk appetite in different tenors. Moreover, extreme high out of money implied volatilities are consequences of risk aversion and fear of the unpredictable.

Besides, our risk neutral risk assessment should be consequent. Otherwise, calendar and butterfly arbitrage opportunities appear;

$$\mathrm{C}(\mathrm{K},\tau_1) < \mathrm{C}(\mathrm{K},\tau_2) \text{ if and only if } \tau_1 < \tau_2 \qquad (1)$$

$$\mathrm{C}(\mathrm{K}_1,\tau) - \frac{K_3-K_1}{\mathrm{K}_3-\mathrm{K}_2}\mathrm{C}(\mathrm{K}_2,\tau) + \frac{\mathrm{K}_2-\mathrm{K}_1}{\mathrm{K}_3-\mathrm{K}_2}\mathrm{C}(\mathrm{K}_3,\tau) > 0 \qquad (2)$$

where $\mathrm{C}(\mathrm{K},\tau)$ represents the price of a European call option with strike K and maturity τ.

Considering the behavior of compound interest the Black-Scholes log-normal model is applicable.

$$\mathrm{dS_t} = \mathrm{rS_t dt} + \sigma \mathrm{S_t dW_t} \qquad (3)$$

Also we have seen the volatility surface is not flat hence this model needs some adjustments. The most straightforward correction leads to the Local Volatility model (Dupire 1994).

$$\mathrm{dS_t} = \mathrm{rS_t dt} + \sigma(\mathrm{S_t},\mathrm{t})\mathrm{S_t dW_t} \qquad (4)$$

However, calculating implied and realized volatilities show that Local Volatility is only an idealized perfect fit, because volatility is stochastic.

$$\mathrm{dS_t} = \mathrm{rS_t dt} + \sigma_t^{Spot} S_t \mathrm{dW_t} \qquad (5)$$
$$\mathrm{d}\sigma_t^{BS} = \mathrm{u(k,t)dt} + \gamma(\mathrm{k,t})\mathrm{dW_t} + \textstyle\sum_{i=1}^{n} \mathrm{v_i(k,t)}\, \mathrm{dW_t^i}$$

where $\mathrm{W}, \mathrm{W}_1, \dots \mathrm{W}_\mathrm{n}$ are independent Brownian motions, k is the log-moneyness and σ^{BS} denotes the Black-Scholes implied volatility.

Schönbucher showed that the spot volatility can not be an arbitrary function of implied volatility, because of the static arbitrage constraints.

$$\sigma^{\mathrm{Spot}} = \frac{-\gamma k}{\sigma^{\mathrm{BS}}(\mathrm{k,T})} \pm \sqrt{\sigma^{BS}(\mathrm{k,T}) + \frac{k^2}{\left(\sigma^{BS}(k,T)\right)^2}\left(\textstyle\sum_{i=1}^{n} \mathrm{v_i^2} - \gamma^2\right)} \qquad (6)$$

Besides the arbitrage constraints, Heston's model sheds more light on implied volatility modeling. Gatheral et

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

al. proposed the so called SVI (Stochastic Volatility Inspired) function to estimate all the implied volatility surface;

$$\sigma^{BS} = \mathrm{a} + \mathrm{b}\left(\rho(\mathrm{k}-\mathrm{m}) + \sqrt{(\mathrm{k}-\mathrm{m})^2 + \sigma^2}\right) \tag{7}$$

where a controls the level, b the slopes of the wings, ρ the counter-clockwise rotation, m the location and σ the at the money curvature of the smile. The SVI model has compelling fitting results, in addition it implies a static arbitrage free volatility surface. The only arguable step in the methodology is the model calibration. Kos et al. (2013) proposed to minimize the square differences between observed and fitted volatility, while Homescu (2011) advised a square difference method. Nevertheless West (2005) applied vega weighted square volatility differences. Zelida system (2009) used total implied variances. An other noticeable approach comes from Gatheral et al. (2013) who minimized squared price differences, but there are further regression based models as well (Romo 2011).

In this article we propose a new absolute price difference based approach to stabilize SVI in illiquid market conditions.

SENSITICITY ANALAYSIS OF SVI

At the money options are more liquid than far out of money options, hence the bid-ask spread widens along the wings. This implies larger price ambiguity of OTM option prices. In order to stabilize the implied volatility surface we have to penalize price ambiguity.

Uniform weights

Highlighting the problem we can apply uniform weights. This approach assumes that all of the information is equally relevant. Thus, we get the usual square distance optimization task (Zelida 2009, Kos 2013);

$$\min_{\sigma^{\mathrm{Fit}}_{\mathrm{K},\tau}\in \mathrm{C}^0} \sum_{\tau\in\mathcal{T}}\sum_{K\in\mathcal{K}}\left(\sigma^{\mathrm{Fit}}_{\mathrm{K},\tau} - \sigma^{\mathrm{BS}}_{\mathrm{K},\tau}\right)^2 \tag{8}$$

where $\mathcal{K}$ and $\mathcal{T}$ represent the sets of the traded strikes and maturities, $\sigma^{\mathrm{BS}}_{\mathrm{K},\tau}$ is the implied and $\sigma^{\mathrm{Fit}}_{\mathrm{K},\tau}$ is the fitted volatility.

1. Note that for fixed maturity $\sigma^{\mathrm{BS}}_{\mathrm{K,T}}$ increases in $|\mathrm{K}-\mathrm{F_T}|$. The volatility bid-ask spread also widens along the wings. This implies that $\sigma^{\mathrm{BS,Ask}}_{\mathrm{K,T}}$ increases faster than $\sigma^{\mathrm{BS,Bid}}_{\mathrm{K,T}}$hence defining the fair value of a deep OTM option from bid and ask price is not straightforward.

2. Furthermore, deep out of money implied volatilities are usually higher than ATM volatilities. Therefore, uniform square penalty overfits the wings and underfits the ATM range.

3. In addition, the set of traded strikes is not stable in time. Therefore, the estimated surface will be unstable in time.

Data truncation

The simplest approach to solve the problem could be just using close ATM prices to fit SVI and then extrapolate along wings. The main drawback of this method is that OTM short dated options contain the market anticipated tail risk information. Truncating the data stabilizes the surface and implies accurate long term fit, but underestimates tail risk hence underprices exotic products.

Square of price differences

The most popular optimization technique is minimizing L^2 distances. The main drawback of this approach is fitting to the mean, instead of the median which implies outlier sensitivity.

Vega weights

In order to deal with the skew and price ambiguity we propose to use a natural Gaussian based weight function. It turned out that truncating the data do not give the appropriate results. Therefore, we have to find a weight function which minimizes L^1 distance, penalizes price ambiguity, but still able to use tail risk information.

$$\min_{\sigma^{\mathrm{Fit}}_{\mathrm{K},\tau}\in \mathrm{C}^0} \sum_{\tau\in\mathcal{T}}\sum_{K\in\mathcal{K}} \mathrm{w}(\mathrm{K},\tau)\mid \sigma^{\mathrm{Fit}}_{\mathrm{K},\tau} - \sigma^{\mathrm{BS}}_{\mathrm{K},\tau}\mid \tag{9}$$

Note that the above optimization problem is still not general enough, because weight is a function of strike and maturity. This incorporates a sticky strike assumption. However, if we add the spot price S_0 as another independent variable to the weight, then we can get more general penalty functions.

$$\min_{\sigma^{\mathrm{Fit}}_{\mathrm{K,S_0},\tau}\in \mathrm{C}^0} \sum_{\tau\in\mathcal{T}}\sum_{K\in\mathcal{K}} \mathrm{w}(\mathrm{K},\mathrm{S_0},\tau)\mid \sigma^{\mathrm{Fit}}_{\mathrm{K,S_0},\tau} - \sigma^{\mathrm{BS}}_{\mathrm{K,S_0},\tau}\mid \tag{10}$$

Our initial problem is to find an implied volatility surface. This means that we would like to penalize observed and fitted volatility differences.

Practitioners need the surface for trading. Hence, they are interested in the dollar amount of the discrepancies between fitted and observed volatilities.

$$\min_{C^{Fit}_{K,S_0,\tau}\in \mathrm{C}^0} \sum_{\tau\in\mathcal{T}}\sum_{K\in\mathcal{K}} \left|C^{Fit}_{K,S_0,\tau} - C^{BS}_{K,S_0,\tau}\right| \tag{11}$$

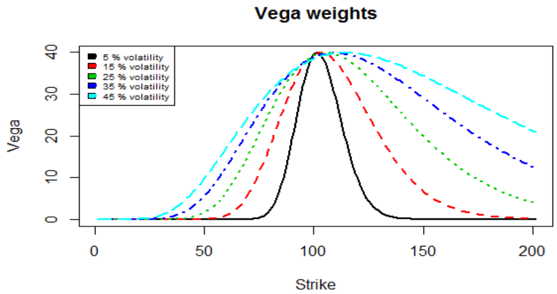
After some calculations in Appendix, we can see that optimizing the price differences is approximately the same as optimizing vega weighted implied volatility differences.

$$\min_{\sigma^{\mathrm{Fit}}_{\mathrm{K,S_0},\tau}\in \mathrm{C}^0} \sum_{\tau\in\mathcal{T}}\sum_{K\in\mathcal{K}} \mathcal{V}^{BS}_{\mathrm{K,S_0},\tau}\mid \sigma^{\mathrm{Fit}}_{\mathrm{K,S_0},\tau} - \sigma^{\mathrm{BS}}_{\mathrm{K,S_0},\tau}\mid \tag{12}$$

Using the definition of $\mathcal{V}^{BS}_{\mathrm{K,S_0,\tau}}$ we get the price difference implied weight function.

$$\mathrm{w(K,S_0,\tau) = S_0 e^{-q\tau}\varphi(d_1)\sqrt{\tau}} \qquad (13)$$

where $\varphi(\mathrm{x})$ represents the standard normal distribution function, d_1 is the standard notation from the Black-Scholes formula and q is the continuous dividends rate.



Figures 1: Vega weights as function of strike, and volatility $\sigma \in [5,15,\dots,45], S_0 = 100,\ \mathrm{K} \in [0,\dots,200]$ and $\mathrm{T} = 1$ year

Supposing that q, r, σ, S_0, τ are fixed and using the definition of d_1 we get that the weight function has a Gaussian shape in log moneyness $K = Fe^k$.

$$w(K) = \mathrm{S_0\ e^{-q\tau}}\ \frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}\left(\frac{\ln\frac{F}{K}+\frac{\sigma^2}{2}\tau}{\sigma\sqrt{\tau}}\right)^2} = \mathcal{O}\left(e^{-k^2}\right) \qquad (14)$$

This implies that we highly penalize fitting discrepancies in the ATM range, while we are lenient with deep OTM fits.
The next step is to fix q, e, S_0, τ and use the first order Taylor approximation of $\sigma(K, S_0, \tau)$ around ATM log moneyness.

$$w\big(\sigma(k)\big) =$$

$$\frac{S_0 e^{-q\tau}}{2\pi} e^{-\frac{1}{2}\left(\frac{-2k+\left(\sigma(0,S_0,\tau)+\Psi(S_0,\tau)k+\mathcal{O}(k^2)\right)^2\tau}{2\left(\sigma(0,S_0,\tau)+\Psi(S_0,\tau)k+\mathcal{O}(k^2)\right)\sqrt{\tau}}\right)^2} \sqrt{\tau} \qquad (15)$$

Hence we get;

$$\mathrm{w\big(\sigma(k)\big)} \approx \mathcal{O}(e^{-k^4}) \qquad (16)$$

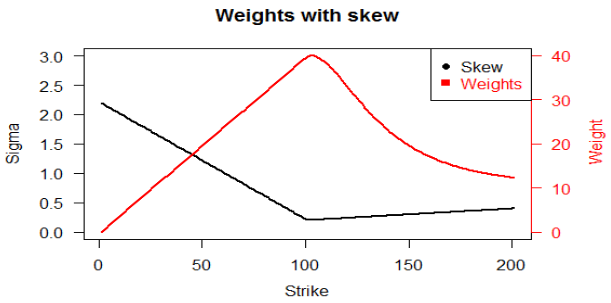
ATM skew is represented by $\Psi(S_0, \tau)$. The asymptotic behavior of $w\big(\sigma(k)\big)$ shows that the vega weighted implied volatility surface would be stable against extreme OTM implied volatilities.
Moreover, it also can be seen that if k is close to zero then for implied volatility skew and smile we get rather flat vega weights.

$$\mathrm{-2k + \big(\sigma(0,S_0,\tau) + \Psi(S_0,\tau)k} + \mathcal{O}(\mathrm{k^2})\big)^2 =$$
$$\mathrm{-2k\big(1 - \sigma(0,S_0,\tau)\Psi(S_0,\tau)\big) + \sigma(0,S_0,\tau)^2} + \mathcal{O}(\mathrm{k^2}) \qquad (17)$$

Dividing with $\sigma(0, S_0, \tau)$ we get:

$$\frac{-2k\big(1-\sigma(0,S_0,\tau)\Psi(S_0,\tau)\big) + \sigma(0,S_0,\tau)^2 + \mathcal{O}(\mathrm{k^2})}{2(\sigma(0,S_0,\tau) + \Psi(S_0,\tau)k + \mathcal{O}(\mathrm{k^2}))}$$

This function is rather constant if k is small. Equation 14. also shows that for bigger $|k|$ values the weight should decrease with approximately $\exp(-k^2)$.



Figures 2: Implied weights as function of strike, parameters: $\mathrm{S_0} = 100, \sigma_{0,100,1} = 0.2, \mathrm{K} \in [0,\dots,200]$, slopes = (2% ,0.2%)

Figure 2. unveils that vega weights take into account wings, but the bigger the $|k|$ the larger the impact of the $\exp(-k^2)$ term which balances the increasing OTM volatility. Hene, vega weights provide a balanced SVI fit.

Empirical test

In order to lend more color to the fragility of fitting methodology we simulated illiquid market environment by picking 5 data points in each slice from SPX 15/09/2015 option data set (Gatheral 2013). To highlight the outlier-sensitivity we stressed the volatility of the last tenor (T=1.75), moneyness k=0.2 point by 10%.

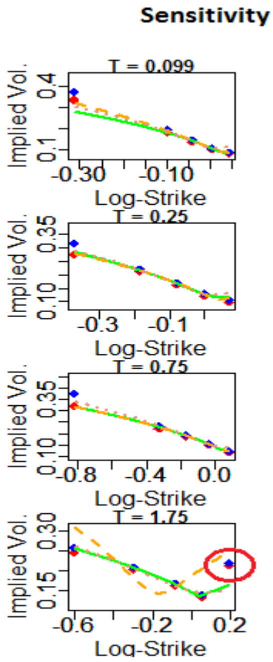


Figure 3: fitted SVI surfaces to filtered SPX 15/09/2015 data, solid lines: price difference based fit (implied vega weights), dashed lines: square price difference based fit, dotted line: square volatility difference based fit

The results show that if we use L^2 optimization techniques then we overpenalize outliers. It also can be seen using absolute difference based optimization makes the SVI fit stable. In illiquid market environment it is crucial because using square difference based fits only one outlier could have a huge impact on the affected slice or even the all surface, thus destabilizing option prices.

CONCLUSION

We showed that the absolute price difference based SVI fitting methodology is able to stabilize the implied volatility surface. Moreover, we shed some lights on the asymptotic behavior of the weights and displayed the connection with vega weights. We also stressed that absolute price difference based optimization do not assume any specific stickiness, hence it can be used in every volatility regime.

ACKNOWLEDGEMENTS

Mihály Ormos acknowledges the support by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. László Nagy acknowledges the support by the Pallas Athéné Domus Scientiae Foundation. This research is partially supported by Pallas Athéné Domus Scientiae Foundation.

REFERENCES

A. Kos; and B. M. Damghani. 2013. “De-arbitraging With a Weak Smile: Application to Skew Risk”. *WILMOTT magazine*

B. Dupire. 1994. “Pricing with a Smile”. *Risk*

C. Homescu. 2011. “Implied volatility surface: construction methodologies and characteristics”. *arXiv:1107.1834v1*

G. West. 2005. “Calibration of the SABR model in illiquid markets”. *Applied Mathematical Finance 12(4):371-385*

J. M. Romo. 2011. “Fitting the Skew with an Analytical Local Volatility Function”. *International Review of Applied Financial Issues and Economics, Vol. 3, pp. 721-736*

J. Gatheral; and A. Jacquier. 2013. “Arbitrage-free SVI volatility surfaces”. *arXiv:1204.0646v4*

L. R. Goldberg; M. Y. Hayes; O. Mahmoud; and Tamas Matrai. 2011. “Vega Risk in RiskManager Model Overview”. *Research insight*

P. J. Schönbucher.1998. “A market model for stochastic implied volatility”

Zeliade Systems. 2009. “Quasi-Explicit Calibration of Gatheral's SVI model”. *Zeliade White Paper*

László Nagy

László Nagy is a PhD student at the Department of Finance, Institute of Business at the School of Economic and Social Sciences, Budapest University of Technology and Economics. His main area of research is financial risk measures and asset pricing. Laszló earned his BSc and MSc in Mathematics with major of financial mathematics at the School of Natural Sciences at Budapest University of Technology and Economics. He is teaching investments and working on his PhD thesis. Before his PhD studies he worked at Morgan Stanley on risk modeling.

Mihály Ormos

Mihály Ormos is a Professor of Finance at the Department of Finance, Institute of Business at the School of Economic and Social Sciences, Budapest University of Technology and Economics. His area of research is financial economics especially asset pricing, risk measures, risk perception and behavioral finance. He serves as one of the contributing editors at Eastern European Economics published by Taylor and Francis. His teaching activities concentrate on financial economics, investments and accounting. Prof. Ormos published his research results in Journal of Banking and Finance, Quantitative Finance, Finance Research Letters, Economic Modelling, Empirica, Eastern European Economics, Baltic Journal of Economics, PLoS One, Acta Oeconomica and Economic Systems amongst others.

APPENDIX

$$\begin{aligned}
&\min_{C^{Fit}_{K,S_0,\tau}\in C^0} \sum_{\tau\in\mathcal{T}} \sum_{K\in\mathcal{K}} |C^{Fit}_{K,S_0,\tau} - C^{BS}_{K,S_0,\tau}| \\
&= \min_{C^{Fit}_{K,S_0,\tau}\in C^0} \sum_{\tau\in\mathcal{T}} \sum_{K\in\mathcal{K}} |S_0 e^{-q\tau}\left(\Phi^{SVI}_{(d_1)} - \Phi^{BS}_{(d_1)}\right) \\
&\qquad - e^{r\tau} K(\Phi^{SVI}_{(d_2)} - \Phi^{BS}_{(d_2)})|
\end{aligned}$$

$$= \min_{C^{Fit}_{K,S_0,\tau} \in \mathrm{C}^0} \sum_{\tau \in \mathcal{T}} \sum_{K \in \mathcal{K}} | \left(S_0 e^{-q\tau} \varphi(d_1^{BS}) \frac{\ln \frac{K}{F_\tau} + \frac{\sigma^{BS}\sigma^{SVI}}{2}}{\sigma^{BS}\sigma^{SVI}\sqrt{\tau}} \right.$$

$$\left. - K e^{-r\tau} \varphi(d_2^{BS}) \frac{\ln \frac{K}{F_\tau} - \frac{\sigma^{BS}\sigma^{SVI}}{2}}{\sigma^{BS}\sigma^{SVI}\sqrt{\tau}} \right) (\sigma^{SVI} - \sigma^{BS})$$

$$+ \mathcal{O}\left(\left(\frac{\sigma^{SVI} - \sigma^{BS}}{\sigma^{BS}\sigma^{SVI}\sqrt{\tau}} \right)^2 \right) |$$

$$= \min_{C^{Fit}_{K,S_0,\tau} \in \mathrm{C}^0} \sum_{\tau \in \mathcal{T}} \sum_{K \in \mathcal{K}} | \left(\mathcal{V}^{BS}_{K,\tau} \frac{\ln \frac{K}{F_\tau} + \frac{\sigma^{BS}\sigma^{SVI}}{2}}{\sigma^{BS}\sigma^{SVI}\sqrt{\tau}} \right.$$

$$\left. - \mathcal{V}^{BS}_{K,\tau} \frac{\ln \frac{K}{F_\tau} - \frac{\sigma^{BS}\sigma^{SVI}}{2}}{\sigma^{BS}\sigma^{SVI}\sqrt{\tau}} \right) (\sigma^{SVI}$$

$$- \sigma^{BS}) + \mathcal{O}\left(\left(\frac{\sigma^{SVI} - \sigma^{BS}}{\sigma^{BS}\sigma^{SVI}\sqrt{\tau}} \right)^2 \right) |$$

$$= \min_{C^{Fit}_{K,S_0,\tau} \in \mathrm{C}^0} \sum_{\tau \in \mathcal{T}} \sum_{K \in \mathcal{K}} | \mathcal{V}^{BS}_{K,\tau} (\sigma^{SVI} - \sigma^{BS})$$

$$+ \mathcal{O}\left(\left(\frac{\sigma^{SVI} - \sigma^{BS}}{\sigma^{BS}\sigma^{SVI}\sqrt{\tau}} \right)^2 \right) |$$

$$\approx \min_{C^{Fit}_{K,S_0,\tau} \in \mathrm{C}^0} \sum_{\tau \in \mathcal{T}} \sum_{K \in \mathcal{K}} | \mathcal{V}^{BS}_{K,\tau} (\sigma^{SVI} - \sigma^{BS}) |$$

Note that $\mathcal{V}\ is\ o(\sqrt{\tau})$, thus options with short expiry are not vega sensitive.

Modelling of provision under new International Financial and Reporting Standard (IFRS 9)

Kádár Csaba
Corvinus University of Budapest
1093, Budapest
E-mail: csaba.kadar@uni-corvinus.com

KEYWORDS
Model, impairment, IFRS9, economic cycle.

ABSTRACT

The impairment recognition in International Financial and Reporting Standard will change significantly in 2018 with IFRS 9. The reason of the update of the current IAS 39 standard is related to the global financial crisis. In this paper I am modelling the possible effect of the new standard to the allowance calculation compared with the previous standard. I set the focus on the two main dimensions of impairment recognition, namely time and amount. These characteristics of the introduced model in IFRS 9 are closely related to the economic cycle. During the analysis, I include the effect of macro environment to the allowance and highlight the expected changes and possible upcoming uncertainty of impairment estimation.

INTRODUCTION

In the last few decades, significant changes appeared in the field of financial regulation. Some of them are related to the increasing complexity and volume of financial deals and exposures, and interdependencies between different sectors and entities. Others are consequences of the last financial crisis, which generated a regulatory dumping all over the world. The sweeping changes of financial infrastructure have remarkable effects to the "real" economy as well. Great parts of the most relevant developments are connected to the reserving capability and reserves of the banks both from prudential - solvency capital - and accounting – allowance - sides. Of course there should be relevant and significant differences between the prudential and accounting regulations (Borio and Tsatsaronis 2005).

The update of International Financial and Reporting Standards (IFRS) are one of the main evolutions at international level. Maybe the biggest impact inside IFRS has the IFRS 9, the new standard related to the classification, measurement and accounting of financial instruments, for financial sector.

SCOPE OF THE PAPER

In my current paper I am examining the effect of impairment model of the upcoming IFRS 9 standard coming into force after 2018. I compare the results with the currently used standard of IAS 39.

The explanation and justification behind the revision of currently used impairment models was that the existing accounting standard recognized the credit loss with delay and less in amount as it is needed. According to these I am examining two hypothesises, one is related to the timing and one is related to the amount of allowance recognition. The first hypothesis is whether the IFRS 9 will recognise the impairment loss earlier and the second hypothesis is that the IFRS 9 will recognise higher impairment amount compared to the IAS 39 one. Existing IAS 39 standard is based on incurred loss model (Tardos 2005; Szabó 2005), which means that only already "incurred" loss could be taken into account in impairment calculation, while according to the new standard there will be an expected loss model (IASB 2014). It means that future losses stemming from or based on expectations on past or current circumstances - with forward looking - should be included as well. This change is intended to cover the timeliness issue between the discrepancies of existing standard. The concerned related to the shortfall in amount is handled with prescription that when there is a significant change after initial recognition of the financial assets than the entity should calculate expected credit loss for the full lifetime of the instrument instead of 12 months' one. This requirement is called as 'staging rules'.

Definition of significant credit risk increase after initial recognition includes qualitative and quantitative criteria as well. Qualitative criteria consist of day past due, work-out, forbearance and early warning indicators, while qualitative criteria are connected to the rating systems as change of rating grades and related probability of default since initial recognition.

In that sense the macroeconomic circumstances are getting an important role in IFRS 9 through the

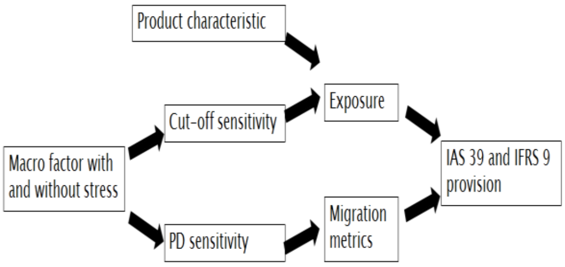
Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

comparison of initial and current credit risk expectations, which could be the driver of impairment as a systematic factor. Because generally there is a big uncertainty in the estimation of macro circumstances, it is worth to analyse the effect of such a macro stress or a shock to the impairment amount.

Currently introduced model is the sub-model of a greater one which examines the IFRS 9 changes with the interaction of Basel III prudential rules especially with countercyclical capital buffer and try to highlight the combined effects of the prudential and accounting rules.

MODELLING PROCESS

In this model I investigate the effects of the upcoming impairment model on a hypotatic portfolio with characteristics - based on reasonable judgement - described in chapter Exogenous variables. The modelling process depicted by the following diagram:



Picture 1: Effects in the model

Modelling process is a deterministic calculation. The outcome of the model is the amount of provision according to IAS 39 and IFRS 9 for sequence of periods with different macroeconomic states. The final provision of a given period is determined by the current and future exposure of the deals in the portfolio multiplied with loss given default and default probabilities coming from migration metrics (see details later in Table 4). Number of deals is given by the cut-off sensitivity – as rejection rule – of the given financial entity to the macro factor. Product characteristics and migration metrics are influenced by the macro factors as well. It is worth to note that the cut-off sensitivity is not a crucial part of the sub-model so that conclusions do not change without it.

Exogenous variables

Before I show the steps of the calculation process I introduce the exogenous variables and simplification used in the model.
Exogenous variables and related simplification are the following:

- Unconditional probability – so where macroeconomic circumstance is still not incorporated - is constant at 10% for all rating grades.
- Unconditional acceptation rate is constant 80%. So rejection rate is 20%.
- Loss given default (LGD): loss recognized after default. Constant value of 10% is used.
- Amortisation of deals exposure (principal balance): Linearly up to the maturity. Default maturity is 5 years. So exposure of the given deal are 100, 80, 60, 40, 20 in the sequence years after draw down.
- Applicant: Number of possible applicant is constant at 100.
- Migration metrics ($M_{0,1}$): during the calculation the following cumulative unconditional migration metrics is used:

Table 1: Migration metrics

t0->t1	Rating A	Rating B	Default
Rating A	0,7	0,9	1
Rating B	0,2	0,9	1
Default	0	0	1

It means that we have only 2 non-defaulted rating grades and 1 defaulted category without recurrence. Cumulative values means that the migration probabilities are cumulated from the first value in each rows. So exclude the effect of different rating grades I set the direct default probability to the same level, which is 0,1 as 1 minus 0,9.

- Staging rules: At initial recognition all the exposure are in rating A. During modelling I will use a simplified staging criteria namely if the exposure migrated to the rating B then the exposure will be in stage 2 and lifetime expected loss needs to be calculated instead of 12 months' one.
- Macro factor: add information about the state of the economy of the given year as a systematic factor. One baseline and one stress scenario are used during modelling with the following standard normally distributed variables:

Table 2: Scenarios

Time period in year	1-9 period	10	11	12	13	14	15-20 period
Baseline scenario	1	1	1	1	1	1	1
Stress scenario	1	-1	-1	-1	0,5	0,5	1

So the economic circumstance in case of baseline scenario is a constant mild expansion - because 1 means that circa 85% of the possible outcomes are worse. After the 9 period there is a shock in the stress scenario where the -1 means that the 85% of the possible outcome are better. After 12 period the state of economic is start to converge to the baseline scenario.

- The cut-off sensitivity (acceptance rate of applicants as a new debtor) set to 10% and migration metrics – so PD – sensitivity set to 10%. During the modelling process the increase of the sensitivity does not change the final results.

Endogenous variables

As described above the major interactions are between PD and macro factor and between acceptation rate and macro factors.
For the dependency of the PD and the acceptation rate Vasicek formula (1) is used:

$$p_A(X) = \Phi\left(\frac{\Phi^{-1}(\bar{p}_A) - \sqrt{\rho_A}X}{\sqrt{1-\rho_A}}\right). \tag{1}$$

Where the X is the macro factor with standard normal distribution and the ρ_A is the correlation between macro and unconditional PD and acceptation rate. $p_A(X)$ refers to the conditional PD and acceptation rate calculated on the same way, but with different correlation factor to the systematic macro factor (Janecsko 2004).
If we have the conditional PD value – defined above -, we can use it to adjust the unconditional migration metrics to get the conditional migration metrics of the given year. The adjustment process of the unconditional migration metric with conditional PD values (2) looks like the following (named as z-shift adjustment):

$$\begin{tikzcd}
\mathrm{M}_{0,1} \arrow[r, "\boxed{\text{Cumulate metrics}}"] & \mathrm{M}_{0,1} \arrow[r, "\boxed{\text{Inverse normal}}"] & \mathrm{M}_{0,1} \arrow[d, "\boxed{\text{Shift}}"] \\
\mathrm{M}_{0,1\,\text{shifted}} & \mathrm{M}_{0,1} \arrow[l, swap, "\boxed{\text{Backward calculation}}"] & \mathrm{M}_{0,1} \arrow[l]
\end{tikzcd} \tag{2}$$

Where $\mathrm{M}_{0,1}$ is the unconditional migration metrics (3) with all the rating grades as showed in Table 1 above:

Table 3: Structure of migration metric

$$\mathrm{M}_{0,1} = \begin{pmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,d} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,d} \\ \vdots & & \cdots & \\ m_{d-1,1} & m_{d-1,2} & \cdots & m_{d-1,d} \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{3}$$

If we have the conditional migration metrics for all the years adjusted with the actual unconditional PD value then we can calculate the state of the exposures in year t (Gruenberger 2012) after initial recognition (4) as:

$$\mathrm{M}_{0,t} = \mathrm{M}_{0,1} \cdot \mathrm{M}_{1,2} \cdot \cdots \cdot \mathrm{M}_{t-1,t} \tag{4}$$

After we get the state of the exposure we get to know that the given exposure is in rating A, rating B or in default. Under IAS 39 in case of non-defaulted rating grades we calculate the impairment value based on the loss of the previous year, because of incurred loss model. Under IFRS 9 if it is in rating A than according to our assumption, we need to calculate impairment for the next 12 months' expected credit loss. If it is in rating B we need to calculate impairment for the whole lifetime. In case of default under both standards I suppose that the exposure will be written down to the appropriate recovery rate - (1-LGD).
So impairment formulas look like the followings:

Table 4: Calculation types

Type of impairment calculation	Calculation formula
IAS 39 Incurred loss	LGD * conditional(t-1)PD * Exposure
IFRS 9 (stage 1) 12 months' expected credit loss	LGD * conditional(t)PD * Exposure
IFRS 9 (stage 2) Lifetime expected credit loss	∑LGD * conditional(t)PD(t) * Exposure(t), where t goes from the current year up to the final maturity of the deal
IFRS 9 expected credit loss	IFRS 9 (stage 1) 12 months' expected credit loss + IFRS 9 (stage 2) Lifetime expected credit loss

RESULT OF THE MODEL

After the calculation with the model introduced in section Modelling process, we got the intended and expected results. According to this the IAS 39 allowance values are less in amount and in case of change of economic circumstances response of impairment increase lags behind IFRS 9 values. So it seems that both hypothesises are proved to be true. Look at the details of the results.

It is worth to mention that if we compare the IAS39 and IFRS 9 requirements without staging rule – so when all the exposures remain in stage 1 - than the IAS 39's values lag behind the IFRS 9 ones. As depicted in Figure 1 the impairment rate – impairment divided by exposure – under IAS 39 starts to increase later at the beginning of the recession and starts to decrease later at the end of the recession compared to IFRS 9. At the beginning it is important, because the loan loss provision appears later in the profit and loss statement and maybe it is resulting that the lending activity is not moderated in time. This feature is illustrated with the higher blue line (IAS39) compared to the purple line (IFRS 9 all exposure in stage 1) in period 11. At the end of it, there is maybe a reverse effect to the lending activity because higher provision goes to the profit and loss statement reducing the bank's willingness to offer loans.

The effect of the increase in amount could be reviled if we compare the red line (IFRS 9) with purple line (IFRS 9 all exposures in stage 1). With it we catch the effect of staging criteria, so in our simple example, where the loans migrate to the rating B, provision for the whole lifetime need to be calculated.

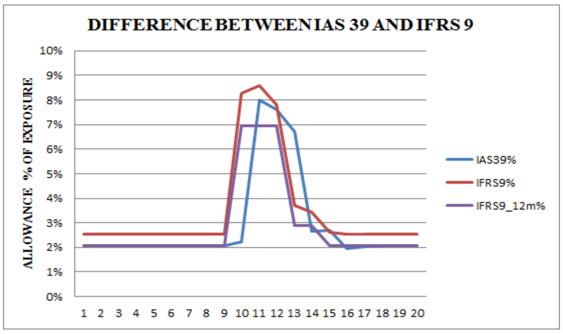


Figure 1: Impairment in IAS 39 and IFRS 9

After the comparison of the result of IAS 39 and IFRS 9 impairment values in different economic states, it is valuable to examine the reaction of the provision to unexpected macroeconomic changes.

To calculate it, I run the model with a baseline and a stress – unexpected – macro factor scenario detailed in Table 2. Because of the fact that the IAS 39 impairment includes only the incurred losses, such changes result no difference in expected and actual provision numbers. But it has effect to the IFRS 9 impairment because of forward looking feature if it. Indeed, it is effective only to the deals which are in stage 2 (rating B), so lifetime expected credit loss need to be calculated. It is illustrated with Figure 2, where red line shows if I do not have information about the upcoming shock and orange line shows, where I have exact information about it. Orange line is higher in its allowance in period 7, 8 and 9. It is because we have already incorporated the higher lifetime provision of the unexpected shock for the remaining exposure of such a deal, which will be still in the portfolio during the recession. So it is clear that the unexpected shock demolish the timeliness effect of the IFRS 9 impairment values.

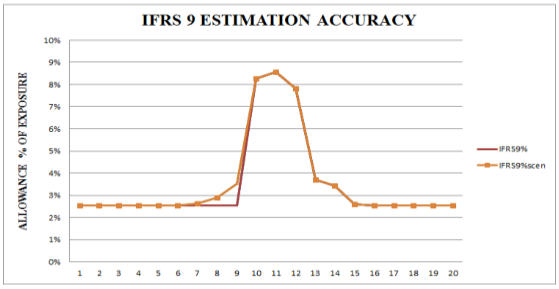


Figure 2: Estimation accuracy

CONCLUSION

This paper has provided an analysis between impairment requirement of the current IAS 39 and upcoming IFRS 9 standards. The focus was on the two main dimensions of impairment recognition, namely time and amount. I set up one hypothesis for each of the dimension. I got the result that in my examination the hypothesis are true, so IFRS 9 recognise loan loss provision earlier and with higher amount then IAS 39. I also show that the timeliness of the provision is demolished if there is an unexpected shock or uncertainty in the economic circumstance.

Clearly, further research is needed to highlight the detailed effect of the IFRS 9's impairment method, maybe with an extended model where prudential regulation is incorporated as well. It means that not only the change and magnitude of loan loss provision are analysed, but together with capital requirement (expected shortfall) and countercyclical capital buffer as well. Some of the issues are already discussed in (Wezel at al. 2012) where dynamic provisioning is analysed as an expected credit loss based method or in (Gruenberger 2012) where capital requirement is incorporated.

REFERENCES

Bank for International settlements. 2011. Basel III:Capital.

Borio C. and Tsatsaronis K, 2005. "Accounting, prudential regulation and financial stability: elements of a synthesis", BIS Working Papers, No 180.

Gebhardt, G. and Z. Nowotny-Farkas, 2011. "Mandatory IFRS Adoption and Accounting Quality of European Banks", Journal of Business Finance & Accounting, Vol. 38(3) & (4), pp. 289-333.

Gruenberger, D. 2012. "Expected Loan Loss Provisions, Business- and Credit Cycles", FMA working paper 2013/01.

IASB, IFRS Foundation Publications Department. 2014..IFRS 9 Financial Instruments. International Financial Reporting Standard, London.

Janecsko, B. 2004. "A Bázel II. belső minősítésen alapuló módszerének közgazdasági- matematikai háttere és a granularitási korrekció elmélete", Közgazdasági Szemle, LI. Évf., 2004. március 218-234.

Szabó, G. 2005. "Bankok és az IFRS-ek". Szakma, 2005/11, 484-486.

Tardos. Á 2005. "Hitelek értékvesztése" IFRS és Basel II. ITCB. Workshop 2005.11.16., Budapest.

Wezel T; J. A. Chan-Lau; and Columba F.,2012. "Dynamic Loan Loss Provisioning: Simulations on Effectiveness and Guide to Implementation", IMF Working Paper, WP/12/110.

AUTHOR BIOGRAPHIES

CSABA KÁDÁR was born in Csongrád, Hungary and went to the Corvinus University of Budapest, where he studied financial risk management and investment analysis and obtained his degree in 2006. He started his PhD in 2015. Before that, he worked as a supervisor at HFSA, risk manager and risk consultant in the financial sector. His research area is prudential and accounting regulation and related financial models. His e-mail address is: csaba.kadar@uni-corvinus.com.

ENHANCING MODEL INTERCHANGEABILITY FOR POWERFLOW STUDIES: AN EXAMPLE OF A NEW HUNGARIAN NETWORK MODEL IN POWERFACTORY AND *e*ASiMOV

Bálint Hartmann

Centre for Energy Research
Hungarian Academy of Sciences
Konkoly-Thege Miklós út 29-33.,
Budapest, 1121, Hungary

E-mail:hartmann.balint@energia.mta.hu

Hüseyin K. Çakmak
Uwe G. Kühnapfel
Veit Hagenmeyer
Institute for Applied Computer Science
Karlsruhe Institute of Technology
Hermann-von-Helmholtz-Platz 1
Eggenstein-Leopoldshafen, 76344, Germany
E-mail: hueseyin.cakmak@kit.edu,
uwe.kuehnapfel@kit.edu, veit.hagenmeyer@kit.edu

KEYWORDS

Hungarian transmission grid, power system simulation, computer modeling, PowerFactory, *e*ASiMOV

ABSTRACT

The article presents a comparison of basic powerflow studies, performed with a market available and a new self-developed power grid analysis software suite. The example of the Hungarian power system is used to show a blueprint process for creating computer models of large-scale electric systems, using dominantly openly available data sources. The first part of the article introduces the Hungarian power systems, and details how topological, consumption and generation data are collected. The second part presents the implementation of these raw data in DIgSILENT PowerFactory and *e*ASiMOV, highlighting the interoperability of the softwares and the additional value of model data conversion.

INTRODUCTION

The fact that all power grid simulation software use their own proprieraty model specification hinders the easy model exchange between various simulation platforms. For model benchmarking purposes the time-consuming remodelling process is a serious problem, especially for large and complex grid models. As a standard for basic model data exchange IEEE CDF (WGoCF 1973), PTI Power Flow Data Format (Portante et al. 1997) and PECO PSAP Format (Christie 1993) are used because of their simplicity and clarity. They are capable of describing grid topology based on nodes and branches with basic network component types. With the Common Information Model (CIM) an ontology for power grid specification has been presented (Uslar 2012). Due to its complexity and being still a work in progress, it is not finally established and also not fully supported by commercial software.

Within the project "Energy System 2050", which is a joint initiative of the research field Energy of the Helmholtz Association (ES2050 2016), the model exchange is essential for the close cooperation of research groups. It will boost the development of new methods for forecasting of loads and power generation with renewable energy sources, uncertainty analysis for power grids, new control algorithms, etc. by providing verified simulation models and validated basic power flow results.

The ES2050 project focuses in detail on the integration of relevant technology elements into the energy system and on the development of solutions for the successful use of the partly strongly fluctuating renewable energy sources in the German and European energy supply systems. Work is aimed at obtaining tangible and exploitable findings and technologies by 2019, which may be used by politics and industry afterwards. The Institute for Applied Computer Science (IAI) at Karlsruhe Institute of Technology (KIT) is the leading partner of Topic 5 “Toolbox with Databases”. In this topic, standardized data formats are to be defined and various data sources are to be combined. It is aimed at developing standardized models for components of the energy system and at conceiving reliable algorithms for planning, operation, and optimization of the energy system. IAI manages the following four work packages:

- Data formats and data quality in database applications;
- Simulation platforms, IT systems architecture, security;
- Prognosis and automated operation planning;
- Instrumentation and control.

Early phase of this work has seen collaboration between IAI and the Centre for Energy Research, Hungarian Academy of Sciences (MTA EK), where similar research is being performed in the field of electricity networks. The joint work of the colleagues of the two institutions aimed to create a new common model for the Hungarian power system that can be implemented in various simulation software and that is based dominantly on openly available datasets.

Grid modelling is traditionally part of the activities of distribution and transmission system operator companies, local and regional utilities and planning enterprises. Thus, the majority of network models are not available for public or research use, due to well-understood interests. The authors of the present contribution aim at negotiating this obstacle by creating a blueprint for modelling regional and national power systems, utilising openly available data to a maximum possible extent. Previous research experience has shown that by identifying the proper data sources (national statistics, the REMIT database, satellite image processing, etc.), such goals can be achieved, especially for high-voltage transmission networks (Çakmak 2015). Another challenge that is addressed is that various power system simulator tools require various input data, making the exchange of models harder and sometimes erroneous. For this reason, two well-known software platforms are chosen for the comparison, the market available DIgSILENT PowerFactory and the IAI in-house developed *e*ASiMOV. The presented research does not only aim to set up the network topology but also to define the behaviour of consumption and generation units of the examined area. This subtask was of primary importance, since proper load-flow analysis requires accurate power data. As the main results of the load-flow calculation are highly dependent on the load at the nodes of the system, the two biggest challenges concerning the distribution of the loads are replicating temporal behaviour and identifying separate groups of consumers. The first issue is necessary to be solved to run a 24-hour time-sweep, while addressing the second issue is necessary to determine the spatial distribution of loads. Due to page limitations, this process is only partially shown in the present contribution, while more focus is put on the compilation of the topological data and the example software environments.

OVERVIEW OF DATA RESOURCES

By the creation of MAVIR Hungarian Transmission System Operator Company Ltd. (hereinafter mentioned as MAVIR) on 1st January 2006, the transmission system operator incorporated the Division for Network Operation from National Power Line Company Ltd. (OVIT) and the Division for Transmission Network from Hungarian Power Companies Plc. (MVM).

Within the transmission activity, it is MAVIR's task to ensure a European-level, safe and balanced electricity supply on its transmission network that enables the improvement of economic life. This transmission network mainly consists of the 750, 400 and 220 kV transmission lines and substations, to connect them and to transmit the electricity.

MAVIR also fulfils the transparency requirements of the Congestion Management Guidelines (Annex to Regulation (EC) 714/2009), thus provides an openly available data on topology, generation and consumption levels, power plants and cross-border exchanges. These data served as the basis for the computer models.

Transmission network

By the end of 2014, MAVIR operated 30 substations (1 x 750/400 kV, 4 x 400/220/120 kV, 1 x 400 kV, 1 x 400/220 kV, 12 x 400/120 kV, 11 x 220/120 kV) (MAVIR 2013a). The total length of MAVIR operated high-voltage transmission lines was 4855 km (268 km of 750 kV, 2978 km of 400 kV, 1393 km of 220 kV, and 199 km of 120 kV). The vast majority of these lines are overhead lines, only 17 km of the network consists of 120 kV cables (MAVIR 2013b, MAVIR 2014a).

Power plants

Due to the economic and financial changes of recent years, the Hungarian power plant portfolio is experiencing a huge setback, both in terms of operating hours and active units. The share of cheap imported electricity is reaching new heights, leading to temporary or final shutdown of power plants. A good example of this process is Gönyű power plant, one of the newest units of the country, which only reached an annual load factor of 34% – in its first years. Another aspect of this process is the lack of physical inertia in the system. During summer off-peak periods, sometimes only three of the major plants are operating, leaving complete regions without inertial support. It is expected though that the picture will change in the near future, the market will rearrange itself, and the present period will only be temporary. To provide a rather proper background for modelling electricity generation in the Hungarian system, the authors have built on the dataset of 2013, the last year that has seen most of the plants operating. In the following, a brief overview is given on the state of the power system (MAVIR 2014b).

The gross installed generation capacity of Hungarian power plants was 9113.1 MW on 31 December 2013. Taking into account constant losses, available capacity was 7521.1 MW. As installed capacity, there is a controllable capacity of 4824.5 MW (or 3780.7 MW considering constant losses) being available via small-scale and large power plants, that were present on the primary, secondary, tertiary and emergency markets in 2013.

Upon the request of AES Tisza II. Power Plant (900 MW), submitted in order to be able to participate in the tender for the "Procurement of Ancillary Services" in 2013, the company regained its operation licence as from 1 January 2013. However, the power plant initiated the suspension of its generation licence. In pursuance to the Resolution, the permission of suspension was valid from 1 July 2013 to 30 June 2016.

Due to the expiry of the Operational Licence of generating units X, XI and XII (645 MW) in block "F" of Dunamenti Power Plant (31 December 2012), the generating units above have become derecognised in the installed capacity of the Hungarian system as from 1 January 2013. The Operational Licence of generating unit XIII (215 MW) in block "F" was extended, later the Power Plant initiated the suspension of the Generation

Licence of this unit with regards to which they signed an agreement with the Transmission System Operator. Accordingly, the generating unit is in "constant non-operational" status as from 1 January 2014.

Referring to unfavourable market conditions, E.ON Hungaria group initiated the suspension of the Generation Licence of Debrecen Combined Cycle Power Plant that was approved by Hungarian Energy and Public Utility Authority. The Licence was valid from 1 July 2013 until 30 June 2016. Furthermore, EON initiated the suspension of operation of Nyíregyháza Combined Cycle Power Plant (49 MW). Pursuant to the agreement signed with the Transmission System Operator, the Power Plant is in "constant non-operational" status with its total installed capacity. The Power Plants intends to suspend its operation until 30 June 2016. Depending on the maintenance periods, Gönyű and Dunamenti G3 Power Plants were available in 2013, however, they were operating only in 10 and 18% of this period, respectively, accounting for a load factor of 7.4% and 13% on an annual basis. This anticipates, depending on future market conditions, the possibility of a "constant non-operational state" of these power plants.

The generation licence of Vértes Power Plant will expire on 31 December 2020. In order to sustain urban area heating, the duration of its availability may be temporarily extended.

Taking into consideration the above mentioned facts, the suggested power system model of Hungary includes 18 power plants (capacity above 50 MW) that were operational in 2013. These units provide approximately 85% of domestic electricity production (26 366 GWh). The largest hydro power plants (Kiskörei and Tiszalöki erőmű), and two wind parks (Bőny and Sopronkövesd-Nagylózs) are also included, however the share of these units in the final portfolio is neglectable. The remaining 15% (2833 GWh) is generated in small power plants, mostly gas engines, but smaller wind parks also belong to this group.

To create a 24-hour representation of the power system, not only the annual generation of the power plants has to be taken into consideration, but one must also examine the generation patterns as well. Based on their behaviour we can define baseload, load-following and peaking power plants. Since the available data is not sufficient to make a good differentiation between the latter two groups, only baseload and load-following plants are modelled. In the Hungarian power system, two power plants provide roughly two-third of the annual generation: Paks nuclear power plant and the Mátrai power plant, running on lignite. During most of the year these two provide baseload capacity, which results in high annual load hours (7685 and 5968 in 2014, respectively), and provides an easy option to model them. In the model, the units are expected to keep a constant output during the day, where the actual output has been determined using the average of the daily load curve and ratio of the plants annual electricity generation compared to total generation. If the resulting baseload output exceeded the available capacity of the power plant (2000 MW for Paks, 920 MW of Mátra), an upper limitation is set. The power output of remaining power plants is calculated in proportion to their share in total annual generation. In several cases the resulting generation curves exceeded the regulation range of the power plants (e.g. running on 20% load), but since their individual share is rather small, this assumption does not create significant errors during load-flow calculations.

Cross-border exchange

Partially as a result of the current situation of power plants, the share of imported electricity in the Hungarian consumption mix has reached new highs. Currently, approximately one-third of the total consumption is based on import, making it theoretically the second biggest "fuel source" in the system. Thus it is necessary to take into consideration the cross-border exchange, when creating the power system model of the country. Since there is no openly available data on the ratio of the loading of individual transmission paths, several assumptions have to be made.

One of the available options is to use annual exchange volumes between Hungary and its neighbouring countries (MAVIR 2014c). The most important tendencies are that majority of imported electricity is transmitted from Slovakia (balance of 8278.3 GWh in 2013), while the main export routes lead to Croatia (balance of -2491.4 GWh in 2013). Since these exchange volumes do not have any temporal characteristic, they can only be utilised to calculate average daily exchange volumes or average export-import power values. However, the temporal behaviour of the exchange is far not constant, and in some cases even the direction of transmitted power may change during a day. For this reason, instead of the previous values, transmission network measurements are taken into consideration in order to determine the cross-border exchange in the model. However, these measurements only indicate total exchange powers on state borders, but provide no information on how this is distributed among the separate transmission paths. To distribute the volume of 6 borders among the 15 transmission paths, nominal parameters of the transmission lines are used namely voltage and maximal loading current, which give us a good assumption on the theoretical transfer capacity of each line. Actual transmitted power on each line is then calculated in proportion to their share in total transfer capacity. The resulting exchange is modelled with a fictive load/generator unit, depending on the net volume of transmitted power.

Consumption

Among the data available on the website of MAVIR, net consumption, net generation and cross-border exchange data are downloaded for two characteristic days of 2014. These two days are the recorded winter and summer peak consumption days, 1 December and 21 July, respectively. The temporal resolution of all data is 15 minutes. Consumption data of the substations, operated by the Hungarian State Railways, are obtained from the existing database of the authors.

To distribute the load among the consumption points of the model, nominal power of substation transformers is used as a basis. In many cases, topology data of power systems includes the ownership of transformers, which allows us to estimate the share of industrial and residential units. However, it would be improper to assume that the industry-owned transformer units are responsible for all industrial consumption, since not all industrial consumers require separate high-voltage feeding points. Smaller industrial consumers are dominantly connected to medium-voltage distribution networks, and thus generate part of the consumption on “residential purpose” transformers.

It is assumed that the daily minimum of the load curve (3480 and 3555 MW in the above examples) is evenly distributed among industrial and residential transformers, in proportion to the ratio of their built-in capacity - as if a constant power level is used for industrial electricity use. To distribute the rest of the load, only residential transformers are taken into consideration.

In Hungary, the power supply of the railway system is using the same 120 kV sub-transmission network as all other consumers. Currently 42 substations provide traction power. Based on the available data, an average daily load pattern is calculated, which varies between 58 and 123 MW, while the average is 100 MW. This load is distributed among the traction substations in proportion to the nominal power of the transformers, installed at each substation.

POWERFACTORY MODEL

The calculation program PowerFactory, as written by DIgSILENT, was collectively chosen by the authors of the present contribution as one of the tools, with the experience being an important factor for the decision.

The constructed network model consists of 120, 220, 400 and 750 kV voltage levels, representing the vast majority of Hungarian transmission and sub-transmission network power lines, substations and power plants.

Transmission system substations are built with a more detailed model, including power transformers as well, while 120 kV substations are created as junction nodes. Net consumption is modelled with the use of general load elements whereas power plants are represented with synchronous machines. Additional shunt reactors (RL units) are also placed at transmission substations for voltage regulation. At the non-Hungarian end of cross-border lines, fictive load/generator units are used to model both import and export power flows.

Due to data handling of PowerFactory being different compared to other modern simulator programs, the authors had two options to model the transmission lines. In PowerFactory, “type” and “element” data are defined. For example, in case of a transmission line, the type data includes per length impedance parameters, while exact length of the line is defined among element data. If the line consists of the same conductor for its total length, and the structure of the pylons is also the same, it is easier to define a set of types, representing the most widely used arrangements. However, if a regional or national system is modelled, it will occur relatively often that actual line parameters show a difference compared to our representative models. Thus it is more practical, to define a new type for every single transmission line, entering total impedance values instead of per length values. This latter solution is used in case of the present contribution.

The model consist of 470 transmission lines, 386 terminals, 307 loads, 64 transformers, 25 generators and 14 shunt reactors. The graphical layout of the model is prepared aiming to replicate the actual layout of network elements. For this purpose, a background layer has been added to the model, which is an official map, issued by MAVIR. National borders are also indicated for better interpretation. The topological map is shown in figure 1. The colours are selected according to the conventional colouring of Hungarian topological maps, where purple, red, green and blue represent 750, 400, 220 and 120 kV, respectively.

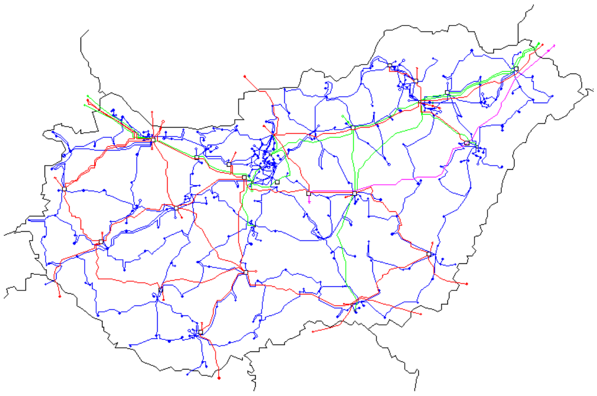


Figure 1: Topology of the Hungarian power system in the PowerFactory model.

To implement the summer and winter peak load days, an additional script is created. The script includes 4 sub-scripts that perform data acquisition from text files. The 4 text files cover generation and load data for the two characteristic days, ordered according to the list of load and generator elements in the model. The main script performs a time-sweep (96 steps in case of 15-minute resolution), at each step setting consumption and generation values, performing a load-flow and storing the results. The script allows the user to decide whether a summer or a winter peak load day is to be simulated. Dunamenti power plant is selected as the system slack. Example results for summer and winter days are shown in figure 2a and 2b, respectively.

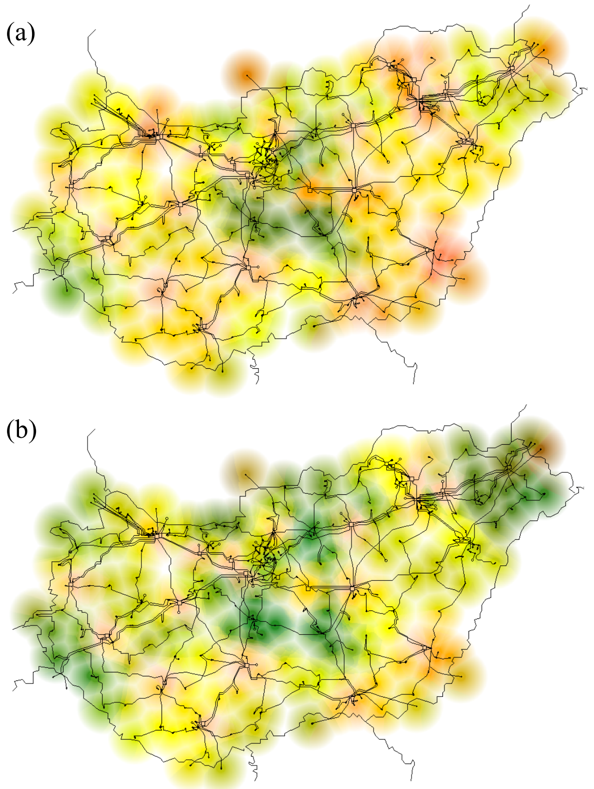


Figure 2: Load-flow results for summer day (a) and winter day (b), colour representation of substation voltages green: 1 p.u., yellow: 1.05 p.u., orange: 1.1 p.u., red: 1.15 p.u.

THE *e*ASiMOV SOFTWARE FRAMEWORK

The new software framework *e*ASiMOV for ***E****lectrical* Grid ***A****nalysis*, ***Si****mulation*, ***M****odelling*, ***O****ptimization* and ***V****isualisation* was initiated as an expandable collection of software tools for interactive power grid modelling and analysis introducing a distributed system architecture (see figure 3).

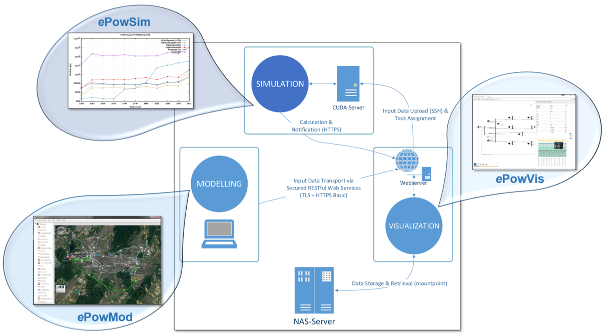

ePowSim
SIMULATION
CUDA-Server
Calculation & Notification (HTTPS)
Input Data Upload (SSH) & Task Assignment
ePowVis
MODELLING
Input Data Transport via Secured RESTful Web Services (TLS + HTTPS Basic)
Webserver
VISUALIZATION
Data Storage & Retrieval (mountpoint)
NAS-Server
ePowMod

Figure 3: Overview of the *e*ASiMOV software framework with the basic software modules *e*PowMod, *e*PowSim and *e*PowVis.

The *e*ASiMOV framework aims at a "One user interface, One simulation model, Multiple simulators" policy. Model data converters enable to export and import models to and from other simulation software. So far direct import and export functionality for Excel based data sheets, and model data for OpenDSS, Matpower and DIgSILENT PowerFactory have been realized. NEPLAN, Simulink and Siemens PSS SINCAL model connectors will be added to *e*ASiMOV in future.

The distributed *e*ASiMOV architecture comprises three main components, which are the Java-based interactive grid modelling editor *e*PowMod and the simulation software module *e*PowSim with the power flow computation engines comprising pure CPU with *Eigen* (Guennebaud 2013), GPU with *Arrayfire* (Malcolm et al. 2012) and a hybrid simulation with *Eigen-CUDA* (Storti 2015) using various solvers.The third component is the Java EE-Server application *e*PowVis that manages external simulators and provides a web based visualisation module for power grid simulation results. It also handles network storages for model and time series data combined with a user database with hierarchical user rights organisation. It supports sharing of simulation models, time series data and simulation results via public hyperlinks.

Also various software modules for analysis and interactive exploration of high rate energy data, web based energy data monitoring, remote software control, model converters as well as support for mobile devices are part of the software suite (Çakmak 2014).

INTERFACES TO POWERFACTORY

Using the PowerFactory Python API (Rueda et al. 2015) and the Java Py4J library (Dagenais 2014) a data and command interface between *e*PowMod and PowerFactory was developed (see figure 4). This enables to import PowerFactory models and to export *e*ASiMOV models. Via the command interface offline simulations can be performed and the results can be imported back to *e*PowVis for visual presentation.

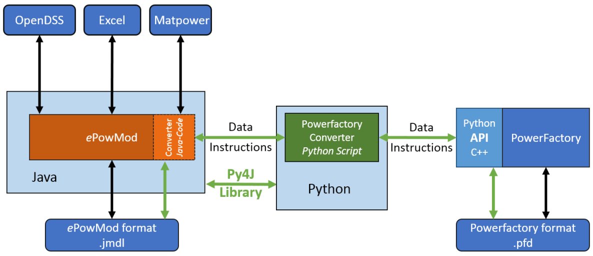


Figure 4: Model conversion and offline simulation in PowerFactory.

After a model in the *e*ASiMOV *jmdl*-format is imported, the bus, line, transformer and generator data arrays are created. The converter then starts the gateway server and calls the phython script and waits for a response in order to stop the gateway server communication. The Python script fetches the grid model data and creates the PowerFactory model dynamically. In a further step the

Python script requests a power flow simulation in PowerFactory and also to export the simulation results. With a similar approach PowerFactory models in *pfd*-format can be imported into *e*PowMod. This enables to create models and execute them in other simulation platforms as Matpower (Zimmerman 2011) or OpenDSS (Montenegro 2012), or simply for data sharing via Excel sheets.

The introduced Hungarian transmission network model in PowerFactory serves as a reference for further optimization of the model conversion algorithms.

THE *e*ASiMOV MODEL

The compiled data resources describing the Hungarian transmission grid are also available as Excel sheets utilizing the IEEE CDF specification.

The *e*PowMod Excel-import filter enables to import the model together with the varying loads as time series data. The 400 kV and 220 kV transmission lines are modeled only roughly, whereas the 120 kV lines are direct node to node connections (see figure 5).

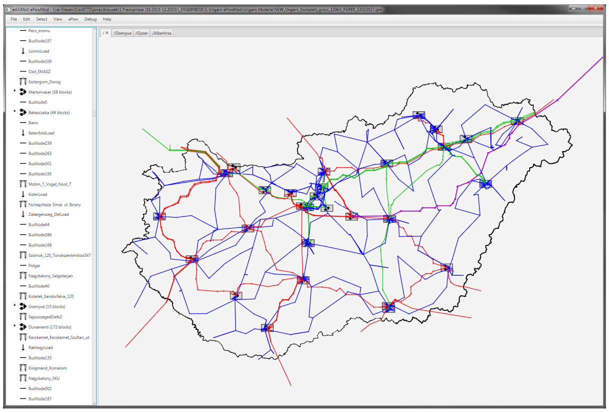


Figure 5: The hungary transmission grid model in *e*ASiMOV-*e*PowMod.

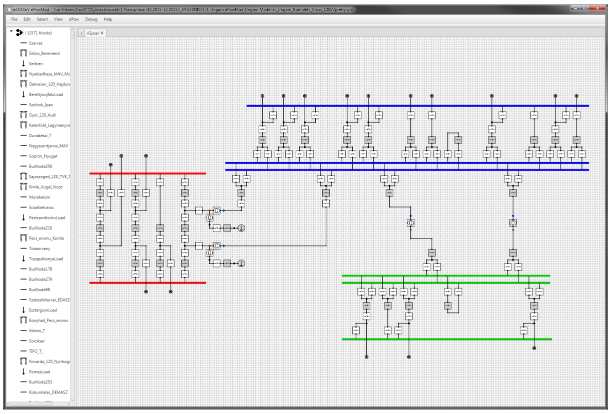


Figure 6: The detail plan of the station Győr showing the buses with the three voltage levels, breakers and transformers.

All 29 substations have been modeled manually according to the MAVIR data (MAVIR 2013a). Figure 6 shows the detailed model of the station Győr with three voltage levels, transformers, switches, loads and connections. The hierarchical model concept in *e*PowMod uses superblocks for a top-down modelling approach that encapsulates local details.

A further benefit of the concept is the ability to store variations of submodels with different levels of details and to compose simulation models easily via configuration scripts.

PARALLEL SIMULATION IN MATPOWER

Since *e*PowMod is fully compatible with the power flow module of Matpower, each of the load cases are exported as individual Matpower simulation files. The Matpower model has a total number of 389 nodes and 683 branches, which may vary from the PowerFactory model, since restrictions in the Matpower model definition needed to be considered. The simulation of 96 load cases takes 10.47 seconds in total, which includes the time for creation of output files and 2.09 second for pure power flow calculation for all load cases. The total number of iterations is 480 to solve the Newton-Raphson algorithm for all cases.

Using a 32 core parallel simulation the total time for 96 load cases is 2.78 seconds including hard disc write operations. All power flow calculations are performed in 0.39 seconds. Thus, the speedup factor is 3.7 for the complete time series simulation and 5.3 for power flow calculations. A significant benefit of parallel simulation is obtained for more complex and large models with several thousands of nodes. The result of one power flow calculation is shown in figure 7 as a heat map.

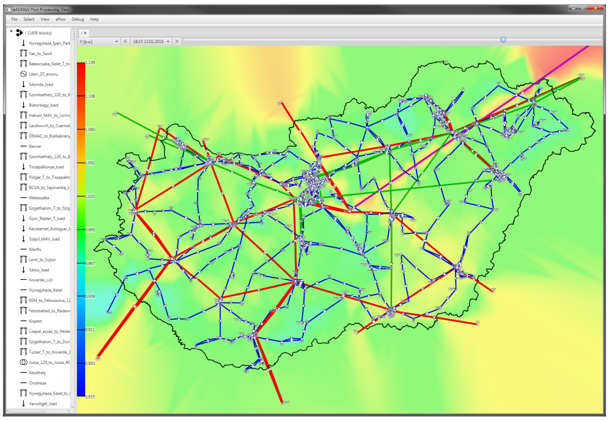


Figure 7: Power flow simulation of a winter day with *e*PowMod using Matpower. The color scale indicates the voltage magnitudes in p.u.

Note that the visualization may differ from those in figure 2 due to the necessary model adaptation, the simplification regarding the lines and the selected load case. In addition, the rendering methods in PowerFactory and in *e*PowMod differ. PowerFactory uses colored circles with radial fading; *e*PowMod uses a Voronoi partitioning (Aurenhammer 1991) and Delaunay triangulation (Fortune 1997). Additionally in figure 7, the superblocks that represent the substations are flattened. The flattening process produces also a non-optimal graphical triangulation due to the premature automated component placement.

FURTHER STUDIES WITH THE HUNGARIAN POWER GRID MODEL

The introduced simulation model is available for various simulation platforms and can be used for further studies. For the optimization of the automated transmission power grid model creation using open source data as the OpenStreetMap data, the model could be used as a further benchmarking model (Çakmak 2015). As part of the R&D at IAI but also as contribution to the ES2050 project (ES2050 2016) and the Energy Lab 2.0 project (Hagenmeyer et al. 2016) the new Hungarian model will be embedded into a detailed EU transmission power grid model. Also selected rural and urban low voltage distribution grids as well as specialized island grids as the KIT campus north 20/0.4 kV power grid with a 1 MW solar park will be integrated.

The EU power grid model will cover renewable energy sources as solar and wind parks besides the traditional power plants and load profiles based on statistical data combined with standardized load profiles.

The introduced Hungary model can also be used for the development of novel forecasting methods for loads (Almeshaiei 2011; Hahn 2009), wind (Monteiro 2009) and solar energy (Ordiano 2016a,b) but also for new stochastic optimal power flow algorithms (Mühlpfordt 2016) and analysis of electricity markets (Keles 2016).

The computation of large and complex power grids can be performed in close cooperation with the SCC (*S*teinbuch *C*entre for *C*omputing) at KIT on the high performance computer ForHLR II with 24.000 computation nodes (Kühnapfel 2016). This facility can also be used to apply genetic and other soft computing algorithms e.g. to power grid scheduling optimization (Blume et Jakob 2009). The management of large-scale data comprising complex simulation models can be handled with generic data services (Süß 2016) using the large scale data facility located at the SCC-KIT (García 2011). The midterm expected extension for *e*ASiMOV will enable parallel dynamic simulation of large power grids as the highly detailed EU model comprising the introduced new Hungarian transmission model with voltage, power and frequency control on the ForHLR II cluster (Kyesswa 2016; Kundur 1994).

CONCLUSIONS

In the present contribution, a novel, detailed model of the Hungarian transmission power grid based on openly available data sources is shown. Power flow simulations with the commercial simulation software DIgSILENT PowerFactory and the *e*ASiMOV software suite, which is developed at the IAI-KIT, are carried out. The value of direct model data conversion in *e*ASiMOV is demonstrated. The Hungarian power grid model can be used for further research e.g. for load and power generation forecasting but also for grid optimization studies. Furthermore, the new model can serve as a basis for future Hungarian power grid expansion planning with renewable energy sources, since only two wind parks are in operation at the moment and more or less no photovoltaic energy is used.

REFERENCES

Almeshaiei, E.; H. Soltan. 2011. “A methodology for Electric Power Load Forecasting”. Alexandria Engineering Journal, 50(2), June 2011, pp. 137-144, ISSN 1110-0168, dx.doi.org/10.1016/j.aej.2011.01.015.

Aurenhammer, F. 1991. “Voronoi diagrams – a survey of a fundamental geometric data structure”. In: ACM Computing Surveys, vol. 23, no.3, pp. 345–405.

Blume, C.; W. Jakob. 2009. “GLEAM - ein Evolutionärer Algorithmus und seine Anwendungen”. KIT Scientific Publishing, Band 32, ISBN 978-3-86644-436-2.

Christie, R. 1993. “The PSAP File Format” in Power Systems Test Case Archive. (15.02.2017). www2.ee.washington.edu/research/pstca/formats/psap.txt.

Çakmak H.K.; H. Maas; F. Bach; and U. Kühnapfel. 2014. “A New Framework for the Analysis of Large Scale Multi-Rate Power Data”. In: KIT Scientific Working Papers 21, Publisher: KIT, Karlsruhe, ISSN: 2194-1629, urn:nbn:de:swb:90-423694.

Çakmak, H.K.; H. Maaß; F. Bach; U. Kühnapfel; and V. Hagenmeyer. 2015. “Ein Ansatz zur automatisierten Erstellung umfangreicher und komplexer Simulationsmodelle für elektrische Übertragungsnetze aus Open-StreetMap-Daten”, at – Automatisierungstechnik, 63(11), pp. 911-925, DOI: 10.1515/auto-2015-0046.

Dagenais, B. 2014. “Py4J - A Bridge between Python and Java”. https://www.py4j.org (15.02.2017).

ES2050. 2016. “Energy System 2050 – A Contribution of the Research Field Energy”,Helmholtz Assoc. (15.2.2017). helmholtz.de/forschung/energie/energie_system_2050.

Fortune, S. 1997. “Voronoi diagrams and Delaunay triangulations”. In: Handbook of discrete and computational geometry, Jacob E. Goodman and Joseph O'Rourke (Eds.). CRC Press, Boca Raton. pp. 377-388.

García, A.; S. Bourov; A. Hammad; J. van Wezel; B. Neumair; A. Streit; V. Hartmann; T. Jejkal; P. Neuberger; and R. Stotzka. 2011. “The large scale data facility: data intensive computing for scientific experiments”. In: 25th IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW) (2011), S. 1467–1474; doi:10.1109/IPDPS.2011.286.

Guennebaud, G. 2013. “Eigen: a c++ linear algebra library. Libraries for scientific computing”. Ecole Polytechnique.

Hagenmeyer, V; H.K. Çakmak; C. Düpmeier; T. Faulwasser; J. Isele; H.B. Keller; P. Kohlhepp; U. Kühnapfel; U. Stucky; S. Waczowicz; and R. Mikut. 2016. “Information and communication technology in energy lab 2.0: Smart energies system simulation and control center with an open-street-map-based power flow simulation example”. In: Energy Technology, 4 (1), pp. 145-162. doi:10.1002/ente.201500304.

Hahn, H; S. Meyer-Nieberg; S. Pickl. 2009. “Electric load forecasting methods: Tools for decision making”. European Journal of Operational Research, Volume 199, Issue 3, 16 December 2009, Pages 902-907, ISSN 0377-2217, http://dx.doi.org/10.1016/j.ejor.2009.01.062.

Keles, D.; J. Scelle; F. Paraschiv; and W. Fichtner. 2016. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. Applied energy, 162, 218–230. doi:10.1016/j.apenergy.2015.09.087.

Kundur, P. 1994. "Power System Stability and Control", McGraw-Hill Education; 1st edition, ISBN-13: 978-0070359581.

Kühnapfel, U.; H.K. Çakmak; D. Piccioni Koch. 2016. "Power Grid Simulation – Using the MATPOWER-Library on bwUniCluster for high-performance parallel Power Flow Timeseries Computation". Workshop HGF-Programm SBD - Schwerpunkt "Energie, 30.6.-1.7.2016.

Kyesswa, M. 2016. "Analysis of power system dynamics". 2016. Doktorandenworkshop der Initiative EnergieSystem 2050, Friedrichsdorf, 7.-8. Dezember 2016.

Malcolm, J.; P. Yalamanchili; C. McClanahan; K. Patel; V. Venugopalakrishnan; and J. Melonakos. 2012. "ArrayFire: a GPU acceleration platform". In: SPIE Defense, Security, and Sensing. pp. 84030A–84030A. International Society for Optics and Photonics.

MAVIR. 2013a. "Transmission network substations of MAVIR Ltd.". MAVIR Hungarian Independent Transmission Operator Company (In Hungarian: A MAVIR Zrt. átviteli hálózati alállomásai).

MAVIR. 2013b. "Transmission lines of MAVIR Ltd.". (In Hungarian: A MAVIR Zrt. átviteli hálózati távvezetékei).

MAVIR. 2014a. "Data of the Hungarian Electricity System".

MAVIR. 2014b. "Medium and lon-term generation capacity development plan of the Hungarian power system". (In Hungarian: A Magyar Villamosenergia-rendszer közép- és hosszútávú forrásoldali kapacitásfejlesztése).

MAVIR. 2014c. "Statistical data of the Hungarian Power System 2013".

Monteiro, C.; R. Bessa; V. Miranda; A. Botterud; J. Wang; and G. Conzelmann. 2009."Wind Power Forecasting: State-of-the-Art 2009", ANL/DIS-10-1, Argonne National Laoratory, http://www.osti.gov/bridge.

Montenegro, D.; M. Hernandez; G.A. Ramos. 2012. "Real time OpenDSS framework for distribution systems simulation and analysis". Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D-LA), Montevideo, 2012, pp. 1-5. doi: 10.1109/TDC-LA.2012.6319069.

Mühlpfordt, T.; T. Faulwasser; and V. Hagenmeyer. 2016. "Solving stochastic AC power flow via polynomial chaos expansion". IEEE Conf. on Control Applications, Buenos Aires, pp. 70-76, doi: 10.1109/CCA.2016.7587824.

Ordiano, J.A.G.; W. Doneit; S. Waczowicz; L. Gröll; R. Mikut; and V. Hagenmeyer. 2016a. "Nearest-Neighbor Based Non-Parametric Probabilistic Forecasting with Applications in Photovoltaic Systems", Proc., 26. Workshop Computational Intelligence, Dortmund, 2016, 2017arXiv170106463A.

Ordiano, J.A.G.; S. Waczowicz; M. Reischl; R. Mikut; and V. Hagenmeyer. 2016b. "Photovoltaic Power Forecasting using Simple Data-driven Models without Weather Data". Computer Science - Research and Development, pp. 1–10.

Portante, E.C.; J.A. Kavicky; J.C. VanKuiken; and J.P. Peerenboom. 1997. "Load Flow Analysis: Base Cases, Data, Diagrams, and Results", AND/DIS/TM-40.

Rueda, J.L.; F.M. Gonzalez-Longatt. 2015. "PowerFactory Applications for Power System Analysis", Springer, ISBN: 978-3-319-12957-0.

Storti, D.; M. Yurtoglu. 2015. "CUDA for Engineers: An Introduction to High-Performance Parallel Computing". Addison Wesley, ISBN: 978-013417741.

Süß, W.; K.-U. Stucky; W. Jakob; H. Maaß; and H.K. Çakmak. 2016. "Generic data services for the management of large-scale data applications". WSEAS transactions on computers, 15, pp. 265-278.

Uslar, M; M. Specht; S. Rohjans; J. Trefke; and J.M. González. 2012. "The Common Information Model CIM", Springer, ISBN 978-3-642-25215-0.

Working Group on a Common Format for the Exchange of Solved Load Flow Data. 1973. "Common Data Format for the Exchange of Solved Load Flow Data". IEEE Transactions on Power Apparatus and Systems, Vol. PAS-92, No. 6, Nov./Dec. 1973, pp. 1916-1925.

Zimmerman, R.; C. Murillo-Sánchez; and R. Thomas. 2011. "MATPOWER: Steady-State Operations, Planning and Analysis Tools for Power Systems Research and Education". IEEE Transactions on Power Systems, vol. 26, no. 1, pp. 12-19.

AUTHOR BIOGRAPHIES

HÜSEYIN KEMÂL ÇAKMAK was born in 1971 in Bolu-Turkey. He holds a Ph.D. in computer sciences from the University Karlsruhe-Germany (2000). He is scientific staff at the Institute for Applied Computer Sciences at the Karlsruhe Institute of Technology. His research interest include modeling, simulation, visualization, 3d/VR/AR, big data analysis, web systems and parallel computing for energy system analysis.
His e-mail address is: hueseyin.cakmak@kit.edu

BÁLINT HARTMANN was born in 1984. He received M.Sc. degree in electrical engineering and obtained his Ph.D. degree from Budapest University of Technology and Economics in 2008 and 2013, respectively. He is senior lecturer with the Department of Electric Power Engineering, Budapest University of Technology and Economics. He is also a part-time research fellow with the Centre for Energy Research, Hungarian Academy of Sciences. His fields of interest include distributed generation, renewable energy sources, energy storage and smart grids.
His e-mail address is: hartmann.balint@energia.mta.hu

UWE G. KÜHNAPFEL is leader of the working group energy systems simulation and analysis at the Institute for Applied Computer Sciences at Karlsruhe Institute of Technology. His main research fields are energy systems (modeling, simulation, monitoring, analysis), mechatronics, virtual and augmented reality.
His e-mail address is: uwe.kuehnapfel@kit.edu

VEIT HAGENMEYER was born in 1971. He is director of the Institute for Applied Computer Sciences at Karlsruhe Institute of Technology. His main field of research is automation technology, control engineering and energy informatics.
His e-mail address is: veit.hagenmeyer@kit.edu

Simulation in Industry, Business, Transport and Services

NO MORE DEADLOCKS – APPLYING THE TIME WINDOW ROUTING METHOD TO SHUTTLE SYSTEMS

Thomas Lienert
Johannes Fottner
Institute for Materials Handling, Material Flow, Logistics
Technical University of Munich
Boltzmannstraße 15, 85748 Garching, Germany
Email: lienert@fml.mw.tum.de, kontakt@fml.mw.tum.de

KEYWORDS
Shuttle systems, Routing, Deadlock handling

ABSTRACT

Autonomous vehicle-based storage and retrieval systems are used in order to supply picking or production areas based on the goods-to-person principle. In one such system, several vehicles move within the same rail system. Hence, routing and deadlock handling is an important issue that has to be resolved carefully to run these systems efficiently and robustly. One possibility for coping with deadlocks is deadlock avoidance by routing with time windows.

In this paper, we present a modelling approach that allows us to apply the time window routing method to shuttle systems. We model the system as a mixed graph and present a concept for moving vehicles safely and efficiently through the storage system.

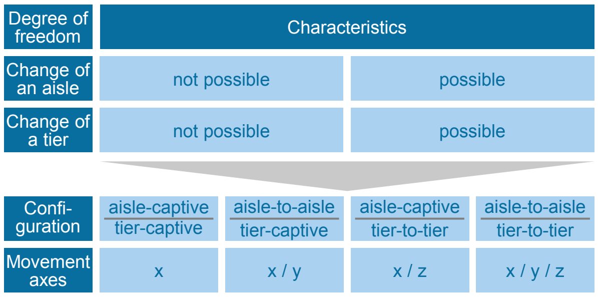
INTRODUCTION

In addition to stacker-crane-based automated storage and retrieval systems (AS/RS), a new technology has been introduced to the market, based on autonomous vehicles. Autonomous vehicle-based storage and retrieval systems (AVS/RS) are used for storing unit loads in order to supply picking and production areas based on the goods-to-person principle (VDI- Richtlinie 2692).

AVS/RS, also known as shuttle systems, are characterized by horizontally operating vehicles. These vehicles travel within a rail system that is integrated into the storage rack. Lifts positioned along the periphery of the storage rack system are used to perform storage and retrieval transactions (Malmborg 2002).

Over the course of recent developments, different system configurations have evolved, which can be classified by the movement space of the vehicles. In the most common configuration, the vehicles are restricted to a single storage aisle and tier. By contrast, in other configurations, the vehicles are able to change tier by using lifts and move between storage aisles by using cross aisles, which are positioned orthogonally to the storage racks.

Figure 1 provides an overview of the four configurations that result from different movement spaces of the vehicles. The x-axis corresponds to the storage aisles, the y-axis to the lifts, and the z-axis to the cross aisles.



Degree of freedom	Characteristics			
Change of an aisle	not possible		possible	
Change of a tier	not possible		possible	
Configuration	aisle-captive tier-captive	aisle-to-aisle tier-captive	aisle-captive tier-to-tier	aisle-to-aisle tier-to-tier
Movement axes	x	x / y	x / z	x / y / z

Figure 1: Shuttle System Configurations

Shuttle systems with aisle- and tier-captive vehicles provide the highest throughput, as the vertical and horizontal movements are completely decoupled from each other. Shuttles hand over the storage units to buffer locations and do not have to wait for the lifts.

But as the number of degrees of freedom increases, the flexibility of the system improves. One important characteristic of shuttle systems with aisle-to-aisle and tier-to-tier vehicles is that every vehicle can reach every single position within the storage system. Therefore it is possible to run the whole system with a single shuttle. If needed, the number of shuttles can be gradually increased in order to achieve a higher throughput. As every storage unit can be delivered to every lift and therefore to every input/output location (I/O location), no merges are required within the pre-storage area. Furthermore, the storage units can be delivered in the desired sequence at every I/O location.

However, such a configuration requires a more complex control strategy in order to run the system robustly and efficiently. The main issues that have to be addressed by the control are dispatching and routing: Where and when should a vehicle travel? And which route should be taken to reach a designated position?

In our research we investigate these questions, focusing on shuttle systems with aisle-to-aisle and tier-to-tier configurations (see figure 2), and evaluating the developed strategies based on simulations.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

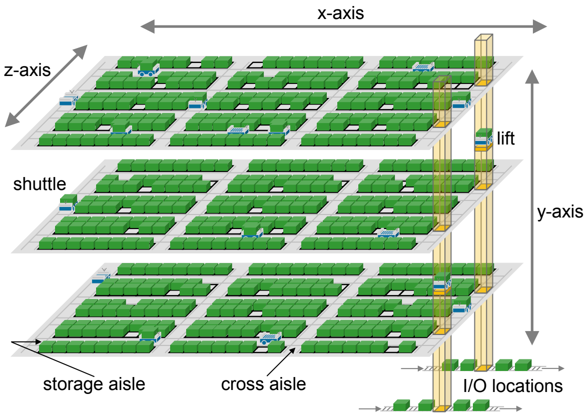


Figure 2: Example of a System

The routing must not only choose the routes themselves, but must also take deadlocks into account, since several shuttles are moving within the same rail system. The notion of deadlock describes a situation where one or more concurrent processes in a system are blocked forever because the requests for resources by the processes can never be satisfied (Kim et al. 1997). In the context of routing vehicles, the processes correspond to the execution of each route and the resources correspond to the layout segments along these routes.

Three generic approaches can be distinguished in deadlock handling: static deadlock prevention; detection and recovery; and dynamic deadlock avoidance, which generally allows the highest resource utilization (Liu and Hung 2001). One possibility for avoiding deadlocks is the time window routing method, which was introduced by Kim and Tanchoco (Kim and Tanchoco 1991). The main idea of this approach consists in modelling the layout as a graph. For each node, the algorithm maintains a list of time windows reserved by routed vehicles and a list of free time windows in which vehicles can be routed.

Since the algorithm is guaranteed to find the fastest deadlock-free path for the vehicles from their start node to their destination node at the specified start time, it is a promising approach. Adapting this algorithm for routing shuttles seems worth investigating and is the object of this paper.

LITERATURE REVIEW

In this section, we give a review of the research into control strategies for tier-to-tier and aisle-to-aisle configurations, followed by a presentation of some applications that implement the time window routing method in different logistical contexts.

Most of the recent research concerning AVS/RS considers tier- and aisle-captive configurations and investigates the performance of the considered system in terms of parameters such as the number of storage bays and tiers or different velocity profiles for the vehicles and lifts. Only a few papers consider control strategies for tier-to-tier and aisle-to-aisle configurations.

Ekren et al. (Ekren et al. 2010) investigated the effect of several design factors on the performance of the storage system. The authors varied the dwell point and the I/O location and used basic dispatching rules to assign storage or retrieval transactions to vehicles.

Roy et al. (Roy et al. 2014) developed simple protocols to cope with deadlocks. If there is a deadlock within a storage aisle, the rearmost vehicle travels to the last available bay location where it waits until the other vehicle completes its task in that aisle. These blocking effects were quantified by an analytical model based on queuing network theory. Their strategy is not robust, as several vehicles can enter into a deadlock, and the applicability is limited to the specific layout that they considered.

Penners (Penners 2015) examined a simplified, isolated tier of an aisle-to-aisle system. He adapted two deadlock-avoiding routing algorithms and compared the performance by conducting a simulation study. He came to the conclusion that the time window routing method, modified by ter Mors et al. (ter Mors et al. 2007), achieves a considerably higher throughput than modified Banker's routing, which was described by Kalinovcic et al. (Kalinovcic et al. 2011). He modelled the tier as a graph, where the nodes represent lifts, lift buffers, storage aisles and crossing aisles. The use of the nodes is exclusive, which means that only one shuttle can occupy any given storage aisle at the same time, which does not take into account the real size of the aisles.

The concept of time window routing, which showed promising results, was first introduced by Kim and Tanchoco (Kim and Tanchoco 1991) for the conflict-free routing of automated guided vehicles in a bidirectional network whose nodes represent important locations such as load transfer stations, parking lots, and battery charging stations. These nodes are interconnected by lanes, which are either unidirectional or bidirectional. The size of the nodes corresponds to a check zone size, which protects vehicles from collisions. The use of the nodes is exclusive, but several vehicles can move along the same edge, as long as head-on and catching-up conflicts are prevented. Maza and Castagna (Maza and Castagna 2005) considered automated guided vehicles and implemented time window routing as proposed by Kim and Tanchoco. They developed a procedure to avoid deadlocks in the presence of interruptions while maintaining the planned routes. They showed that the absence of deadlocks is guaranteed if the node's crossing order of the vehicles based on the conflict-free scheduled date is fulfilled, even if the arrival dates are not.

Busacker (Busacker 2004) developed an time-window-based routing algorithm for optimizing aircraft taxi traffic at airports. The airport is modelled as a graph whose edges correspond to taxiways, parking positions, or any other locations that aircraft might occupy. The edges are weighted by the travel time of the airplanes and connected by nodes that do not have any physical size. As the use of the edges is exclusive, long edges are divided into several shorter edges in order to obtain a higher utilization of resources. The size of the airplanes is not modelled.

Stenzel (Stenzel 2008) used the time window routing approach to route automated guided vehicles within container terminals. The moving area is modelled by a grid graph. Routes are computed in two steps. Firstly, a time window route is calculated, followed by a readjustment that takes into account the real size of the vehicles.

Ter Mors (ter Mors 2010) presented a generic model for routing agents through an infrastructure graph. To calculate the route, a resource graph is generated whose nodes correspond to the nodes and edges of the infrastructure graph. The edges of the resource graph can be interpreted as a successor relation. His version achieves better worst-case performance than the original algorithm by Kim and Tanchoco and calculates a solution in real time.

In summary, routing and deadlock handling for aisle-to-aisle configurations have so far only been considered in a few papers, and only in idealized terms that do not allow the developed strategies to be applied efficiently to real systems. It has been shown that the time window routing method is a promising approach that has so far been successfully applied to different contexts, and diverse answers have been given to the question of how the infrastructure of the considered system can be modelled as a graph. We will apply the time window routing method to shuttle systems. We therefore adapt the generic version by ter Mors. His approach is described in the following section.

TIME WINDOW ROUTING METHOD

The time window routing method is used to obtain deadlock-free routes for vehicles moving through infrastructure modelled by a graph. For each node, the algorithm maintains a list of free time windows through which vehicles can be routed. Each free time window is defined by the end of the preceding reserved time window and the beginning of the subsequent reserved time window, apart from the final free time window at each node, which is endless (see figure 3).

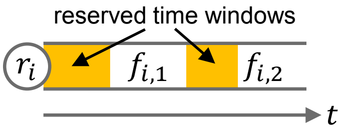


Figure 3: Free Time Windows f_i on the Node r_i

The algorithm inputs consist of the start node r_1, the start time, and the destination node r_n. The output is a plan π.

$$\pi = (\{r_1, [t_1, t'_1]\}, \dots, \{r_n, [t_n, t'_n]\}) \qquad (1)$$

The plan π (1) contains all the nodes along the fastest path $r_1, \dots, r_n$ and the corresponding time windows, which are defined by the entry times t_i and the exit times t'_i at the nodes r_i.

The algorithm consists of two consecutive steps that are executed iteratively. These steps are:

1. Investigate the reachability of all free time windows on all neighbouring nodes.
2. Select the most promising time window for the next iteration.

Starting with the initial time window on the start node, each iteration of the algorithm investigates the reachability of all free time windows on all neighbouring nodes (see figure 4). This procedure is called time window expansion.

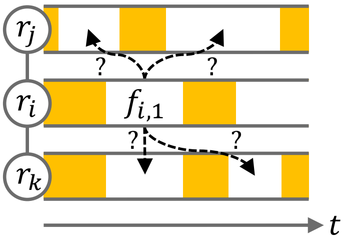


Figure 4: Time Window Expansion

In order for a free time window to be reachable, some conditions must be met. A free time window on a neighbouring node is reachable from the current free time window if:

- the free time window is larger than the minimal duration required for a vehicle to enter, traverse, then exit the resource again,
- both free time windows overlap,
- the current free time window can be completely exited before it ends,
- the remaining duration after entering the free time window is large enough to traverse and exit the resource.

If a free time window is reachable, it is added to a list called open list, which contains the time windows for further expansions.

In the next iteration, the most promising time window is selected from the open list and removed. The most promising time window is the one that allows the final destination to be reached in the theoretical minimal time. This time $y(f)$ is the sum of earliest possible exit time $c(f)$ from the current node and the estimated time $h(f)$ required to complete the routing to the destination:

$$y(f) = c(f) + h(f) \qquad (2)$$

The algorithm terminates as soon as a free time window belonging to the destination node is selected for the next iteration. Once this happens, the plan π is constructed using back pointers and the corresponding time windows are reserved.

The example in figure 5 shows a graph and the fastest path through the free time windows from the start node r_1 to the destination node r_5. The overlapping between free time windows is illustrated. Note that the fastest path might visit a node twice.

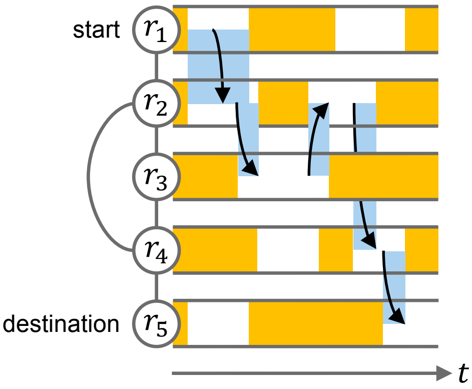


Figure 5: Fastest Path from the Start Node r_1 to the Destination Node r_5

If a free time window needs to be selected for the next iteration from the open list but the open list is empty, there is no route available. This can happen if the time window on the start node is not endless or if vehicles are allowed to dwell on nodes indefinitely.

The search for the fastest path through the free time windows is an application of the A* algorithm, which is used to find a shortest path in a graph (Hart et al. 1968). Each free time window corresponds to a node in a so-called free time window graph. The arcs between these nodes represent the reachability between pairs of free time windows. The algorithm runs in polynomial time in the size of this time window graph. Hence, assuming a feasible number of nodes in the resource graph and a feasible number of vehicles, the algorithm returns a solution in real time.

As a closing remark, we should note that the sequence in which vehicles are routed matters. The algorithm finds an optimal solution for a single vehicle under the given reserved time windows but does not find the global optimum (e.g. the makespan). In the shuttle system, the most commonly encountered process is that of a single shuttle that occasionally arrives on a storage tier needing to be routed. Apart from initialization, it is unlikely that a larger number of shuttles will need be routed at the same time. Hence the presented approach can be used for our purposes.

In the remainder of this paper, we model the tier of the shuttle system as a graph in order to apply the time window routing method, which allows high resource utilization to be achieved. We must therefore answer the question of what the nodes and edges should represent, and how their sizes should be determined.

Below, we modify the time window routing method proposed by ter Mors (ter Mors 2010) in order to use it to route the shuttles on a storage tier.

Finally, we present a control strategy that allows the calculated routes to be executed even if there are delays and uncertainties.

MODELLING THE SYSTEM

The components of a shuttle system are the storage rack, the rail system, the lifts, and the shuttles. In order to apply the time window routing method, the rail system must be modelled as a graph. We apply the concept of the resource graph, which means that the edges simply represent a successor relation and do not have any physical size. The nodes are divided into the following types (see figure 6):

- Storage aisle nodes
  Storage aisle nodes are placed within a storage aisle. In order to access a storage location, shuttles must occupy these nodes.
- Cross aisle nodes
  Cross aisles are positioned orthogonally to the storage racks. The nodes are placed between the crossing nodes and are used to travel between storage aisles.
- Crossing nodes
  Crossing nodes interconnect at least three aisle nodes and allow a shuttle to perform a 90 degree turn in order to move from a cross aisle into a storage aisle or vice versa.
- Buffer nodes
  Buffer nodes are placed at the edges of storage aisles or cross aisles and are used for buffering idle shuttles. This prevents idle shuttles from blocking other shuttles that are performing a storage or retrieval transaction.
- Lift buffer nodes
  Lift buffer nodes are placed on both sides in front of the lifts and are used as input and output buffers on every storage tier.
- Lift nodes
  Lift nodes represent the lifts. They can be entered only if the lift is empty and is currently located on that tier.

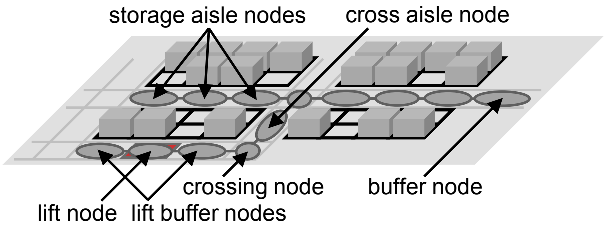


Figure 6: Different Types of Nodes on a Storage Tier

Except for the lift buffer nodes, all nodes can be traversed in both directions. We therefore use a mixed graph in which only the edges incident to the lift buffer nodes are directed, since lifts are accessed from one side and exited from the other side, as this allows a higher throughput to be achieved.

Every node has an attribute that represents the axis along which the vehicle is travelling at that node. For instance, aisle nodes allow only movement along the x-axis and cross aisle nodes allow only movement along the z-axis. Crossing nodes allow both orientations. Whenever a time window is expanded to a time window on a crossing node, the orientation of the vehicle entering the crossing node is stored. If the subsequent node has a different orientation, a 90-degree turn on the crossing node is necessary.

In order to model the rail system, we enforce the following conditions:

1. The use of each node is exclusive.
2. The size of any given node must not be smaller than the size of the shuttle.

The first condition ensures that no sub-routing is required on single nodes, as only one shuttle at a time is allowed to move on the node. The second condition avoids having to take into account more than two nodes for the reachability check, which would considerably increase the complexity of the algorithm. To achieve high resource utilization, the nodes should be as small as possible.

The size of the buffer nodes, lift nodes and cross nodes corresponds to the layout data, as we assume only one shuttle can occupy these nodes. The length L_{Aisle} of both the storage aisles and cross aisles between two crossing intersections is also given, as well as the length of the shuttles $L_{Shuttle}$, which could be replaced by the minimum node length. The number N_{nodes} of nodes within a storage or cross aisle and their lengths L_{nodes} are calculated as follows:

$$N_{nodes} = floor(\frac{L_{Aisle}}{L_{Shuttle}})$$

$$L_{nodes} = L_{Shuttle} + \frac{L_{Aisle} - N_{nodes} * L_{Shuttle}}{N_{Nodes}}$$

If a shuttle needs to be buffered or needs to use a lift, it will be routed to the corresponding node. By contrast, the aisle nodes do not represent destinations in the system themselves.

In order to fulfil a storage or retrieval request, a shuttle must be routed to a storage location. Since the size of the shuttles exceeds the size of the storage units and therefore the size of the storage locations, the size of the storage locations is smaller than the length of the aisle nodes.

In the example shown in figure 7, the storage aisle consists of four nodes. In order to access a storage location, for some storage locations it is sufficient to reserve a single node (e.g. storage location no. 2), whereas for some storage locations two nodes must be reserved (e.g. storage location no. 9).

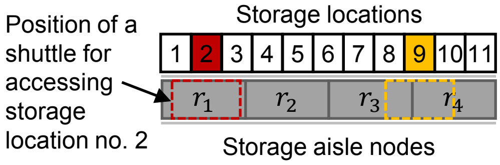


Figure 7: Storage Aisle Nodes

The nodes that have to be reserved in order to perform a storage or retrieval transaction can be easily identified by considering the shuttle's exact position when accessing the location.

MODIFICATIONS TO THE ALGORITHM

In order to use the described time window routing method, some modifications are necessary, which we will describe below.

1. We select the most promising time window for the next iteration in a slightly different way, as we cannot assign an earliest possible exit time to a time window.
2. The algorithm does not necessarily stop as soon as the destination node is reached for the first time, since the time window must guarantee a minimal remaining size after entrance.
3. Instead of a single destination node, we use a set of destination nodes, as it might be necessary to reserve several nodes in order to access a storage location.

Recall that the most promising time window is identified by the value

$$y(f) = c(f) + h(f) \tag{2}$$

We cannot associate an earliest possible exit time with each time window, because this time also depends on the orientation of the neighbouring node. If a 90-degree turn is necessary, the exit time is postponed. Hence, for the expression $c(f)$, we use the arrival time at the node, which describes the moment when the shuttle is located at the centre of the node. Furthermore, to estimate the remaining travel time $h(f)$, we use the fastest possible path from the current node (and orientation) to the destination node without any reservations within the system.

The described version of time window routing terminates as soon as a time window belonging to the destination node is selected. For our purposes, in some cases the time window must be large enough to guarantee a minimal remaining length. Shuttles travel across each storage tier in order to perform storage or retrieval transactions. Each of these transactions requires a certain amount of time. Therefore, if a storage or retrieval transaction needs to be fulfilled at a destination node, the remaining size of the time window must allow the performance of the transaction and full exit from the node.

Instead of a single destination node, we use a set of destination nodes called the destination set. This set usually contains a single node, but when routing towards the storage locations, it contains all the nodes that must be reserved in order to perform the storage or retrieval transaction. Hence, the set contains either one or two nodes. As soon as a time window belonging to the destination set is selected for the next iteration, the algorithm checks whether there is another node in the set. If there is, there must be a free time window on that node that is reachable from the current time window. The overlapping of these two time windows must not only allow the movement of the vehicle from one node to the other, but also must be long enough to allow the storage or retrieval transaction to be performed.

If no time window is reachable on the other node, or the remaining size of the time window does not allow the storage or retrieval transaction to be performed, the search through the free time windows is continued until a later free time window is found that guarantees these conditions.

IMPLICATIONS FOR THE CONTROL

The result of the routing process is a list with the nodes that have to be visited in order to reach all designated destinations on a tier and the corresponding time windows during which shuttle will move to and occupy these nodes. As the route is calculated using idealistic times and neither accelerations and decelerations are fully taken into account, it is not sufficient to move the shuttles according to their reserved time windows. Furthermore shuttles are routed to the lifts. But when the routing is calculated, it is not obvious when the shuttle will enter the lift and exit the previous node, as the control of the lifts is decoupled from the routing. Consequently, delays will be passed from one node to another on each storage tier.

Therefore it is not possible to navigate the shuttles by the time windows alone; only according to the sequence of reserved time windows on the nodes. As Maza and Castagna (Maza and Castagna 2005) showed, the absence of deadlocks is guaranteed as long as the nodes' crossing order is preserved.

Hence we introduce the concept of claiming nodes. A shuttle is allowed to enter a node only if it has previously claimed that node. A node can be claimed by a shuttle if and only if the earliest reserved time window on that node was reserved by the shuttle. Therefore a node cannot be claimed by several shuttles simultaneously.

In order to clarify this process, we consider the following simplified example, shown in figure 8, with three shuttles that must be routed.

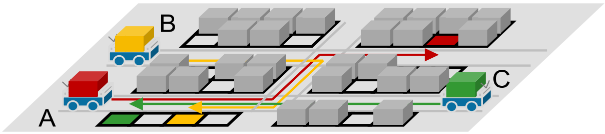


Figure 8: Routing on a Storage Tier

Firstly, shuttle A routes and reserves its time windows, followed by shuttle B and shuttle C. Note that shuttle B has to wait a certain period of time on node r_{14} before it can enter node r_{15}, and shuttle C also has to wait on node r_6.

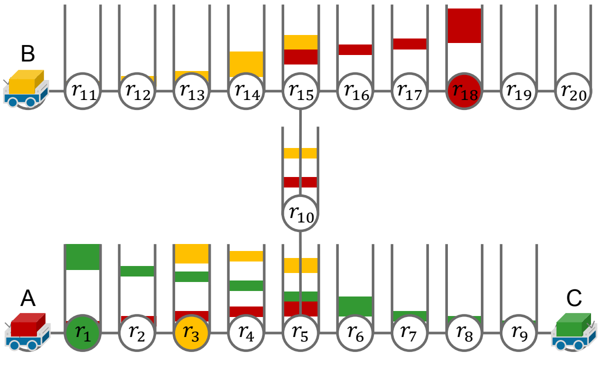


Figure 9: Reserved Time Windows by the Shuttles

For each node, we know which shuttle will occupy the node during each time period. We therefore also know the sequence of shuttles that will cross this node. If we consider the reservation list of crossing node 5, we get the following information.

Table 1: Reserved Time Windows on Node 5

Shuttle	Entry Time	Exit Time
A	00:00:08	00:00:12
C	00:00:12	00:00:15
B	00:00:17	00:00:21

At this node, the sequence of shuttles is A, C, B. This sequence has to be established so that deadlocks can be avoided even if the shuttles are late for some reason.

Whenever a shuttle starts moving, we identify the nodes that could be traversed by the shuttle. These nodes are then claimed by the shuttle. In the example, the following nodes are claimed.

Table 2: Nodes Claimed by the Shuttles

Shuttle	Claimed Nodes
A	$r_1, r_2, r_3, r_4, r_5, r_{10}, r_{15}, r_{16}, r_{17}, r_{18}$
B	$r_{11}, r_{12}, r_{13}, r_{14}$
C	r_9, r_8, r_7, r_6

Shuttle A was able to claim its whole route, whereas shuttle B and C could claim only part of their routes. If a shuttle was able to claim at least one other node, it starts moving and stops as soon as it finishes a claimed segment. It will then start claiming the next segment of nodes again.

If a shuttle cannot claim a single node, it will register as waiting shuttle at this node. Whenever a shuttle exits a node completely, not only is the reserved time window deleted, but the algorithm also checks whether any other shuttle is registered as waiting for that node. If another shuttled is registered on the node, it is triggered and will claim its next segment.

The concept of claiming nodes is necessary. Examining how far a shuttle is allowed to travel is insufficient, since another shuttle can reserve the earliest time window at any moment. The shuttle that had previously reserved the earliest time window would then no longer be allowed to enter that node. Therefore, whenever the earliest time window is reserved on a node, the algorithm checks whether that node has already been claimed by another shuttle. If it has, the node is released by that shuttle, as well as all subsequent nodes along the route that have been already claimed.

SUMMARY

In this paper we described shuttle systems as a technology for storing small unit loads. We focused on systems with a tier-to-tier and aisle-to-aisle configuration, which provide high flexibility. In these systems, every vehicle can reach every storage location. From the perspective of control, routing becomes an important issue due the possibility of deadlocks among the shuttles, which must be dealt with.

As a concept for routing and handling deadlocks, we referred to the time window routing method that has already been successfully applied in different logistical contexts. We adapted the time window routing method and modelled the tier of the shuttle system as a graph in order to apply the method. Finally, we described a concept that enables the vehicles to execute the calculated routes.

The time window routing method was implemented in a generic simulation model kit for shuttle systems, as well as the concept of claiming nodes. In future work, these concepts will be evaluated by simulation experiments.

The concept of claiming nodes could be expanded. It might be possible to allow nodes to be claimed by shuttles that did not reserve the earliest time window on these nodes if this improves the overall efficiency and the absence of deadlocks can be guaranteed. Furthermore, fully integrating the lifts into the time window routing scheme might be interesting.

REFERENCES

Busacker, T. 2005. *Steigerung der Flughafen-Kapazität durch Modellierung und Optimierung von Flughafen-Boden-Rollverkehr – Ein Beitrag zu einem künftigen Rollführungssystem*. Dissertation. Technische Universität Berlin.

Ekren, B. Y., Heragu, S. S., Krishnamurthy A. and Malmborg C. J., 2010. "Simulation based experimental design to identify factors affecting performance of AVS/RS." *Computers & Industrial Engineering* 58*,* No.1, 175-185.

Hart, P. E., Nilsson, N. J. and Raphael, B. 1968. "A Formal Basis for the Heuristic Determination of Minimum Cost Paths" *IEEE Transactions of Systems Science and Cybernetics* 4*,* No.2, 100-107.

Kalinovcic, L. , Petrovic, T. , Bogdan, S. and Bobanac V., 2011. "Modified banker's algorithm for scheduling in multi-agv systems." *Automation Science and Engineering* (Trieste, Italy, Aug. 24-27), 351–356.

Kim C. W. and Tanchoco J. M. A., 1991. "Conflict-free shortest-time bi-directional AGV routing." *International Journal of Production Research* 29*,* No.12, 2377-2391.

Kim C. W., Tanchoco J. M. A. and Koo P., 1997 "Deadlock Prevention in Manufacturing Systems with AGV Systems: Banker's Algorithm Approach." *Journal of Manufacturing Science and Engineering* 119, No.4, 849-854.

Liu F. and Hung P., 2001. "Real-time deadlock-free control strategy for single multi-load automated guided vehicle on a job shop manufacturing system." *International Journal of Production Research* 39*,* No.7, 1323-1342.

Malmborg C. J., 2002. "Conceptualizing tools for autonomous vehicle storage and retrieval systems." *International Journal of Production Research* 40*,* No.8, 1807-1822.

Maza, S. and Castagna, P., 2005. "A performance-based structural policy for conflict-free routing of bi-directional automated guided vehicles." *Computers in Industry* 56, No.7, 719-733.

Penners L. T. M. E., 2015. *Investigating the effect of layout and routing strategy on the performance of the Adapto system*. Master's Thesis. Eindhoven University of Technology.

Roy, R., Krishnamurthy, A. and Heragu, S. S. 2014. "Blocking Effects in Warehouse Systems With Autonomous Vehicles." *IEEE Transactions on Automation Science and Engineering* 11, No. 2, 439-451.

Stenzel, B. 2008. *Online Disjoint Vehicle Routing with Application to AGV Routing*. Dissertation. Technische Universität Berlin.

ter Mors, A. W., Zutt, J. and Witteveen C., 2007. "Context-Aware Logistic Routing and Scheduling." In *Proceedings of the Seventeenth International Conference on Automated Planning and Scheduling* (Providence, USA, Sep. 22-26), 328-335.

ter Mors, A. W. 2010. *The world according to MARP*. Dissertation. Technische Universiteit Delft

VDI-Richtlinie 2692 Blatt 1, 2015. Automated vehicle storage and retrieval systems for small unit loads. Berlin: Beuth.

THOMAS LIENERT has been working as a research assistant at the Institute for Materials Handling, Material Flow and Logistics, Technical University of Munich, since 2014. His research deals with the development of control strategies for autonomous vehicle-based storage and retrieval systems. His email address is: `lienert@fml.mw.tum.de`.

JOHANNES FOTTNER is professor and head of the Institute for Materials Handling, Material flow, Logistics at the Technical University of Munich.

THE WORKER ALLOCATION PLANNING OF A MEDICAL DEVICE DISTRIBUTION CENTER USING SIMULATION MODELLING

Kittikhun Iamsamai and Thananya Wasusri
Logistics Management Program, Graduate School of Management and Innovation
King Mongkut's University of Technology Thonburi
126 Prachautid Rd., Bangkok, Thailand
E-mail: kittikhun.iam@mail.kmutt.ac.th, thananya.was@kmutt.ac.th

KEYWORDS
Worker allocation, Simulation, Worker workload
Medical device distribution center, Inbound logistics process

ABSTRACT

Medical Devices are essential for medical services. It is very necessary to manage them to be ready for quick responding on the demand of the patients that can be urgent. The case study is a Medical Device Distribution Center in Thailand and its major task is to manage medical devices for its clients starting from receiving to shipping. For the inbound logistics process, there are two types of receiving products. The first type is Goods Receive, which is to receive new products. The second type is Goods Return, which is to receive the returned products from customers. The received products will be sent to five different departments. Each department has a different amount of incoming products for each day. The inequality workload clearly affects the inefficiency of the inbound logistics process. This research is then conducted to reduce the lead time of the inbound logistics process. Discrete event simulation modeling using ARENA was utilized to allocate workers in the inbound process. The new worker allocation plan can reduce the average total inbound logistics process time for all product groups. Moreover, the average utilization of each worker is about 30%-40%.

INTRODUCTION

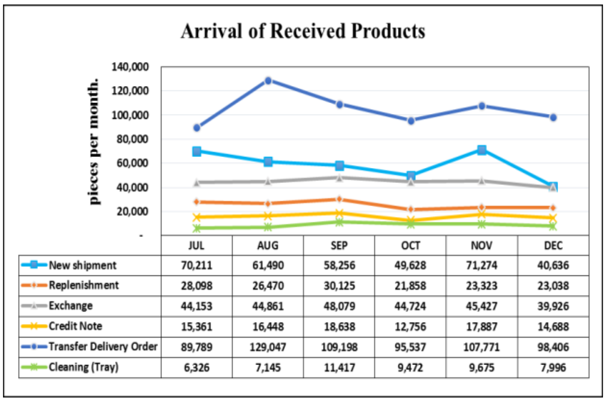
A Medical Device Distribution Center is operated 24 hours a day in order to respond with urgent needs that may exist. Its clients are the owners or agents of medical devices. The medical devices consist of four products that are Cardiovascular, Orthopedic, Ophthalmic and Dental. The case study's major tasks are inbound logistics, warehouse management and outbound logistics. The case study receives medical devices from its clients and conducts warehouse management for those devices received. Once there is a need from hospitals or clinics, the case study will conduct the outbound logistics process in order to send the devices to customers effectively. As the medical devices can be urgent according to the patients' need, lead time reduction is a major issue for the case study. For the inbound logistics process of the case study, there are two product types to receive. The first type is called 'Goods Receive' and it means to receive new products. It can be divided into two groups.

1. Newshipment is a new product from customers.
2. Replenishment is to receive and store products from the main Distribution Center.

The second type is called 'Goods Return' and it means the returned products by the customers. It is divided into four groups.

1. Exchange is the product that the customers want to exchange to other product specifications. Ophthalmic products are mostly found for exchanging.
2. Credit note is the returned product for debt reduction.
3. Transfer Delivery Order is the returned product after the operation (Implant) and it is often orthopedic products.
4. Cleaning is the instruments for the operation. The instruments are returned for cleaning.

The arrival of received products are highly fluctuated for each month as shown in figure 1. It can be seen that Transfer Delivery Order is fluctuated between 80,000 and 130,000 pieces per month. While Credit note is spreading around 12,000-19,000 pieces per month.



	JUL	AUG	SEP	OCT	NOV	DEC
New shipment	70,211	61,490	58,256	49,628	71,274	40,636
Replenishment	28,098	26,470	30,125	21,858	23,323	23,038
Exchange	44,153	44,861	48,079	44,724	45,427	39,926
Credit Note	15,361	16,448	18,638	12,756	17,887	14,688
Transfer Delivery Order	89,789	129,047	109,198	95,537	107,771	98,406
Cleaning (Tray)	6,326	7,145	11,417	9,472	9,675	7,996

Figure 1: Arrival of received products (JUL - DEC 2015)

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

The received products will be sent to five different departments. Each department has a different number of assigned workers. Long lead-time is found in the inbound logistics process due to both fluctuated orders and process times. The exchange group consumes the highest lead time at 562.22 minutes per document. While the lead time of Newshipment and Cleaning group are 413.67 and 345.96 minutes per document respectively. Moreover, the lowest utilization is 8% founded in putaway department. The receive department has the highest utilization at 58%. The purposes of this research are then to balance the number of workers in each department and reduce the average total time of inbound logistics process.

LITERATURE REVIEW

Worker allocation is one of the strategies that can improve workers' performance of the case study. There are several studies focused on worker allocation to improve operational efficiency in many businesses. For example, Spry and Lawley (2005) presented a simulation model for efficiency improvement of a pharmacy department. The objective is to reduce turnaround time so that patients get medication quickly. The results show that adding four technicians in the evening can improve performance and reduce turnaround time. Rong and Grunow (2009) studied workers' planning to manage air freight transportation by using simulation. The steps of managing and sorting products from containers (UID) were studied. The worker allocation plans were suggested to reduce the cost of workers and increase the utilization of workers. Zeng *et al.* (2012) studied to improve the emergency department at a community hospital at Lexington, Kentucky by using simulation. They found that adding one CT scanner and two nurses to take care of the patients for six rooms can reduce waiting times for the patients and utilization of the nurses. Bank and Emery (2013) studied the improvement of the operation in warehouses by simulation modelling. The simulation focused on the improvement of the storage location of goods, the picking of goods, and the allocation of resources in warehouse. The objective is to reduce average queue time of incoming lorries waiting for unloading. The results show that separation of the workers to unload the goods from more than one lorry at a time is the best method to reduce the average queue time of the lorries and also improve the resource utilization. Süer *et al.* (2013) used mathematical modelling to study workforce planning in manufacturing. The authors suggested the flexibility strategy by sharing workers between operations in Cellular Manufacturing in order to maximize output rates and minimize total tardiness.

It can be seen that many authors are interested in exploring simulation tools and techniques to improve worker allocation. The method of each research is adding the workers, resources and sharing the workers together towards some key performance measures such as time, workload and cost. This research is then constructed from a current situation and concerns both time and workload or utilization. The techniques investigated are sharing or allocating the workers in each product group and balancing workload of workers. The next part is research methodology that is used to conduct this research.

RESEARCH METHODOLOGY

Discrete event simulation, ARENA, is applied as a tool to study for this research. Inbound logistics process is selected to study because of the different amount of incoming products for each day. Two types of receiving products are studied that are Goods Receive and Goods Return.

1) Goods Receive

1.1 Newshipment

The inbound logistics process for Newshipment group will start from creating queue, checking products (batch, code, the amount of product, expired date and physical conditions), placing sticker, placing barcode UID, placing both sticker and barcode or not placing anything, creating PO or non-creating PO, creating inbound document (using serial or non-using) and dispatching. It can be written as a flow chart shown in figure 2.

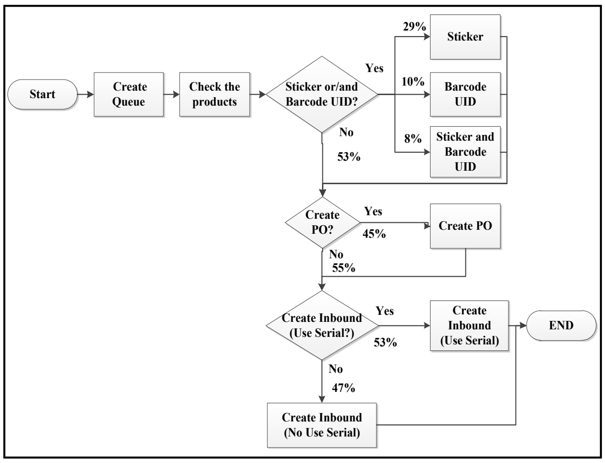


Figure 2: Newshipment process

1.2 Replenishment

The inbound logistics process of the Replenishment product group will start from creating queue, checking products (batch, code, the amount of product, expired date and physical conditions), and dispatching. The replenishment process is shown in figure 3.

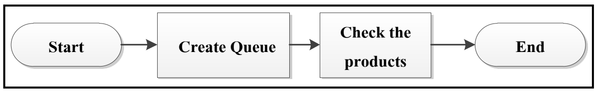


Figure 3: Replenishment process

2) Goods Return

2.1 Credit note

The inbound logistics process of Credit note group will start from creating request documents, receiving products from the customers and returning them to the Distribution Center. When the product is back to Distribution Center, it will start from checking product process (batch, code, amount of product (actual), expired date and physical conditions), keying in received document into the computer system and dispatching. Figure 4 shows the credit note steps.

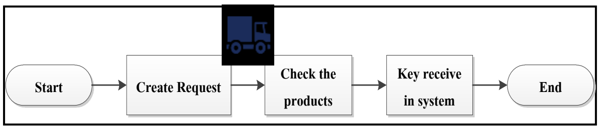


Figure 4: Credit note process

2.2 Exchange

The inbound logistics process of Exchange group is the process to exchange the products within the same product group at the same price and quantity. The process will start from checking the products (batch, code, amount of product, expired date and physical conditions), keying in received document into the computer system (divided into 3 clients; client A, client BR and client W due to different ways), sorting and dispatching. The exchange process is concluded in figure 5.

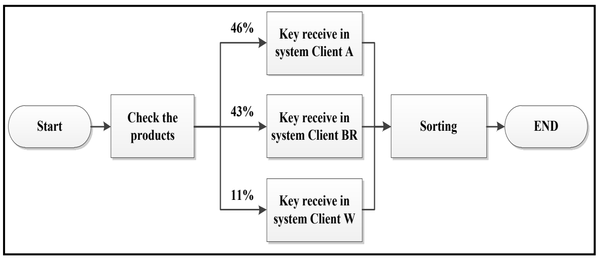

Start
Check the products
46%
Key receive in system Client A
43%
Key receive in system Client BR
11%
Key receive in system Client W
Sorting
END

Figure 5: Exchange process

2.3 Transfer Delivery Order

The inbound logistics process of Transfer Delivery Order group will start from scanning the products returned, checking the products (amount of product and physical conditions), keying in received document into the computer system (divided into 3 clients; client B, client BB and client S due to different ways) and dispatching. The Tranfer delivery order process is shown in figure 6.

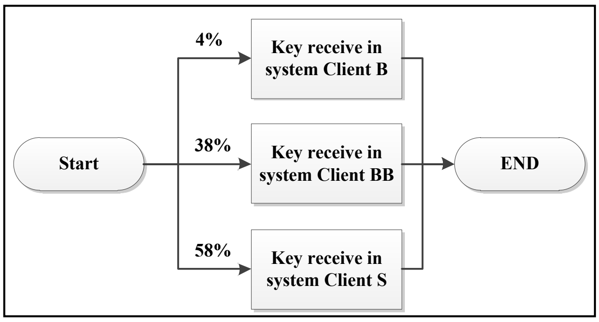

Start
4%
Key receive in system Client B
38%
Key receive in system Client BB
58%
Key receive in system Client S
END

Figure 6: Transfer Delivery Order process

2.4 Cleaning

Cleaning group is divided into 2 clients that are client BB and client S. The inbound logistics process for the two clients is similar. It will start from checking the products (amount of products and physical products), keying in received document into the computer system (before washing), washing the product by machine, key receiving document into the computer system (after washing) and dispatching. The cleaning process is described in figure 7.

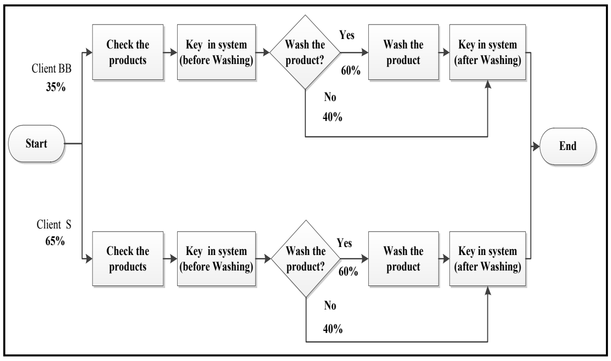


Figure 7: Cleaning process

The inbound logistics process is divided into 5 departments. Each department will have responsibility for each product group as shown in table 1. The priority of the process when the product come in together, the workers in the receiving department must do their works by creating inbound (newshipment) and creating queue (newshipment and replenishment). Creating queue is high priority because it spends a short time in working. In the part of the return department, product exchanging is high priority. The process of working is more complex than that of credit note. For the putaway department, the product for replenishment is high priority. The workers do their work 8 hours per day from 8.00 a.m. to 17.00 p.m. with a one-hour lunch break at 12 a.m. For a part-time job, they start working from 18.00 p.m. to 20.00 p.m. from Monday to Saturday (6 days per week).

Table 1: Tasks and number of workers

Department	Tasks	Number of worker
1.Receive	Check the products, Sticker and Barcode UID (Newshipment)	1
	Create Inbound (Newshipment) and Create Queue (Newshipment and Replenishment)	2
2.Return	Check the products (Credit note and Exchange)	2
	Key receiving document into system (Credit note and Exchange)	4
3.Putaway	Check the products (Replenishment) and Sorting (Exchange)	4
4.TDO-IN	Receive the products (Transfer Delivery Order)	4
5.Cleaning	Receive the products Client S (Cleaning)	2
	Receive the products Client B (Cleaning)	2

Simulation

The simulation model of inbound logistics process from creating queue until printing the putaway report was constructed. The information which is used to conduct the simulation is from October – December 2015. The data is divided into 2 parts. Firstly, the primary data is gathered by interviewing the specialists of the case study and the processing time records. In each process, it has different time units such as time per document, time per line item and time per piece. The secondary data is obtained from the company' s databases that are the arrival rate of each product group, the quantity of products and the number of line items. The data will be analyzed by using Input Analyzer to find the best distribution fit. The example of the arrival rate and the process time of Newshipment group is shown in table 2-3. Both the arrival rate and the process time of each product group are controlled by Common Random Numbers (CRN) to reduce the variation of the data between replications.

Table 2**:** Arrival rate of Newshipment

Period	Arrival rate of Newshipment (Document)
8 p.m – 9 p.m	DISC (0.722, 0, 0.861, 1, 0.937, 2, 0.962, 3, 1.000,4)
9 p.m – 10 p.m	DISC (0.747, 0, 0.785, 1, 0.873, 2, 0.911, 3, 0.949, 4, 0.987, 6, 1.000, 8)
10 p.m – 11 p.m	DISC (0.671, 0, 0.899, 1, 0.962, 2, 1.000, 3)
11 p.m – 12 p.m	DISC (0.772, 0, 0.899, 1, 0.924, 2, 0.962, 3, 0.987,4, 1.000, 8)
12 p.m – 13 a.m	DISC (0.810, 0, 0.949, 1, 0.975, 2, 0.987, 4, 1.000, 5)
13 a.m – 14 a.m	DISC (0.747, 0, 0.924, 1, 0.949, 2, 0.975, 3, 0.987,5, 1.000, 6)
14 a.m – 15 a.m	DISC (0.418, 0, 0.595, 1, 0.747, 2, 0.848, 3, 0.873,4, 0.899, 5, 0.924, 7, 0.937, 8, 0.962, 9, 0.987, 10, 1.000, 11)
15 a.m – 16 a.m	DISC (0.696, 0, 0.797, 1, 0.886, 2, 0.924, 3, 0.937, 4,0.949, 5, 0.962, 6, 0.975, 7, 1.000, 8)
16 a.m – 17 a.m	DISC (0.633,0 , 0.823,1 ,0.911 ,2 ,0.949 , 3, 0.975,5, 1.000, 7)
17 a.m – 8 p.m	0 (None)

Table 3: Distribution of process time (Newshipment)

Operation	Distribution of process time
Create Queue	18 + 78 * BETA(1.17, 1.05)(Sec /Doc)
Check the products	TRIA(4.28, 6.64, 9) (Sec /Piece)
Sticker	TRIA(7.15, 9.65, 14.9)(Sec /Piece)
Barcode UID	TRIA(5.03, 6.66, 8.91)(Sec /Piece)
Sticker & Barcode UID	TRIA(14, 15.5, 17.9)(Sec /Piece)
Create PO	TRIA(3.17, 5.19, 9.92)(Sec /Line item)
Create Inbound(serial)	TRIA(13, 17.6, 23)(Sec /Piece)
Create Inbound(No serial)	18 + 16 * BETA(1.33, 1.27)(Sec /Line item)

This study does not include warehouse management tools such as scanners and computers as those tools are sufficient. Resources of the study are then mainly workers.

Verification and Validation

Verfication has been conducted for each module in order to assure that the simulation program can perform as required. The average documents per month of 5 product groups are used to validate with real life performances. It was found that the document of each product group obtained by simulation is within the minimum and maximum of actual performances as shown in table 4.

Table 4: Output Validation

Product	Actual (Doc/Month)			Simulation (Doc/Month)		
	Average	Min	Max	Average	Min	Max
Newshipment	166	126	250	239.32	233.83	244.81
Replenishment	113	92	138	129.16	124.48	133.84
Credit Note	2787	1888	3433	2099.39	1997.10	2201.68
Exchange	3034	2244	3733	2449.31	2344.61	2554.01
Transfer Delivery Order	1922	1214	2747	2111.57	2023.51	2199.63
Cleaning	2910	2453	3223	2699.45	2591.71	2807.19

To identify run length and run replication, equation 1 was used (Kelton et al. 2007)

$$n \cong n_0 \frac{h_0^2}{h^2} \qquad (1)$$

The expected number of replications is represented by n. The desired half width is h and n_0 represents the number of replications for the pilot run. The half width obtained from the pilot run is h_0. The simulation is conducted based on terminating system as the case study starts working from Create Queue and finishing printing putaway. The number of replications is 260 and run length is 26 days.

RESULTS

Current situation

The As-is process of individual department has been simulated. It was found that the longest lead time in product groups is the Exchange group. The exchange group consumes the highest average total time at 562.22 minutes per document. It is followed by that of Newshipment and Cleaning group at 413.67 and 345.96 minutes per document respectively. The lowest average total time is found in the Replenishment group at 30.49

minutes per document. Table 5 shows the average total time for each product group.

Table 5: Current Total Time per Document

Product group	Average total time (Min/Document)	
	Average	Half-Width
Newshipment	413.67	19.90
Replenishment	30.49	2.07
Credit note	167.16	19.60
Exchange	562.22	48.62
Transfer Delivery Order	105.20	16.50
Cleaning	345.96	32.21

In each process of the product groups it was found that washing for client S has the highest average total waiting time at 394.72 minutes per document. In the meantime, the checking products of Exchange, Credit note and Newshipment group have the average total waiting time at 101.13, 164.03 and 162.01 minutes per document, relatively. Besides this, the process of Sticker, Barcode UID and Sticker & Barcode UID have the average total waiting time at 195.93, 171.96 and 146.21 minutes per document, consecutively. The key receiving document process of Credit note group is 0.76 minutes per document. More details are shown in table 6.

Table 6: Current total wait time for each process

Product group	Operation	Average total wait time (Min/Document)	
		Average	Half-Width
New shipment	Create Queue	26.67	1.52
	Check the products	101.13	10.60
	Sticker	171.96	13.45
	Barcode UID	146.21	13.43
	Sticker & Barcode UID	195.93	16.91
	Create PO	42.28	2.25
	Create Inbound(serial)	123.32	3.77
	Create Inbound(No serial)	37.07	2.20
Exchange	Check the products	164.03	23.06
	Key receiving document A	3.42	0.30
	Key receiving document BR	2.04	0.19
	Key receiving document W	3.54	0.38
	Sorting	23.90	1.29
Credit note	Create Request	20.66	4.73
	Check the products	162.01	19.59
	Key receiving document	0.76	0.07
Replenishment	Create Queue	18.51	2.02
	Check the products	0.83	0.19
Transfer Delivery Order	Receiving product B	84.38	16.78
	Receiving product BB	76.23	16.19
	Receiving product S	100.54	16.72
Cleaning	Check the products B	20.13	1.82
	Key receiving document B(before)	9.80	1.84
	Washing B	58.92	5.29
	Key receiving document B(after)	64.73	3.52
	Check the products S	66.64	8.79
	Key receiving document S(before)	66.31	8.79
	Washing S	394.72	92.48
	Key receiving document S(after)	79.37	2.15

The result of the current workload leads to imbalance of worker utilizations. The putaway department has the lowest average utilization about 8%. For the receive department (Worker B and C), worker utilizations are high at 58%. Other worker utilizations are shown in table 7.

Table 7: Current worker utilization

Worker	Utilization (%)	
	Average	Half-Width
Receive Worker A	41.75	1
Receive Worker B	58.86	1
Receive Worker C	58.16	1
Return Worker D	46.97	1
Return Worker E	46.97	1
Return Worker F	18.37	1
Return Worker G	18.38	1
Return Worker H	18.32	1
Return Worker I	32.48	1
Putaway Worker J	8.55	0
Putaway Worker K	8.52	0
Putaway Worker L	8.54	0
Putaway Worker M	8.50	0
TDO Worker N	49.73	2
TDO Worker O	49.82	2
TDO Worker P	49.66	2
TDO Worker Q	49.81	2
Cleaning B Worker R	22.58	1
Cleaning B Worker S	22.52	1
Cleaning S Worker T	48.05	2
Cleaning S Worker U	48.01	2

Alternative worker allocation strategies

From the current situation, it was found that the average total time of each product group is signicantly different. The waiting time of each process is high and the workload of workers has fluctuated greatly. The promising method to gain a better performance on the inbound logistics process is to adjust the worker allocation plan. Thus, two guidelines were proposed.

1. To manage group of the workers who are working the same tasks together.
2. To adjust workers to balance workers' utilization.

The first alternative

- Grouping workers in checking product; receive department (A) and return department (D,E) together for checking in newshipment, exchange and credit note because their works are similar. To move 1 worker from putaway (M) to work checking product.
- Reducing the workers in putaway to three workers (J,K,L) and get them help in receive for placing stickers and bacode UID (when they are available).
- Adding 3 workers (F,G,H) to create request documents.
- Adding 1 worker (S) who moves from cleaning department works in Transfer Delivery Order group.
- Grouping workers in cleaning department (client B (R) and client S (T,U)) work and use cleaning machines together in cleaning process.

The second alternative

- Grouping workers in checking product; receive department (A), return department (D,E) and replenishment department (J,K,L) together for checking in newshipment, exchange, credit note and replenishment. Beside this, they place stickers and bacode UID.
- Reducing the workers in putaway to three workers (J,K,L).
- Adding 3 workers (F,G,H) to create request documents.
- Adding 1 worker (M) who moves from putaway department works in Transfer Delivery Order group.
- Grouping workers in cleaning department (client B (R,S) and client S (T,U)) work and use cleaning machines together in cleaning process.

The conclusion of improvement is shown in table 8.

Table 8: Current and Alternatives worker allocation

Product group	Operation	As-Is	To-be 1	To-be 2
New shipment	Create Queue	B C	B C	B C
	Check the products	A B C	A D E M	A D E J K L
	Sticker		B C J K L	
	Barcode UID			
	Sticker and Barcode UID			
	Create PO	B C	B C	B C
	Create Inbound(serial)			
	Create Inbound(No serial)			
Exchange	Check the products	D E	A D E M	A D E J K L
	Key receiving document A	F G H I	F G H I	F G H I
	Key receiving document BR			
	Key receiving document W			
	Sorting	J K L M	J K L	J K L
Credit note	Create Request	I	F G H I	F G H I
	Check the products	D E	A D E M	A D E J K L
	Key receiving document	F G H I	F G H I	F G H I
Replenishment	Create Queue	B C	B C	A D E J K L
	Check the products	J K L M	J K L	J K L
Transfer Delivery Order	Receiving product B	O N P Q	O N P Q U	O N P Q M
	Receiving product BB			
	Receiving product S			
Cleaning	Check the products S	R S Mc-S	R S T Mc-S Mc-B	R S T U Mc-S Mc-B
	Key receiving document S (before)			
	Key receiving document S (after)			
	Check the products B	T U Mc-B		
	Key receiving document B (before)			
	Key receiving document B (after)			

From the comparison of two alternatives with the current situation, we can reduce the average total lead time in every product groups. However, the second alternative can perform better than the others. Most of the product groups can have a shorter lead time when alternative 2 was employed . It is shown in table 9. For the workers' utilization, both of the alternatives can improve and balance the workes' utilization better than the current situation.

Table 9: Average total lead time

Product group	Average total lead time (Min/ document)		
	As-Is	To-Be 1	To-Be 2
	Average	Average	Average
Newshipment	413.67	315.50	290.01
Replenishment	30.49	31.06	29.76
Credit note	167.16	59.29	56.71
Exchange	562.22	378.94	389.35
Transfer Delivery Order	105.20	67.82	68.19
Cleaning	345.96	300.49	249.95

For the balancing of workers, the second alternative is seemed to be better than the first alternative as the utilizations are ranged from 30% to 40%, while those of the first alternative are more variable. The details are shown in table 10.

Table 10: Worker Workload of Current and Alternative

Worker	Workload (%)		
	As-Is	To-Be 1	To-Be 2
	Average	Average	Average
Receive Worker A	41.75	35.56	34.02
Receive Worker B	58.86	43.73	38.58
Receive Worker C	58.16	42.96	38.20
Return Worker D	46.97	35.66	34.53
Return Worker E	46.97	35.54	34.45
Return Worker F	18.37	21.89	21.90
Return Worker G	18.38	21.86	21.89
Return Worker H	18.32	21.82	21.88
Return Worker I	32.48	21.83	21.84
Putaway Worker J	8.55	21.54	37.95
Putaway Worker K	8.52	21.18	37.76
Putaway Worker L	8.54	20.75	37.28
Putaway Worker M	8.50	35.38	39.53
TDO Worker N	49.73	39.82	39.83
TDO Worker O	49.82	39.91	39.56
TDO Worker P	49.66	39.83	39.83
TDO Worker Q	49.81	39.68	39.67
Cleaning BWorker R	22.58	46.79	35.27
Cleaning BWorker S	22.52	39.64	35.19
Cleaning SWorker T	48.05	46.74	35.12
Cleaning SWorker U	48.01	46.56	35.15

CONCLUSION

From the results, it can be concluded that the second alternative is the best alternative as it can reduce the average total lead time for every product groups. The highest reduction on the average total lead time is the Exchange group. The average total lead time can be reduced to 176.87 minutes per document (30.75%). It was followed by the Newshipment group as the average total lead time can be shrinked to 123.66 minutes per document(29.89%). The Replenishment group can minimize the average total lead time to 0.73 minutes per document (2.39%). For the workers' workload, the overall of utilizations are about 30%-40%. Although the alternatives are conducted by grouping the workers who are working the same tasks or processes together, each process such as credit note and replenishment can be different regarding to client's requirement. On The Job training is necessary for every process to assure that the workers can successfully perform their work for every clint. In addition, the alternatives suggested may dissatisfy the workers who have to work more. The executives may need to add some incentives by evaluating performance on the total output of the inbound logistics process. If the teamworks can reduce the average total inbound process time; the productivity of the inbound process will be increased. The teamworks will then get more incentives. Then, collaboration of the workers can achieve win-win situation.

REFERENCES

Bank, A.A. and Emery. 2013. "Using Spatial Simulation Modeling to Improve Warehouse – Logistics Operations Management." *International Conference on Control, Engineering & Information on Technology*, Vol. 1, 47-53.

Kelton, W.D.; R.P. Sadowski. and N.B. Zupick. 2015, *Simulation with Arena, 6th Edition.*, McGraw-Hill, New York.

Rong, A. and Grunow, M. 2009. "Shift designs for Freight Handling Personnel at Air Cargo Terminals". *Transportation Research*, Part E, Vol. 45,725-739.

Spry, C.W. and Lawley, M.A. 2005. "Evaluating hospital pharmacy staffing and work scheduling using simulation." *Proceedings of the 2005 Winter Simulation Conference*, 2256-2263.

Süer, G.A., Kamat, K., Mese, E. and Huang, J. 2013. Minimizing total tardiness subject to manpower restriction in labor-intensive manufacturing cells. *Mathematical and Computer Modelling*, 57, 741 -753.

Zeng, Z., Ma, X., Hu, Y., Li, J. and Bryant, D. 2013. "A Simulation Study to Improve Quality of care in the Emergency Department of a Community Hospital." *Journal of Emergency Nursing*, Vol. 38, No. 4(Jul), 322-328.

AUTHOR BIOGRAPHIES

KITTIKHUN IAMSAMAI obtained a Master degree of Science in Logistics Management at King Mongkut's Unversity of Technology Thonburi, Thailand. His e-mail address is: kittikhun.iam@mail.kmutt.ac.th

THANANY WASUSRI graduated with a PhD from the University of Nottingham, England in the field of manufacturing engineering and operations management. She is working at King Mongkut's University of Technology Thonburi. Her research of interest is logistics management, supply chain management and inventory management whilescienctific tools are simulation, optimization and multivariate statisitcs. Her e-mail address is : thananya.was@kmutt.ac.th

SIMULATION OF A QUEUEING MODEL USEFUL IN CROWDSOURCING

Srinivas R. Chakravarthy
Department of Industrial and
Manufacturing Engineering
Kettering University
MI 48504, Flint, USA
E-Mail: schakrav@kettering.edu

Serife Ozkar
Department of Statistics
Istanbul Medeniyet University
Istanbul, Turkey
E-mail: serife.ozkar@medeniyet.edu.tr

KEYWORDS
Crowdsourcing, *MAP* arrivals, phase type services, multi-server queue models, simulation.

ABSTRACT

In this paper we study a multi-server queueing model in the context of crowdsourcing useful in service sectors. There are two types of arrivals to the system such that one group of customers after getting service from the system may serve (with a certain probability) another group. Through simulation we point out, even for a small value of this probability, the significant advantage in considering crowdsourcing by offering more traffic load (through increasing the rate of online customers without violating the stability condition) to the system resulting in more customers served (which in turn increasing the revenues when cost/profits are incorporated). Also, the role of correlation, especially a positive one, present in the inter-arrival times in the system performance measures is highlighted.

INTRODUCTION

The concept of crowdsourcing has been used in different domains gaining significant exposure in many service sectors. We refer the reader to a recent survey paper by (Hosseini et al. 2015 and Evans et al. 2016). The meaning and interpretation of crowdsourcing is varied and despite its popularity among companies in many sectors it remains little understood. We refer the reader to (Howe 2008) for a number of examples related to crowdsourcing in various sectors.

The literature on the quantitative analysis of the crowdsourcing with the help of mathematical models is very small even though the literature on qualitative nature of crowdsourcing dealing with various definitions and classification is huge. The quantitative models for crowdsourcing will benefit business and service industries to better understand the system when underlying parameters change. For example, with the help of survival analysis and using the dataset from MTurk, (Wang et al. 2011) analyzed the completion time of crowdsourcing campaigns. The material flow of crowdsourcing processes in manufacturing systems was studied by using stochastic Petri nets in (Wu et al. 2014). Only recently stochastic models, more specifically queueing models, useful in crowdsourcing in the context of service sectors have been studied. To the best of our knowledge the first queueing model using one type of customers as possible servers for another group was studied by (Chakravarthy and Dudin 2017). The authors studied a queueing with corowdsourcing of *M/M/c*–type using matrix-analytic methods.

It should be pointed out that the model studied in (Bernstein et al. 2012) deals with retaining a select few workers as "backup" servers on call for helping the system. These workers are allowed to tend to other tasks until a request for their help is made by the system. Thus, in their model the customers arriving at the system are never considered as servers for the system.

In this paper, we generalize the model studied in (Chakravarthy and Dudin 2017) by considering a more versatile point process, namely, Markovian arrival process (*MAP*) for Type 1 arrivals, Poisson arrivals for Type 2, and phase type (*PH*-distribution) for services. That is, we study queueing model of *MAP/PH/c*–type with crowdsourcing and resort to simulation for the analysis since the state space of the queueing model grows exponentially with the number of servers, phase of the arrival, and the phase of the service processes.

It is well-known (Neuts 1975) that a PH-distribution is obtained as the time until absorption in a finite-state Markov chain with an absorbing state. Realizing the limitations of Poisson processes and exponential distributions in spite of their nice mathematical properties, Neuts (Neuts 1979) first developed the theory of phase type distributions and *MAPs*. The *MAP* is a rich class of point processes that not only generalize many well-known processes such as Poisson, *PH*-renewal processes, and Markov-modulated Poisson process but also provides a way to model *correlated arrivals.* For further details on *MAP* and their usefulness in stochastic modelling, we refer to (Lucantoni et al. 1990; Lucantoni 1991; Neuts 1992) and for a review and recent work on *MAP* we refer the reader to (Artalejo et al. 2010; Chakravarthy 2001; Chakravarthy 2010).

MODEL DESCRIPTION

We consider a *c*- server queueing system in which two types, say, Type 1 and Type 2, of customers arrive. We assume that Type 1 customers arrive according to a *MAP* with representation (D_0, D_1) of order m. An arriving Type 1 customer finding the server idle will get into service immediately. Otherwise, the customer will enter a finite

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

buffer of size L, $1 \leq L < \infty$, to be served on a First-Come-First-Served ($FCFS$) basis when the server becomes free. Thus, it is possible for a Type 1 customer to be lost at the time of arrival due to the buffer being full. Let D, defined by $D = D_0 + D_1$, govern the underlying Markov chain of the *MAP* such that D_0 accounts for the transitions corresponding to no arrival; D_1 governs those corresponding to an arrival of a Type 1 customer. By assuming D_0 to be a nonsingular matrix, the interarrival times will be finite with probability one and the arrival process does not terminate. Hence, we see that D_0 is a stable matrix. Let λ_1 denote the arrival rate of Type 1 customers. The arrivals of Type 2 customers are assumed to follow a Poisson process with rate λ_2. There is no restriction on how many Type 2 customers can be in the system. That is, there is an infinite buffer space for Type 2 customers.

While Type 1 customers are to be served by one of the *c* servers, Type 2 customers may be served by a Type 1 customer having already been served and also available to act as a server or by one of the *c* (system) servers. For example, Type 1 customers visit the store to buy items while Type 2 customers order over some medium such as Internet and phone, and expects them to be delivered. The store management can use the in-store customers as couriers to "serve" the other type of customers. Not all in-store customers may be willing and in some cases not possible to act as servers for the store. Hence, a probability is introduced for Type 1 customers to opt for servicing Type 2 customers.

A Type 2 customer getting serviced by a Type 1 customer depends on the following conditions. First, that Type 1 customer should have just finished getting a service and opts to service a Type 2 customer. Secondly, at the time of opting to serve there is at least one Type 2 customer waiting to get a service. We assume that a served Type 1 customer will be available to act as a server for a Type 2 customer under the conditions mentioned above with probability p, $0 \leq p \leq 1$. With probability $q = 1 - p$, the served Type 1 customer will leave the system without opting to serve a Type 2 customer. Upon completion of a service a free server will offer service to a Type 1 customer on a *FCFS* basis; however, if there are no Type 1 customers waiting, the server will serve a Type 2 customer if there is one present in the queue. If a Type 1 customer decides to serve a Type 2 customer, for our analysis purposes that Type 2 customer will be removed from the system immediately. This is due to the fact that the system no longer needs to track that Type 2 customer.

We assume that all system servers offer services to either type on a *FCFS* within the type; however, Type 1 customers have non-preemptive priority over Type 2 customers. The service times are assumed to be of phase type with representation $(\boldsymbol{\beta}, S)$ of order *n* with mean $1/\mu = \boldsymbol{\beta}(-S)^{-1}\boldsymbol{e}$, where $\boldsymbol{e}$ is a column vector of 1's of order *n* here and will be of dimension of appropriate dimension in the sequel.

The arrival rate, λ_1, is given by $\lambda_1 = \boldsymbol{\delta} D_1 \boldsymbol{e}$, where $\boldsymbol{\delta}$ is the stationary probability vector of the irreducible generator D, and is the unique (positive) probability vector satisfying $\boldsymbol{\delta} D = \boldsymbol{0}, \boldsymbol{\delta e} = 1.$

The model outlined above can be studied as a Markov process by keeping track of (a) the number, $K_1(t)$, of Type 2 customers in the queue; (b) the number, $K_2(t)$, Type 1 customers in the queue; (c) the number, $K_3(t)$, of servers busy serving Type 1 customers; (d) the number, $K_4(t)$, of servers busy with Type 2 customers; (e) the phase , $J_r(t)$, of the r^{th} server, and (f) the phase, $J(t)$, of the arrival process at time *t*. The process $\{(K_1(t), K_2(t), K_3(t), K_4(t), J_1(t), \cdots, J_{K_3(t)+K_4(t)}(t), J(t): t \geq 0)\}$ is a continuous-time Markov chain with state space given by

$$\Omega = \{(i_1, i_2, i_3, i_4, j_1, \cdots, j_c, k): i_1 \geq 0,\ \ 0 \leq i_2 \leq L,$$
$$0 \leq i_3 \leq c,\ \ 0 \leq i_4 \leq c,\ \ 0 \leq i_3 + i_4 \leq c,\ \ 0 \leq j_r \leq n$$
$$1 \leq r \leq c, 1 \leq k \leq m\}.$$

Note that we take $J_r = 0$ when the r^{th} server is idle. The generator of this Markov process can be set up with the help of Kronecker products and sums of matrices. However, it is clear that the analysis of this model analytically requires a large state space to account for all the states described above. These are currently work-in-process and the results will be reported elsewhere. However, our goal in this paper is to see how the impact of introducing crowdsourcing in the context of multi-server queueing system with *MAP* arrivals through simulation. The rest of the paper is based on simulating the crowdsourcing queueing model described here with the help of ARENA (Kelton et al. 2010). Unless otherwise mentioned, we ran our simulation models using 3 replications and for 1,000,000 units (which in our case is minutes) for each replicate.

VALIDATION OF THE SIMULATED MODEL

In any simulation work, it is important to validate the simulated model by comparing the results with any known analytical results. Hence, we will do that in this section. The only cases for which analytical results are available for the model under study are for *M/M/c* (Chakravarthy and Dudin 2017). This is due to recent interests to study crowdsourcing from queueing theory perspective.

We will list four key system performance measures among many for our illustration.

- The probability, *PLOS*, that an arriving Type 1 customer is lost due to the buffer being full.
- The probability, *PT*2*L*1, that a Type 2 customer leaves (served) with a Type 1 customer.
- The mean, *MN*1*Q*, number of Type 2 customers waiting time in the queue.
- The mean, *MN*2*Q*, number of Type 2 customers waiting time in the queue.

M/M/c crowdsourcing

In (Chakravarthy and Dudin 2017), the authors studied an *M/M/c* type queueing models with crowdsourcing. We will compare our simulated results against their numerical ones, which were generated through analytical study by employing matrix-analytic methods. Specifically, we fix $\lambda_1 = 1$, $\mu = 1.1, L = 10$, and vary other parameters as follows: $p = 0, 0.5, 1$, $\rho = 0.8, 0.99$, $c = 1, 2, 5, 10$. The values of λ_2, which depend on the (fixed) value of ρ (see (Chakravarthy and Dudin 2017)), are displayed in Table 1 below.

Table 1: Values of λ_2 for various scenarios for *M/M/c*

	$\rho = 0.8$			$\rho = 0.99$		
c	$p = 0$	$p = 0.5$	$p = 1$	$p = 0$	$p = 0.5$	$p = 1$
1	0.1232	0.5016	0.8800	0.1524	0.6207	1.0890
2	0.9602	1.3601	1.7600	1.1882	1.6831	2.1780
5	3.6000	4.0000	4.4000	4.4550	4.9500	5.4450
10	8.0000	8.4000	8.8000	9.9000	10.3950	10.8900

The error percentage, which is calculated as $\{|Analytical - simulated|\ /\ Analytical\}\ 100\%$, for various scenarios are displayed in Table 2.

By looking at Table 2 we notice that our simulated results are very close to the ones obtained using analytical results presented in (Chakravarthy and Dudin 2017) for all except a couple of scenarios. For these scenarios we ran the simulation again but with 10,000,000 minutes and 3 replicates and found the error percentages for these cases drop significantly.

Table 2: Error percentages (%) of analytical and simulated models for *M/M/c*

		MN1Q		*MN2Q*		*PLOS*		*PT2L1*	
c	p \ ρ	0.8	0.99	0.8	0.99	0.8	0.99	0.8	0.99
1	0	0.13	0.20	0.66	7.24	0.00	0.00	0.00	0.00
	0.5	0.42	0.22	1.04	11.98	0.39	0.93	0.11	0.13
	1	0.30	0.16	0.10	22.47	1.16	0.00	0.10	0.07
2	0	0.10	0.05	0.51	7.05	0.00	0.00	0.00	0.00
	0.5	0.17	0.07	0.88	8.68	0.00	0.00	0.00	0.03
	1	0.09	0.00	0.07	3.05	0.00	0.00	0.02	0.02
5	0	0.22	0.18	0.28	0.80	0.00	0.00	0.00	0.00
	0.5	0.07	0.00	0.43	7.02	0.00	0.00	0.12	0.10
	1	0.07	0.05	0.20	2.77	0.00	0.00	0.07	0.11
10	0	0.22	0.10	0.19	1.16	0.00	0.00	0.00	0.00
	0.5	0.00	0.00	0.12	0.59	0.00	0.00	0.00	0.21
	1	0.20	0.00	0.26	1.00	0.00	0.00	0.35	0.00

SIMULATED RESULTS FOR *MAP/PH/c* CROWDSOURCING

For the arrival process, we consider the following five sets of values for D_0 and D_1 as follows.

Erlang distribution (*ERLA*):

$$D_0 = \begin{pmatrix} -2 & 2 \\ 0 & -2 \end{pmatrix}, \quad D_1 = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}$$

The exponential distribution (*EXPA*):

$$D_0 = (-1), D_1 = (1)$$

The hyper-exponential distribution (*HEXA*):

$$D_0 = \begin{pmatrix} -1.90 & 0 \\ 0 & -0.19 \end{pmatrix}, \ D_1 = \begin{pmatrix} 1.710 & 0.190 \\ 0.171 & 0.019 \end{pmatrix}$$

The MAP with negative correlation (*MNCA*):

$$D_0 = \begin{pmatrix} -1.00222 & 1.00222 & 0 \\ 0 & -1.00222 & 0 \\ 0 & 0 & -225.75 \end{pmatrix},$$

$$D_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0.01002 & 0 & 0.9922 \\ 223.4925 & 0 & 2.2575 \end{pmatrix}$$

The MAP with positive correlation (*MPCA*):

$$D_0 = \begin{pmatrix} -1.00222 & 1.00222 & 0 \\ 0 & -1.00222 & 0 \\ 0 & 0 & -225.75 \end{pmatrix},$$

$$D_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0.9922 & 0 & 0.01002 \\ 2.2575 & 0 & 223.4925 \end{pmatrix}$$

These *MAP* processes will be normalized to have a specific arrival rate. However, these are qualitatively different in that they have different variance and correlation structure. The first three arrival processes, namely, *ERLA*, *EXPA*, and *HEXA*, have zero correlation for two successive inter-arrival times. The arrival processes labeled *MNCA* and *MPCA*, respectively, have negative and positive correlation for two successive inter-arrival times with values -0.4889 and 0.4889. The ratio of the standard deviation of the inter-arrival times of these five arrival processes with respect to *ERLA* are, respectively, 1, 1.41421, 3.17450, 1.99335, and 1.99335.

For the service times $(\boldsymbol{\beta}, S)$ we consider the following three PH-distributions.

The Erlang distribution (*ERLS*):

$$\beta = (1 \quad 0), S = \begin{pmatrix} -2 & 2 \\ 0 & -2 \end{pmatrix}$$

The exponential distribution (*EXPS*):

$$\beta = (1), S = (-1)$$

The hyper-exponential distribution (*HEXS*):

$$\beta = (0.9 \quad 0.1), S = \begin{pmatrix} -1.90 & 0 \\ 0 & -0.19 \end{pmatrix}$$

Notice that these three *PH*-distributions all have the same mean but are qualitatively different in that the variations in the distributions are different. The ratio of the standard deviation of these three service times with respect to

ERLS are, respectively, 1, 1.41422, and 3.17454. These distributions will be appropriately normalized to attain a specific mean in the numerical examples.

Having validated our simulated crowdsourcing model for known cases in the previous section, we will now present a few illustrative examples to bring out qualitative nature of *MAP/PH/c* crowdsourcing model under study. We will discuss three examples here. In the sequel, we let Y denote the waiting time in the system of a Type 2 customer. In addition to the measures listed in the validation section we consider the following ones.

- The probability, *PIDL*, the system is idle.
- The probability, *PBUS*1, that the system is busy serving Type 1 customers.
- The probability, *PBUS*2, that the system is busy serving Type 2 customers.
- The mean, *MWTS*$_1$, waiting time in the system of Type 2 customer.
- The mean, *MWTS*$_2$, waiting time in the system of Type 2 customer.
- The fraction, *FATH*, of Type 2 customers whose waiting time in the system exceeds r, $r \geq 2$, times the average service time by one of the system servers. That is, $P\left(Y > \frac{r}{\mu}\right), r \geq 2$. Since there is no analytical expression available for the measure, *FATH*, dealing with a specific tail probability of the waiting time in the system of a Type 2 customer, we used the simulated result instead. It should be pointed out that one can compute algorithmically the tail probability for classical single-server model, *MAP/PH/1* using the matrix-analytic methods (Neuts 1981) but for a multi-server case it is highly complicated and hence we resort to simulation only for this particular measure.

The purpose of our next example is to investigate the level of such effect *MAP/PH/c* case using simulated results. We will look at the ratio of a few of the measures under study here. The ratio, $\frac{\eta(p>0)}{\eta(p=0)}$, will be of interest for a given measure η.

EXAMPLE 1: The effect of crowdsourcing is studied in this example by comparing the models: (a) $p = 0$ that corresponds to having two independent arrival processes and (b) $p > 0$ that corresponds to the crowdsourcing model in the context of a multi-server system. In the latter case there is a possibility for Type 1 customers to act as servers for Type 2. We fix $\lambda_1 = 1, \lambda_2 = 2, \mu = 1.1, c = 3, L = 10$ and vary p on the interval $(0, 1]$ under different combinations of arrival and service distributions.

In Figure 1 below we display the ratios of all but *PBUS*1 since the ratio for *PBUS*1 for all scenarios are almost 1. Some key observations are summarized as follows. First observe that the smaller the ratio the better the system in terms of all measures except *PIDL* in which case it should be the larger the better. Having more idle time for the system will enable the management to use that time for other activities without having to lower the quality of service provided to the customers.

In Figure 1, we display the ratios of the key measures under different scenarios for *ERLS* and *HEXS* services. Note that for lack of space we display only selected combinations; however, our observations summarized below are valid for other combinations not displayed in this figure.

A quick look at these figures reveals the following interesting and important observations.

- The ratio for the measure, *PIDL*, is greater than 1 and increases as p is increased indicating that the server becomes idle more often when $p > 0$ when compared to that of $p = 0$. This is true for all combinations of arrivals and services.
- The ratios of all other measures are less than 1 and decrease as p is increased. The rate of decrease depends on the type of arrivals and services.
- The ratio for *MWTS*$_2$ decreases significantly as p increases. This is true for all cases. This is very important from both management's as well as customers' points of view.
- The ratio for *MWTS*$_1$ decreases as p increases but not as significantly as that of *MWTS*$_2$. This is somewhat surprising since one would expect the ratio to be close to 1 since Type 1 service is not affected by the value of p due to non-preemptive nature of services. However, as p increases, Type 1 customers have a higher probability of serving Type 2 customers resulting in relatively fewer Type 2 customers to be served by one of the system servers and hence a reduction in the mean waiting time in the system.
- The ratio for *FATH* is decreasing at a significantly higher rate as p increases (for all scenarios) which is again very important from both management as well as customers' points of view and can also be used by the management to guarantee some kind of a guarantee on the service times of Type 2 customers.

The above observations show the significant advantage in introducing this type of variants, namely, crowdsourcing, to the classical queueing models.

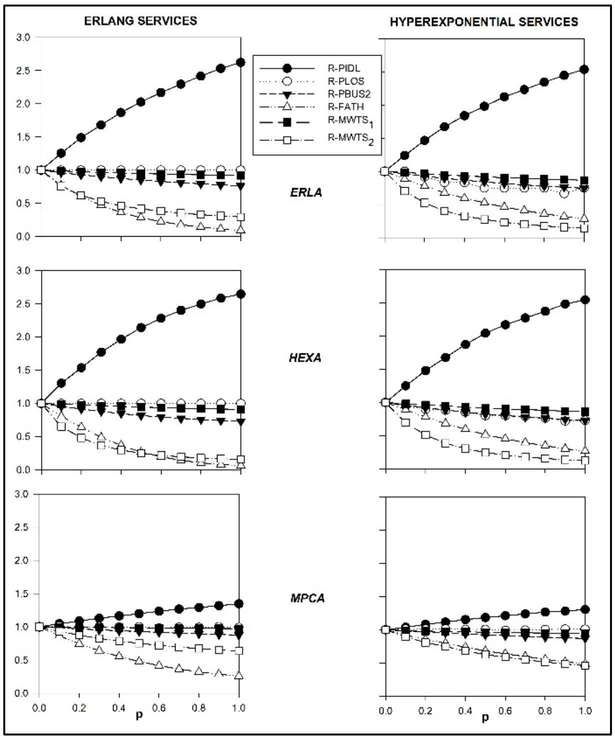


Figure 1: Ratios of various measures under different scenarios with *ERLS* and *HEXS* services

In (Chakravarthy and Dudin 2017), it was shown that even in the case of small p there is a significant advantage in considering crowdsourcing by offering more traffic load through increasing the rate of Type 2 customers into the system. The rate of increase in the offered load to the system is much higher for small values of c. Here, we will investigate a similar advantage from a different point of view by considering the cases when $L = 0$ and $L = 1$. In the former case Type 1 customers are allowed only when at least one server is idle. Thus, the maximum number of Type 1 customers that can be present at any time in the system is c and $c + 1$, respectively.

EXAMPLE 2: This example is very similar to Example 1 except that we look at the cases when (a) $L = 0$ and (b) $L = 1$. Note that in these two cases the model can be considered as a slight variation of *M/PH/c* model since Type 2 arrivals arrive to a multi-server system with phase type arrivals and occasionally Type 1 customers are allowed to enter into the system. Hence, it will be interesting to see how having only a small number of Type 1 customers, namely, c and $c+1$ when $L = 0$ and $L = 1$, respectively, at any given time will have an impact on the selected system performance measures. Towards this end, we will fix $\lambda_1 = 1$, $\lambda_2 = 2$, $\mu = 1.1$, $c = 3$, and vary p on the interval $(0,1]$ under different combinations of arrival and service distributions. Note that the queue is stable for all combinations under these values. In order to properly compare, we now look at the ratio $\frac{\zeta(L=1)}{\zeta(L=0)}$ where $\zeta(L = r) = \frac{\eta(p>0,L=r)}{\eta(p=0,L=r)}$, $r = 0, 1$.

In Figure 2 below we display the ratios for selected measures and for representative scenarios. First observe that the smaller the ratio the better the system with $L = 1$ as compared to $L = 0$ in terms of all measures except *PIDL* in which case it should be the larger the better. Having more idle time for the system will enable the management to use that time for other activities without having to lower the quality of service provided to the customers.

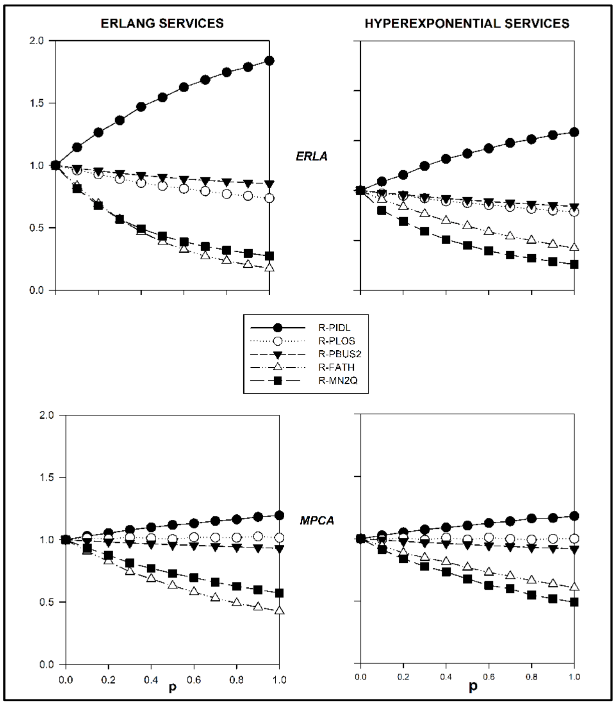


Figure 2: Comparison of the ratio under different scenarios with *ERLS* and *HEXS* services

A quick look at this Figure reveals a very significant advantage of having extra Type 1 customers in the system. This is the case for all scenarios (even the ones not displayed here for lack of space) and for all values of p. So, having Type 1 customers even if they are willing to offer services rarely plays a significant role in crowdsourcing applications.

In the next example, we try to find the optimum c^* such that the proportion of Type 2 customers whose waiting time in the system exceeds r, $r \geq 2$, times the average service time by one of the system servers does not exceed a pre-determined value, such as 5%. Recalling that Y denotes the waiting time in the system of a Type 2 customer, then for any $c \geq c^*$ when all other parameters are fixed, the following holds good.

$$P\left(Y > \frac{r}{\mu}\right) < 0.05, r \geq 2.$$

The purpose of this is to identify the regions where for a given p the minimum value of c that will guarantee that only certain (pre-determined) percentage of Type 2 customers has longer than a (pre-determined) multiple of the average service time. Similarly, we can fix c and identify if there is any p that will yield a similar guarantee.

EXAMPLE 3: Here we fix $\lambda_1 = 1$, $\mu = 1.1$, $L = 10$, and vary other parameters as follows: $\rho = 0.8, 0.9$, $p = 0, 0.5, 1$.

In Table 3 we display the optimum c^* for various combinations. We ran our simulation for c up to 50 and if an optimum is not found in that range, we will denote this by simply displaying with "> 50".

Table 3: Optimum c^* values

		ERLS			*EXPS*			*HEXS*		
ρ	*MAP*	*p*=0	*p*=0.5	*p*=1	*p*=0	*p*=0.5	*p*=1	*p*=0	*p*=0.5	*p*=1
0.8	*ERLA*	3	2	2	4	4	3	>50	>50	>50
	EXPA	3	2	2	5	4	3	>50	>50	>50
	HEXA	3	3	2	5	4	3	>50	>50	>50
	MNCA	3	2	2	5	4	3	>50	>50	>50
	MPCA	3	2	2	20	4	3	>50	>50	>50
0.99	*ERLA*	3	3	2	4	4	3	>50	>50	>50
	EXPA	3	3	2	10	4	3	>50	>50	>50
	HEXA	3	3	2	5	4	3	>50	>50	>50
	MNCA	3	3	2	4	4	3	>50	>50	>50
	MPCA	3	3	2	20	4	3	>50	>50	>50

A quick look at this table suggests that for *HEXS*, which has a large variation compared to the other two service distributions, one needs c to be larger than 50 for all values of p. Only in the case of positively correlated arrivals and with non-Erlang service times we see a relatively large c when Type 1 customers are not willing to serve Type 2 customers (i.e., when $p = 0$) which is not surprising as positively correlated arrivals are known to show such "odd" behavior with regard to other system performance measures in the literature (see e.g., (Chakravarthy 2010)).

CONCLUSION

In this paper we considered a queueing system useful in crowdsourcing. Specifically, we considered a multi-server queueing model in which one type of customers may be available to serve another type of customers leading to more efficiency as well as to help the management to increase their productivity and hence revenues. Through illustrative numerical examples obtained via simulation, we showed the significant benefits in introducing this type of variants to the classical queueing models.

Even for the small value of the probability introduced for in-store customers to opt for servicing the other type, we point out the significant advantage in considering crowdsourcing by offering more traffic load (through increased the rate of online customers without violating the stability condition) to the system resulting in more customers served (which in turn increasing the revenues when cost/profits are incorporated). Also, the role of correlation, especially a positive one, present in the inter-arrival times in the system performance measures is highlighted.

The model considered in this paper can be improved as follows. The assumption that Type 2 customers may be served singly by Type 1 customers can be relaxed to include batch services.

REFERENCES

Artalejo, J.R.; A. Gomez-Correl; and Q.M. He. 2010. “Markovian arrivals in stochastic modelling: a survey and some new results.” *SORT*, 34(2), 101-144.

Bernstein, M.S.; D.R. Karger; R.C. Miller; and J. Brandt. 2012. Analytic methods for optimizing realtime crowdsourcing. *Proceeding of Collective Intelligence Conference* held at MIT.

Chakravarthy, S.R. 2001. “The batch Markovian arrival process: A review and future work.” Advances in Probability Theory and Stochastic Processes. A. Krishnamoorthy, N. Raju, and V. Ramaswami. (Eds.) Notable Publications Inc. NJ, 21-39.

Chakravarthy, S.R. 2010. “Markovian arrival processes.” *Wiley Encyclopedia of Operations Research and Management Science.*

Chakravarthy, S.R. and A.N. Dudin. 2017. “A Queueing Model for Crowdsourcing.” *Journal of Operational Research Society,* 68(3), 221-236. doi:10.1057/s41274-016-0099-x.

Evans, R.D.; J.X. Gao; S. Mahdikhah; M. Messaadia; and D. Baudry. 2016. “A review of crowdsourcing literature related to the manufacturing industry.” *Journal of Advanced Management Science,* 4(3), 224-231.

Hosseini, M.; A. Shahri; K. Phalp; J. Taylor; and R. Ali. 2015. “Crowdsourcing: A taxonomy and systematic mapping study.” *Computer Science Review*, 17, 43-69.

Howe, J. 2008. *Crowdsourcing- Why the power of the crowd is driving the future of business*. Three Rivers Press, New York.

Kelton, W.; R. Sadowski; and N. Swets. 2010. *Simulation with ARENA*. Fifth ed. McGraw-Hill, New York.

Lucantoni, D.; K.S. Meier-Hellstern; and M.F. Neuts. 1990. “A single-server queue with server vacations and a class of nonrenewal arrival processes.” *Advances in Applied Probability*, 22, 676-705.

Lucantoni, D.M. 1991. “New results on the single server queue with a batch Markovian arrival process.” *Stochastic Models*, 7, 1-46.

Neuts, M.F. 1975. “Probability distributions of phase type”, Liber Amicorum Prof. Emeritus H. Florin, Department of mathematics, University of Louvain, Belgium, 173-206.

Neuts, M.F. 1979. “A versatile Markovian point process.” *Journal of Applied Probability*, 16, 764-779.

Neuts, M.F. 1981. *Matrix-geometric solutions in stochastic models: An algorithmic approach*. The Johns Hopkins University Press, Baltimore, MD. [1994 version is Dover Edition]

Neuts, M.F. 1992. “Models based on the Markovian arrival process.” *IEICE Transactions on Communications* Vol.E75-B, No.12, 1255-1265.

Wang, J.; S. Faridani; and P.G. Ipeirotis. 2011. “Estimating the completion time of crowdsourced tasks using survival analysis models.” *Workshop on Crowdsourcing for Search and Data Mining*, Hong Kong, China.

Wu, D.; D.W. Rosen; and D. Schaefer. 2014. “Modelling and Analyzing the Material Flow of Crowdsourcing Processes in Cloud-Based Manufacturing Systems Using Stochastic Petri Nets.” *The ASME 2014 International Manufacturing Science and Engineering Conference*, June 9-13, Ann Arbor, Michigan, USA.

AUTHOR BIOGRAPHIES

SRINIVAS R. CHAKRAVARTHY is professor of Industrial Engineering and Statistics in the Departments of Industrial and Manufacturing Engineering & Mathematics at Kettering University (formerly known as GMI Engineering & Management Institute), Flint, Michigan. His research interests are in the areas of algorithmic probability, queuing, reliability, inventory, and simulation. He has published more than 100 papers in leading journals and made more than 90 presentations at national and international conferences. He co-organized the First International Conference on MAMs in Stochastic Models in 1995 held in Flint. His recognitions and awards include Rodes Professor, Kettering University, Kettering University Distinguished Research Award, Kettering University/GMI Alumni Outstanding Teaching Award, GMI Outstanding Research Award, and GMI Alumni Outstanding Teaching Award, and Educator of the Year Award by IEOM Society, 2016. Srinivas Chakravarthy has significant industrial experience by consulting with GM, FORD, PCE, and UPS. His professional activities include serving as (a) Area Editor for the journal, Simulation Modelling Theory and Practice; (b) Associate Editor for the journal IAPQR TRANSACTIONS-Indian Association for Productivity, Quality & Reliability; (c) Advisory Board Member for several other journals and International Conferences; and (d) Reviewer for many professional journals.

SERIFE OZKAR is a research assistant pursuing her doctoral studies in the Department of Statistics, Istanbul Medeniyet University. This paper was written while she was a visiting research scholar in the Department of Industrial and Manufacturing Engineering, Kettering University, Flint, Michigan. Serife Ozkar’s research interests are in the areas of queuing systems and stochastic processes.

3D SIMULATION MODELING OF APRON OPERATION IN A CONTAINER TERMINAL

Jingjing Yu
Guolei Tang*
Da Li
Baoying Mu
Faculty of Infrastructure Engineering
Dalian University of Technology
Dalian 116024, China
*E-mail: tangguolei@dlut.edu.cn

KEYWORDS

Container Terminal; Quay Crane; 3D Simulation Modeling; Operation Efficiency.

ABSTRACT

In response to the phenomenon that some container terminals are excess capacity and some others are overloaded because of the imbalance of transportation development, this paper proposed a 3D simulation model of operation system for the container terminal apron. First, we study the characteristics of operation system for container terminal apron. Second, a 3D simulation model of terminal apron operation is implemented, which includes sub-model of layout, sub-model of setting variables and parameters, sub-model of ship arriving, sub-model of loading and discharging operation, and sub-model of horizontal transport and yard operation, as well as sub-model of 3D animation. Finally, the implemented model is applied in practice to examine the impact of the time of single operation of the quay crane, the number of trucks allocated for each working path and the number of yard cranes equipped for each block on the operation efficiency of terminal apron. And the results show the proposed simulation model performs well that can provide experience and reference for exploring terminal apron effectively.

1. INTRODUCTION

As is known to all, the port enterprise is a service industry with large capital investment. From the early stage of land acquisition to construction and the purchase of large port handling machinery, every stage needs to take economy into consideration. For this reason, the operators make the best use of the equipment in the port to increase the port economic interests. However, for peak hour of ship arrival, the equipment may be inadequate for the ships, and the ships have to wait for idle equipment, which leads to the losses for both the port authority and the ship owners. On the contrary, if the port authority aims at providing a high level of service for the ship, it is inevitable to increase the number of equipment and improve production capacity, which would decrease the waiting time of the ship but increase the idle time of berths and equipment when the arriving ships are fewer. Therefore, to design and operate a successful container terminal, an effective model is needed to help planners to evaluate and explore the operation efficiency and utilization of equipment.

In the recent years, many scholars at home and abroad have made great progress in the modeling and simulation of the container terminal operating system, and developed many new modeling and simulation techniques. The foreign scholars mainly focus on the operational simulation. For example, Sun et al. (2012) introduced a general simulation platform, named MicroPort, which aims to provide an integrated and flexible modeling system for evaluating the operational capability and efficiency of different designs of seaport container terminals. And Bruzzone et al. (2013) presented an advanced high level architecture federation of simulators. The federation of simulators is used for operators' training in terminal containers. The federation includes multiple container handling equipment simulators. In addition, Azab and Eltawil (2016) developed a discrete event simulation model to study the effect of various truck arrival patterns on the truck turn time in container terminals. And the result showed the influence of the arrival patterns on the turn time of external trucks. In China, the development and application of the container terminal simulation system also received much attention and obtained a series of research and application results. Zhang et al. (2016) put forward the idea that the operation efficiency by planning the block length reasonably based on the analysis of the operation of container terminals. They built the container terminal operation simulation model based on the computer simulation technology and research the effect of block length on the operation efficiency of container terminals through simulation experiments. Ren (2011) used the software FlexSim to study the quantitative relations of container ships, cranes, berths, and container trucks Lin (2011) studied the loading and discharging process of super large container ships by using the software WITNESS. And they obtained optimization plan according to the analysis on the parameters of loading and discharging

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

process. Ji et al. (2007) developed an optimization model aimed at minimizing the operation time of trucks. And they also designed an algorithm to solve the optimization model and carried on digital simulation experiments to obtain the valid numeric results.

This paper presents a 3D simulation model of operations on the container terminal apron using the software AnyLogic (2017). And the model emphasizes on the analysis on the influencing factors of operation efficiency and utilization of equipment and 3D visual operation process for container terminal apron, which are the main contributions of this paper.

2. CONTAINER TERMINAL APRON OPERATION

The container terminal apron is the seaside of a terminal which involves in the loading and unloading of vessels. On arrival, a vessel docks at a free berth. Most container terminals in China use quay cranes for the (un)loading operations of containers (from) onto vessels. And internal and external trucks for the horizon transport between quay and storage yard, as well as between storage yard and landside interfaces.

As shown in Figure 1, the operation process of container terminal apron includes the ship berthing, loading and discharging, as well as the ship departing. And among these, the container loading and discharging process dominates the operation process of terminal apron. And the scheduling of terminal apron mainly includes the berth allocation and the quay crane assignment, both of which have important influence on the efficiency of the container terminal.

3. 3D SIMULATION MODEL OF CONTAINER TERMINAL APRON OPERATION

To evaluate the performance of terminal apron operation, we establish a 3D simulation model to simulate the process of terminal apron operation visually using AnyLogic software (2017).

According to the characteristics of terminal apron operation and interactions with yard operation, the simulation model proposed in this paper includes 6 sub-modules, including terminal layout and setup sub-module, ship behavior sub-module, quay crane operation sub-module, horizontal transport and yard operation sub-module and 3D animation sub-module.

3.1 General Assumptions

(1) Reshuffle operation is not considered in quay crane operation;

(2) The acceleration and deceleration processes of trucks are not implemented in this simulation model.

(3) The yard cranes would return the original location after loading and discharging operation.

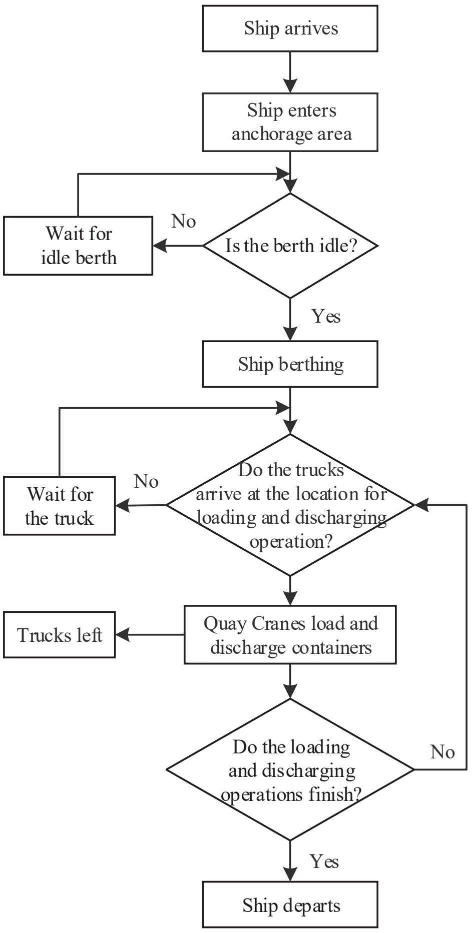

Ship arrives
Ship enters anchorage area
Wait for idle berth
No
Is the berth idle?
Yes
Ship berthing
Wait for the truck
No
Do the trucks arrive at the location for loading and discharging operation?
Trucks left
Quay Cranes load and discharge containers
Do the loading and discharging operations finish?
No
Yes
Ship departs

Figure 1: Logical Model of Operation for Container Terminal Apron

3.2 Terminal Layout and Setup Sub-Module

This sub-module is used to visualize the layout of a specific container terminal, and set the working paths and points of container trucks.

(1) Container terminal Layout. To realize the layout of the container terminal, we use the software AutoCAD to draw the masterplan of the terminal layout and then import it to AnyLogic. The plan is consistent with the actual size of the container terminal to facilitate the analysis on the operation process in terminal apron.

(2) The transfer routes and transfer points for yard trucks are set according to the layout plan of the container terminal. In this sub-model, the transfer routes are the paths for horizontal transport, and the transfer points refer to the specific locations for loading and discharging operation.

(3) Entities and service resources. In the simulation system, the entities mainly include containers and ships. And the resources include container trucks, quay cranes and yard cranes. Correspondingly, the parameters for entities are the type and number of containers, the type and number, as well as the arrival regulation of ships. And the parameters of service resources include the

number and speed of trucks, the number and handling time of quay cranes and yard cranes, as well as the yard capacity of the container terminal. This sub-model is also used to setup values of variables and parameters.

3.3 Ship Behavior Sub-Module

This sub-model is used to simulate the process of ship arriving, ship berthing and ship departing.

(1) Ship arriving: The ship arrives at the port according to ship arrival pattern. And the operator of the port makes a berthing plan for the ship according to the arrival information in advance. In this sub-model, the "ShipArrival" module is used to generate the arriving ships in accordance with regulation set in sub-module of set variables and parameters.

(2) Ship berthing: If the assigned berth is occupied by other ships, the ship should wait in anchorage area. And when the assigned berth is idle, the ship starts berthing operation according to the guidance of operators. In this process, the Module "SelectOutput5" is used to evaluated whether is assigned berth is idle. If the assigned berth is occupied with other ships, the Module "Delay" is applied to prevent the ships from berthing. On the contrary, if the berth is idle, the ship starts berthing operation, and the Module "Source" is used to generate containers based on the "Shiptype" function defined in the sub-module of Terminal Layout and Setup sub-module. Then the "SelectOutput5" module is used to allocate quay cranes and call the trucks to move to the terminal apron with the purpose of accomplishing the loading and discharging operation.

3.4 Quay Crane Operation Sub-Module

This sub-module is used to realize the loading and discharging operations. Figure 2 shows the logic model of loading and discharging operations on container terminal apron.

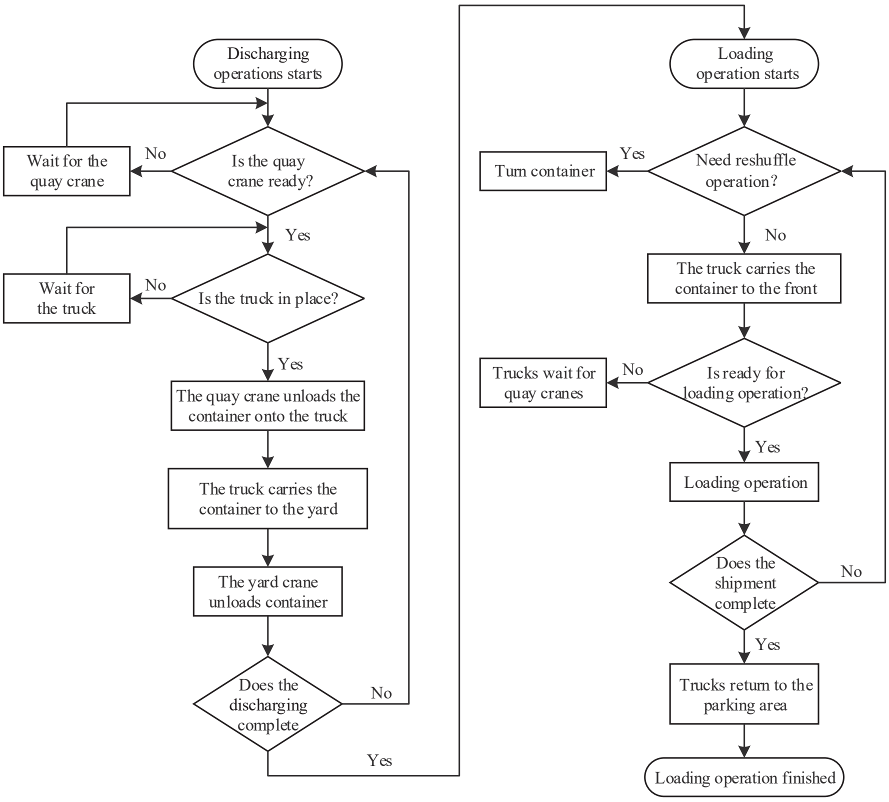

Discharging operations starts
Wait for the quay crane
No
Is the quay crane ready?
Yes
Wait for the truck
No
Is the truck in place?
Yes
The quay crane unloads the container onto the truck
The truck carries the container to the yard
The yard crane unloads container
Does the discharging complete
No
Yes
Loading operation starts
Turn container
Yes
Need reshuffle operation?
No
The truck carries the container to the front
Trucks wait for quay cranes
No
Is ready for loading operation?
Yes
Loading operation
Does the shipment complete
No
Yes
Trucks return to the parking area
Loading operation finished

Figure 2: Flow Chart of Loading and Unloading Ship

(1) Discharging Operation: we use the "Seize" module to realize the request of container trucks and quay cranes. And the "Lifting-Translation-Dropping" module is used to simulate the operating animation and record operating time of quay cranes. The "Queue" module is applied to simulate the queue of container trucks, and after unloading the container on to the truck, the quay crane will be released by "Release" module. In addition, the "combine" module is used to output loaded container trucks, which enter the yard operation system by "SelectOutput5" module.

(2) Loading Operation: in this system, we focus on the process that the trucks retrieve containers from different blocks and transport them to the terminal apron where the containers are loaded on to ships by quay cranes. And when the loading and discharging tasks are finished, the quay cranes are released and the ship deberths, then the berth will be released.

3.5 Horizontal Transport and Yard Operation Sub-Module

The horizontal transport studied in this paper includes horizontal transport for loading operation and discharging operation. The movements of trucks in horizontal transports are realized by "MoveTo" module. In the horizontal transport for discharging operation, the loaded trucks from "Combine" module enter the yard by "SelectOutput5" module, carry the containers unloaded from the ship by quay cranes, and move to the specified blocks to start yard operation by "MoveTo" module.

In the horizontal transport for loading operation, the trucks transport the containers from the yard to the terminal apron by "SelectOutput5", and then the containers are loaded on the ship by quay cranes.

3.6 3D Animation Sub-Module

This sub-module is used to realize a 3D animation of container terminal apron operation, which helps the planners to identify the bottleneck that may be encountered during the operation of container terminals.

(1) Animation of ship waiting for berth and berthing: When the ship arrives, the system can evaluate whether the assigned berth is idle according to the current state of the berth. If it is idle, the ship will berth at the assigned berth, as shown as in Figure 3. If not, the ship will wait in anchorage area until the berth is idle, as shown in Figure 4.

(2) Animation of quay cranes operation: Figure 5 shows the 3D animation of containers being lifted by spreader of the quay crane. The containers are moved on to ships from trucks by quay cranes according to the steps of "Lifting-Translation-Dropping" in loading process. And in discharging process, following the same steps, the containers are moved on to trucks from ships by quay cranes by quay cranes. With the handling process of quay cranes, the number of containers on ships would change.

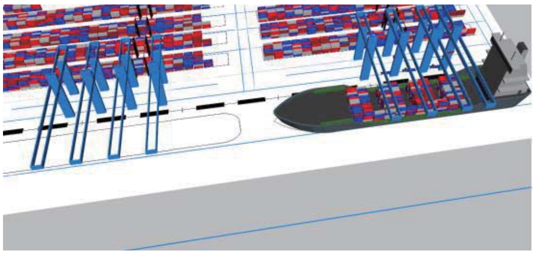


Figure 3: Ship at Berth

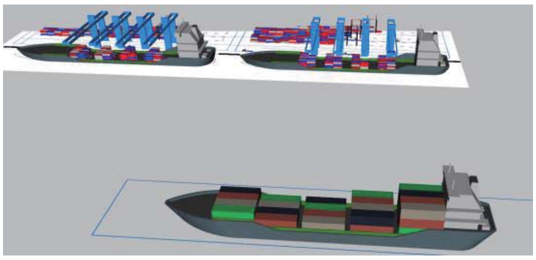


Figure 4: Ship Waiting for Berth

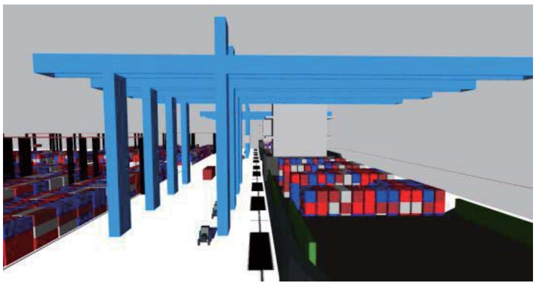


Figure 5: Loading and Discharging Process by Quay Cranes

3.7 Model Verification and Validation

This model is verified and validated to confirm that it is correctly implemented with respect to the process of terminal apron operation. We consider the design ship is 70000-DWT, the number of quay cranes is 4 for each berth, the number of yard cranes is 18, and the number of trucks allocated for one quay crane is 5. After the model is performed for one week, we get the average utilization of quay cranes. The output result from this model is 64%, and the actual utilization of quay cranes is within 62%~77%. Therefore, the simulation model proposed in this paper can be used to simulate the processes of terminal apron operations.

4 CASE STUDY

The case study considers a container terminal with two 70000-DWT berths in the north of China. As shown in Figure 6, the berths are arranged along the shore, and the length of the two berths is 680m. The quay cranes are used to loading and discharging operation in the terminal apron, and the rubber-tired gantry cranes (RTGs) are used for yard operation. And the internal and external trucks are for horizon transport between quay and storage yard, as well as between storage yard and landside interfaces.

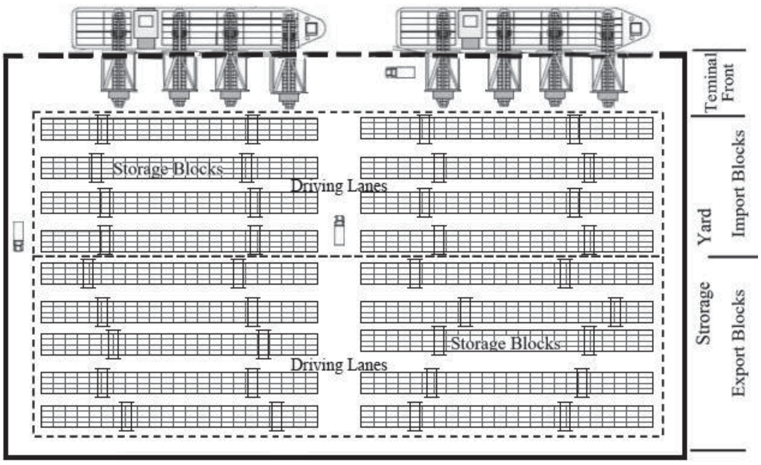


Figure 6: The Layout of Container Terminal in This Case Study

4.1 Model Parameters

(1) Ship arrival: the interval of ship arrival follows the negative exponential distribution with $\lambda = 4$, and the numbers of containers handled in this terminal obey the uniform distribution within 2000 ~ 2250TEU.

(2) Quay cranes: each berth is equipped with 4 quay cranes, considering the instability of practical operation, the time for the single operation of the quay crane (T) follows the uniform distribution of 80~120s and 120~160s respectively for two parallel experiments.

(3) Yard cranes: the yard cranes are RTGs of SRTG5223S, with the dimensions of 26.5m long, 26.5m high and 15.2m wide, and the maximum lifting height of 18.2m. In each block, the number of RTGs (*Nyc*) is 1 and 2 respectively for two parallel experiments.

(4) The containers in the case are standard 40-foot containers, and the external dimensions are 2.9m (height) * 2.44m (width) * 12.2m (length).

(5) There are 20 internal trucks, and according to MTPRC (Ministry of Transport of the People's Republic of China, 2014), the speed limit is 35km/h for empty trucks, and 25km/h for loaded truck. Based on the number of trucks (N_t) from 4 to 8 allocated for each working path, 5 parallel experiments are conducted.

4.2 Simulation Experiments

This experiment mainly changes the following three parameters: the time for the single operation of the quay crane (T), the number of trucks allocated for each quay crane (*Nt*) and the number of yard cranes equipped for each block (*Nyc*). And according to these parameters, 20 simulation schemes are evaluated to estimate the apron performance. For each scheme, similar simulations are performed 10 times for a period of a week.

4.3 Results and Discussion

Running the 20 simulation schemes, we can obtain the average utilization ratio of yard cranes and the average handling efficiency of quay cranes. As shown in Figure 7 and 8, we can draw the following conclusions:

(1) The utilization of quay cranes

Figure 7 compares the utilization of quay cranes of different *T*, N_t and N_{yc}. And the results show that the three variables are the important factors for the utilization of quay cranes. For the condition with the same *T* and N_{yc}, with the increase of N_t, the utilization of the quay crane increases at first and then decreases. And for the condition with the same N_t and $N_{yc,}$, with the increase of *T*, the utilization of the quay crane increases; And for the condition with the same N_t and T, with the increase of $N_{yc,}$, the utilization of the quay crane increases.

Taking the combination of *T*=140s, $N_{yc,}$=2 as an example, we can find that when N_t =5, the utilization of quay cranes is 60%. If the N_t increases to 6, the utilization of quay cranes would increase to 70%. However, if the N_t increases to 7, the utilization of quay cranes would decrease to 65%. Therefore, when *T* and $N_{yc,}$ are determined, the N_t is not the more the better for a container terminal. In each scenario, there is an optimal N_t for getting the maximum utilization of quay cranes.

From the view point of *T*, we take the combination of N_t =6, $N_{yc,}$ =2 as the example. For *T* =100s, the utilization of quay cranes is 52%, and for *T* =140s, the utilization of quay cranes is 70%. This result can be explained that the completed time of loading and discharging the same ship would become less with the operation level improved, and there would no need to equip more quay cranes in the port. Therefore, if the workers with high operation level are employed, the cost of purchasing the machines can be saved.

Moreover, from the view point of $N_{yc,}$, taking the combination of N_t =6, T =100s as the example, we can find that when $N_{yc,}$ =1, the utilization of quay cranes is 40%, and when $N_{yc,}$ =2, the utilization of quay cranes is 60%. It is concluded that the efficiency of the yard has an important effect on the utilization of quay cranes. If the efficiency of the yard is neglected, the yard operation would become the bottleneck of improving the utilization of quay cranes.

(2) The efficiency of quay cranes

Figure 8 shows the efficiency of quay cranes of different T, N_t and N_{yc}. And the results demonstrate that the three variables are the important factors for the efficiency of quay cranes. For the condition with the same T and N_{yc}, with the increase of N_t, the efficiency of the quay crane increases at first and then decreases. And for the condition with the same N_t and $N_{yc,}$, with the increase of T, the efficiency of quay cranes decreases; And for the condition with the same N_t and T, with the increase of $N_{yc,}$, the efficiency of quay cranes increases.

Taking the combination of T=140s, $N_{yc,}$=2 as an example, we can find that when N_t =5, 6 and 7, the efficiency of quay cranes is 50, 54, 52 TEU/h respectively. Therefore, with the similar conclusion of the utilization of quay cranes, when T and $N_{yc,}$ are determined, the N_t is not the more the better for a container terminal. In each scenario, there is an optimal N_t for getting the maximum efficiency of quay cranes.

From the view point of T, taking the combination of N_t =5, $N_{yc,}$ =2, we can find that when T =100s, the efficiency of quay cranes is 60 TEU/h, and when T =140s, the efficiency of quay cranes is 50 TEU/h. This result can be explained that as the operation level is improved, the efficiency of loading and discharging operation can be improved naturally. In addition, from Figure 8, we can also find that if $N_{yc,}$ =1, the optimal efficiency of quay cranes for T =100s and T =140s could be obtained when N_t = 6 and 5 respectively. And if $N_{yc,}$ =2, the optimal efficiency of quay cranes for T =100s and T =140s could be obtained when N_t = 7 and 6 respectively. For this reason, with the improvement of operation level of the workers, the container trucks allocated for the quay cranes should be increased appropriately according to the actual condition of the container terminal.

Furthermore, from the view point of N_{yc}, we take the combination of N_t =7, T =100s as the example. And we can find that when N_{yc} =1, the efficiency of quay cranes is 42%, and when N_{yc} =2, the efficiency of quay cranes is 62%. Therefore, in order to increase the efficiency of operation in the terminal front, the efficiency of yard operation should be taken into consideration, as the operation system of a container terminal is an integrated system, of which all the sub-systems are interrelated and mutual restraint.

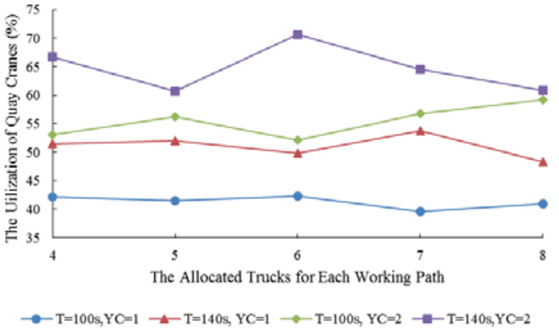


Figure 7: The Simulation Results of the Utilization of Quay Cranes

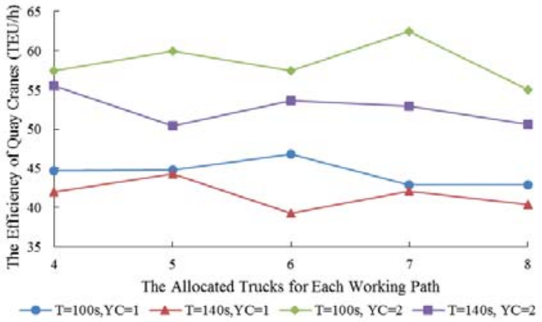


Figure 8: The Simulation Results of the Efficiency of Quay Cranes

5. CONCLUSIONS

The main contribution of this paper is to provide a 3D simulation model of container terminal apron operation. By controlling the exact coordinates of various processes, the proposed simulation model can simulate the real situation of container terminal apron, including ships arriving, berthing, as well as loading and discharging, etc. Based on this 3D simulation model, a good reference for operation planning of a container terminal apron can be provided.

ACKNOWLEDGEMENTS

This paper is a partial result of the project supported by the National Natural Science Foundation, China (No. 51579035) and Support High-Level Talents Innovation and Entrepreneurship Projects, Dalian, China (No. 2016RQ024).

REFERENCES

AnyLogic: AnyLogic. http:// www.anylogic.com (Accessed from April 2017)

Azab A.E, Amr B, Eltawil. 2016. "A Simulation Based Study of the Effect of Truck Arrival Patterns on Truck Turn Time in Container Terminals". *In Proceedings of the 30th European Conference on Modelling and Simulation*, 80-86.

Bruzzone A.G; Longo F. 2013. "3D Simulation as Training Tool in Container Terminals". *The Transports Simulator Journal of Manufacturing Systems*, Vol 32, No.1, 85-98.

Ji MJ, Jin ZH. 2007. “The Integrated Optimization Model for Berth Allocation and Container Truck Assignment”. *Journal of Fudan University (Natural Science)*, Vol 46, No.4, 476-480+488.

Lin W. 2011. “The Research on Optimization Model on Traffic Organization of Loading and Discharging Process for Super Large Container Ships”. *China Water Transport*, Vol 11, No.1, 41-42+45.

Ministry of Transport of the People’s Republic of China (MTPRC). (2014) Design code of general layout of sea ports, JTS 165-2013. Beijing: China Communication Press.

Ren S. 2011. “The Research on the Optimized Simulation of the Container Terminal Key Operation Equipment”. Dalian Maritime University, Dalian, China.

Sun Z et al. 2012. “A General Simulation Platform for Seaport Container Terminals”. *Advanced Engineering Informatics*, Vol 26, No.1, 80-89

Zhang YH et al. 2016. “Effect of Block Length on Operation Efficiency of Container Terminal”. *Port & Waterway Engineering*, No.11 (Nov), 94-98.

AUTHOR BIOGRAPHIES

JINGJING YU was born in Chaoyang City, Liaoning Province, China, and went to Dalian University of Technology, where she majored in port and waterway engineering and obtained the Bachelor’s Degree in 2015. Now, she is studying for a Doctor’s Degree in the field of simulation for port and waterway engineering analysis. Her e-mail address is: yigaoyujingjing@mail.dlut.edu.cn, and her Webpage can be found at http://port.dlut.edu.cn.

GUOLEI TANG was born in Yantai City, Shandong Province, China, and went to Dalian University of Technology, where he obtained the Doctor’s Degree in Hydrology and Water Resources in 2009. He worked for a couple of years for the simulation modeling in engineering, and now, he is leading a large research group in the field of simulation for port and waterway engineering analysis. His e-mail address is: tangguolei@dlut.edu.cn, and his Webpage can be found at http://port.dlut.edu.cn.

DA LI was born in Weifang City, Shandong Province, China, and went to Harbin Engineering University to study port and waterway engineering. In 2016, he obtained the Bachelor’s Degree. Now, he is studying for a Master’s Degree in the field of simulation for port and waterway engineering analysis in Dalian University of Technology. His e-mail address is: dlutllida@mail.dlut.edu.cn, and his Webpage can be found at http://port.dlut.edu.cn.

BAOYING MU was born in Tangshan City, Hebei Province, China, and went to Harbin Engineering University, where she majored in Port & Coastal and Waterway Engineering and obtained the Bachelor's Degree in 2016. Now she is majoring in Port & Coastal and Offshore Engineering for a Master's Degree in Dalian University of Technology. Her e-mail address is: mubaoying@mal.dlut.edu.cn, and her webpage can be found at http://port.dlut.edu.cn.

CONTAINER TERMINALS CAPACITY EVALUATION CONSIDERING PORT SERVICE LEVEL BASED ON SIMULATION

Ningning Li

Dalian Neusoft University of Information
Dalian 116032, China

Jingjing Yu
Guolei Tang*
Da Li
Yong Zhang

Faculty of Infrastructure Engineering
Dalian University of Technology
Dalian 116024, China
*E-mail: tangguolei@dlut.edu.cn

KEYWORDS

Container Terminal, Throughput Capacity, Port Service Level, Simulation and Modeling.

ABSTRACT

Throughput capacity is the production capacity of port enterprise under constant exotic environment, and plays a significant role in production control. According to the analysis on influencing factors of throughput capacity and characters of operation system, this paper proposes a simulation model of container terminal operation system based on port service level. By changing input parameters, different simulation schemes can be obtained with the objective to define the relationship between port service level and throughput capacity of container terminals. And some reasonable suggestions can be given to improve the throughput capacity of container terminals.

1. INTRODUCTION

With the development of economy and the expansion of foreign trade, container throughputs of coastal ports have increased rapidly in China. So new container berths should be planned but with limited shoreline resources. The throughput capacity is an important issue involving respective interests from governments and enterprises. To save coastal port resources, and provide scientific and reasonable port resources for shipping lines, the throughput capacity of container terminal should be evaluated reasonably to determine an optimal berth numbers. In China, a mandatory Design Code of General Layout for Sea Port (MTPRC 2014) provides a set of procedures to evaluate container terminal capacities. However, the existing methods evaluate the throughput capacity without considering port service level and interactions between subsystems. Therefore, a practical approach to evaluate the throughput capacities of Chinese container terminals based on MTPRC (2014).

Many researches have been carried out on throughput capacity of container terminals. These achievements are mainly divided into the following three categories. The first one studies the port capacity using queue theory (Wang et al. 2008; Lee et al. 2014). For example, Wang et al. (2008) used Stochastic Petri Net to establish both a hierarchical model of the container terminal capacity and a dynamic model of subsystems to determine the capacities of the subsystems and detect the bottleneck of port system. However, as the data and the queue configuration are more sophisticated, the researchers have to resort to simulation (Demirci 2003, Quy et al. 2008, Imai et al. 2001, 2005, Wanke 2011, Tang et al. 2016, Azab et al., 2016). So the second one focuses on throughput capacities of specific container terminals/berths using simulation (Wang et al. 2004; Wu et al. 2013). For example, Wu et al. (2013) built a simulation model for barge berths of Kwan Chung container terminals to examine the relationship between berthing capacity and service level in terms of vessel waiting time. And the last one covers the impact analysis of the factors on container terminal capacities (Xie 2008; Liu 2009; Ding 2010; Zhang 2013) For example, Ding (2010) established a simulation model to estimate the throughput capacities of a container terminal under different combination patterns of the types of arriving ships. These researches provide invaluable information and insights regarding methodologies how to describe the stochastics characteristics of ship arrivals and berth service, how to evaluate terminal's service level, and how to simulate the ship-berth link planning operation.

Therefore, considering the stochastic and dynamic characteristics of port system (Demirci 2003, Quy et al. 2008, Tang et al. 2016), in this paper, on the basis of Chinese mandatory Design Code (MTPRC 2014), we establish a simulation model of container terminal operations, to estimate the throughput capacity in terms of port service level for container terminals using Arena simulation software (Arena 2017). And the deduced relationships between port service level and throughput capacity of container terminals which are main contributions of this paper, will provide some reasonable suggestions to container terminal planning.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

2. CONTAINER TERMINAL CAPACITY

According to Chinese mandatory Design Code of General Layout for Sea Ports (MTPRC 2014), when planning a container terminal, an important consideration is to provide a sufficient annual container-handling capability (terminal capacity). Obviously, the acceptable level of service provided by a terminal is not considered when evaluating terminal capacity. Therefore, in this paper, we define the container terminal capacity as the capacity of the container terminal, in terms of containers (Twenty-foot Equivalent Unit, TEU) that can be handled per year with an adequate service level.

The chosen indicator for port service level is the average waiting time / average service time ratio, expressed as S=AWT/AST, in which, AWT represents the average waiting time of ships in the port and AST represents the average service time required for loading and discharging a ship under normal circumstances.

3. SIMULATION MODELING FOR CONTAINER TERMINAL OPERATIONS

3.1 Basic Assumptions

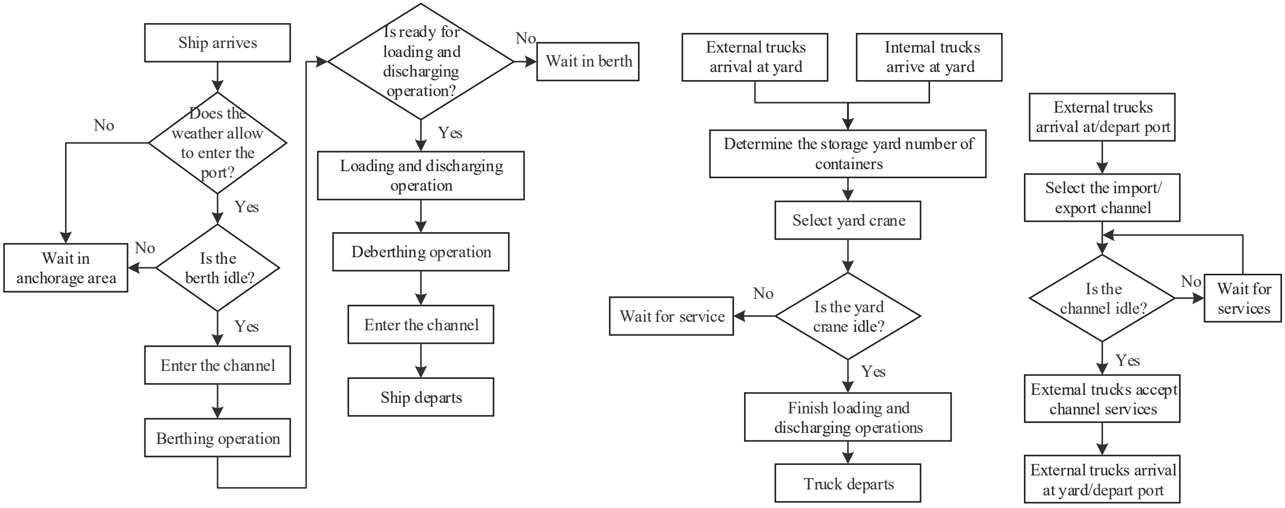
(1) Container ships are served on first-come-first-served basis (FCFS).

(2) The loading operation at apron starts once the discharging operation is finished.

(3) The quay cranes and yard cranes are equipped according to the berth tonnage, and based on the number of quay cranes, the internal trucks can be allocated with the principle of shortest path.

3.2 Simulation Model Implementation

In most container terminals in China, the orientation of the storage blocks is parallel to the shore. And rubber-tired gantry cranes (RTGs) are used for the yard operations, and internal and external trucks for the horizon transport between quay and storage yard, as well as between storage yard and landside interfaces (Ji et al. 2010). According to the logic model illustrated in Figure 1, we implement a simulation model to simulate the processes of Chinse container terminal, which covers five sub-models, which are ship berthing and deberthing sub-model, ship loading and discharging sub-model, horizontal transport sub-model, yard operation sub-model and gate operation sub-model.


Ship arrives
Does the weather allow to enter the port?
No
Yes
Wait in anchorage area
No
Is the berth idle?
Yes
Enter the channel
Berthing operation
Is ready for loading and discharging operation?
No
Wait in berth
Yes
Loading and discharging operation
Deberthing operation
Enter the channel
Ship departs
External trucks arrival at yard
Internal trucks arrive at yard
Determine the storage yard number of containers
Select yard crane
No
Wait for service
Is the yard crane idle?
Yes
Finish loading and discharging operations
Truck departs
External trucks arrival at/depart port
Select the import/ export channel
Is the channel idle?
No
Wait for services
Yes
External trucks accept channel services
External trucks arrival at yard/depart port

(a) Container terminal Front Operation (b) Yard Operation (c) Gate Operation

Figure 1: The Related Logical Model of Container Terminal Operation

We establish the model using Arena software, and the Figure 2 shows the simulation model of container terminal operations using Arena. The modeling processes are described as follows:

(1) Ship berthing and deberthing sub-model simulates the process that the inbound ship travels through the channel from the anchorage and arrives at the berth, and after the berth service time for discharging and loading containers, the outbound ship deberths, travels through the channel and leaves the port. In this sub-model, the entity is the ship with some attributions, such as the dimension, tonnage and single ship loading and discharging capacity.

(2) Ship loading and discharging operation sub-model includes two processes: loading and -discharging. The loading process is that the internal trucks transport containers to the terminal apron, and then the assigned quay cranes load containers onto the ship. The discharging process is to unload containers onto trucks from the ship. In this sub-model, the entity is the container with assigned type and dimensions.

(3) Horizontal transport sub-model simulates the horizontal movement between the berth and yard or gate. In this sub-model, we set the container and yard truck as the entities, and its attributes of truck include the assigned yard block, and travel route.

(4) Yard operation sub-model simulates the loading and discharging process when containers are transported to the yard by trucks. In this sub-model, the entity is the container, and the resources are the bays for stacking containers and RTGs in the block of destination.

(5) The gate operation sub-model provides entrance roads for external trucks. The entity in this sub-model is the trucks, and the resources are the access roads.

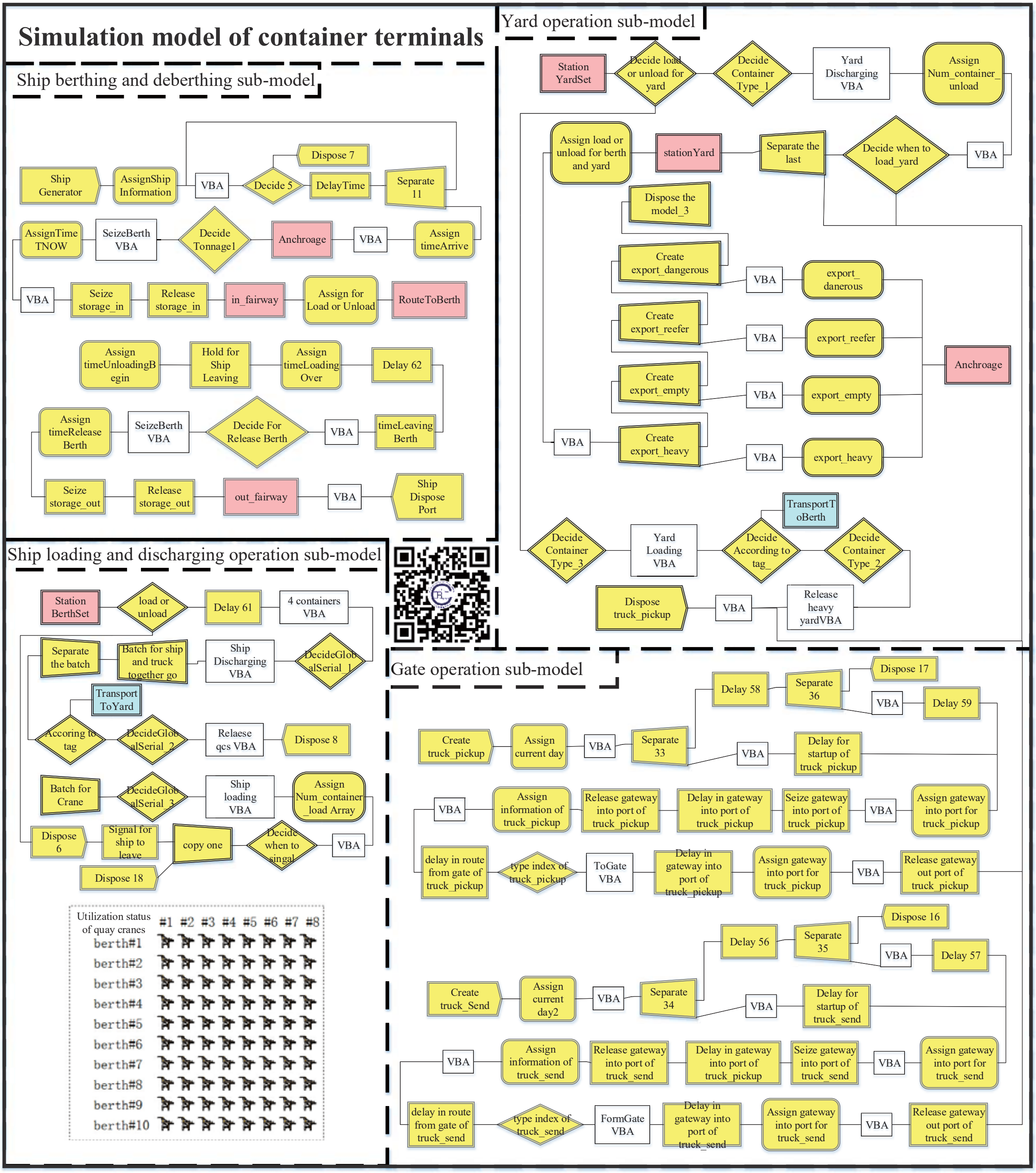


Figure 2: The Simulation Model of Container Terminal

3.3 Model Verification and Validation

To verify and validate the implemented simulation model that it is correctly implemented with respect to the process of throughput capacity, we take three effective measures: Firstly, the model is developed in stages and through sub-models in which each stage is individually examined by a subject-matter expert. Secondly, tracing approach is used throughout the model development phase. Via tracing, we compare the simulation results with manual calculations to check if the logic implemented in the model is as intended. Finally we take Yantian International Container Terminal (YICT) as the example. The actual throughput of YICT is 8.62 million TEU in 2010. And model parameters can be obtained according to the actual operation situation of YICT in 2010. Then the throughput calculated by using this model is 8.74 million thousand TEU, with the discrepancy of 1.39%. So the implemented simulation model is reliable and can be used for further study.

4. RELATIONSHIP BETWEEN PORT SERVICE LEVEL AND THROUGHPUT CAPACITY OF CHINESE CONTAINER TERMINALS

4.1 Simulation Setup

Container terminal characteristics include the number of berths and their tonnages, and distribution of berth service time. In this study, the simulation experiments evaluate 8 classes of n-DWT berths i.e., n = {10000t, 20000t, 30000t, 50000t, 70000t, 100000t, 120000t, 150000t}, and 6 options for the number of each class of berths, n_{bth} = {1, 2, 3, 4, 5, 6}, totaling 48 container terminal scenarios to be investigated.

The values or distributions of the simulation model parameters, are determined according to Chinese mandatory Design Code of General Layout for Sea Ports (MTPRC 2014). For example, the ships arrive rates follow Poisson distribution with the daily number of ship arrivals varying within certain ranges. Other parameters' values are listed in Table 1 and 2.

Table 1: Some Model Parameters

Model parameters		Value
Time (h)	Auxiliary operation and berthing time	3~5
Efficiency of yard handling equipment (TEU/h)	Heavy container yard	40
	Empty container yard	60
	Dangerous container yard	40
	Refrigerated container yard	40
Gate inspection time (s)	Ingate empty trucks	TRIA(20,25,30)
	Ingate loaded trucks	TRIA(30,40,50)
	Outgate empty trucks	TRIA(5,10,15)
	Outgate loaded trucks	TRIA(30,40,50)

Table 2: Model Parameters of equipment

Tonnage DWT (t)	Design efficiency of quay crane(TEU/h)	Number of quay crane per berth	Number of trucks per quay crane	Number of RTGs per quay crane
10000	40	2	14	8
20000	40	2	14	8
30000	48	3	21	12
50000	48	4	28	16
70000	48	4	28	16
100000	60	5	35	20
120000	80	5	35	20
150000	80	5	35	20

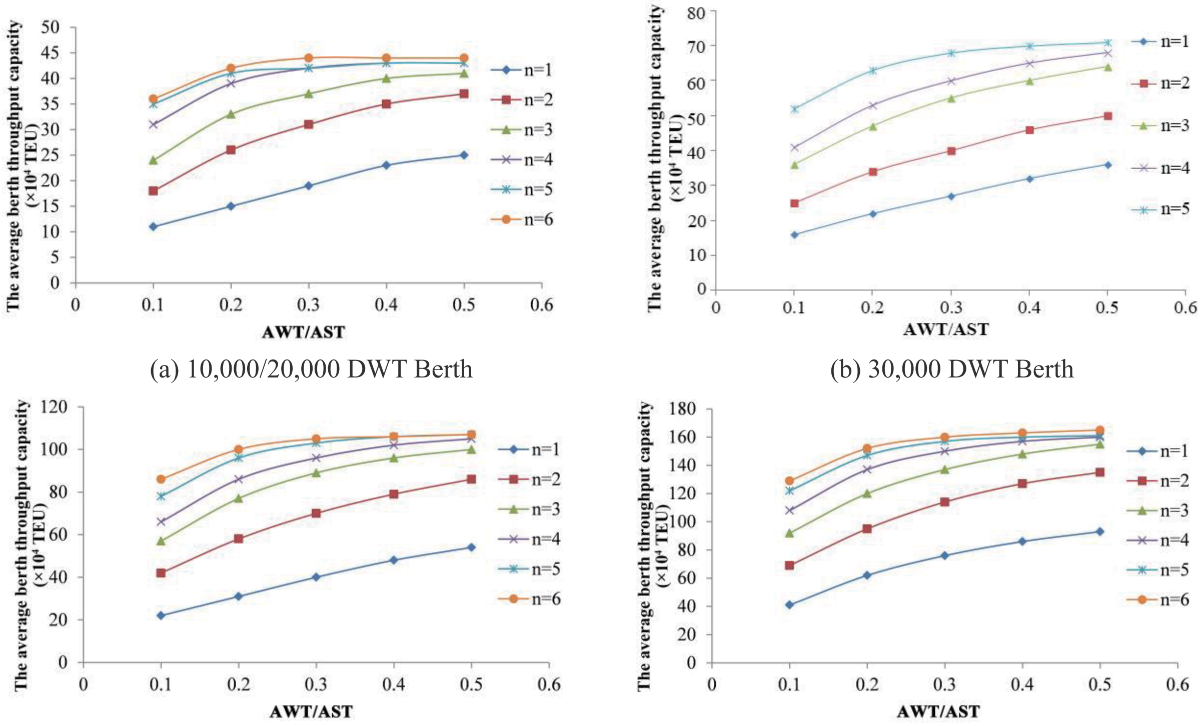
4.2 Analysis and Discussion

Running the simulation model, the relationship between port service level (AWT/AST) and average berth capacity (P_t) of container terminals with the corresponding of change in the number of berths for different tonnages of berths can be obtained:

(1) As shown in Figure 3, given the same number of berths, the terminal throughput capacity increases with the values of AWT/AST. And the relationship between terminal capacity and AWT/AST follows an exponential function with monotone increasing.

(2) As shown in Figure 4, given the same value of AWT/AST, when the number of berths is larger than 3~5, the throughput capacity of marginal berths drops rapidly. Therefore, a terminal with 3~5 berths is relatively economic and reasonable design.

(3) Given the same number of berths, container berth can be divided into 4 classes (10,000/20,000 DWT), 30,000 DWT, (50,000/70,000 DWT), and 100,000 DWT based on throughput capacity with a certain port service level. The recommended terminal throughput capacity in terms of AWT/AST are listed in Table 3, which are used to evaluate the terminal capacity and determine the number of new berths.



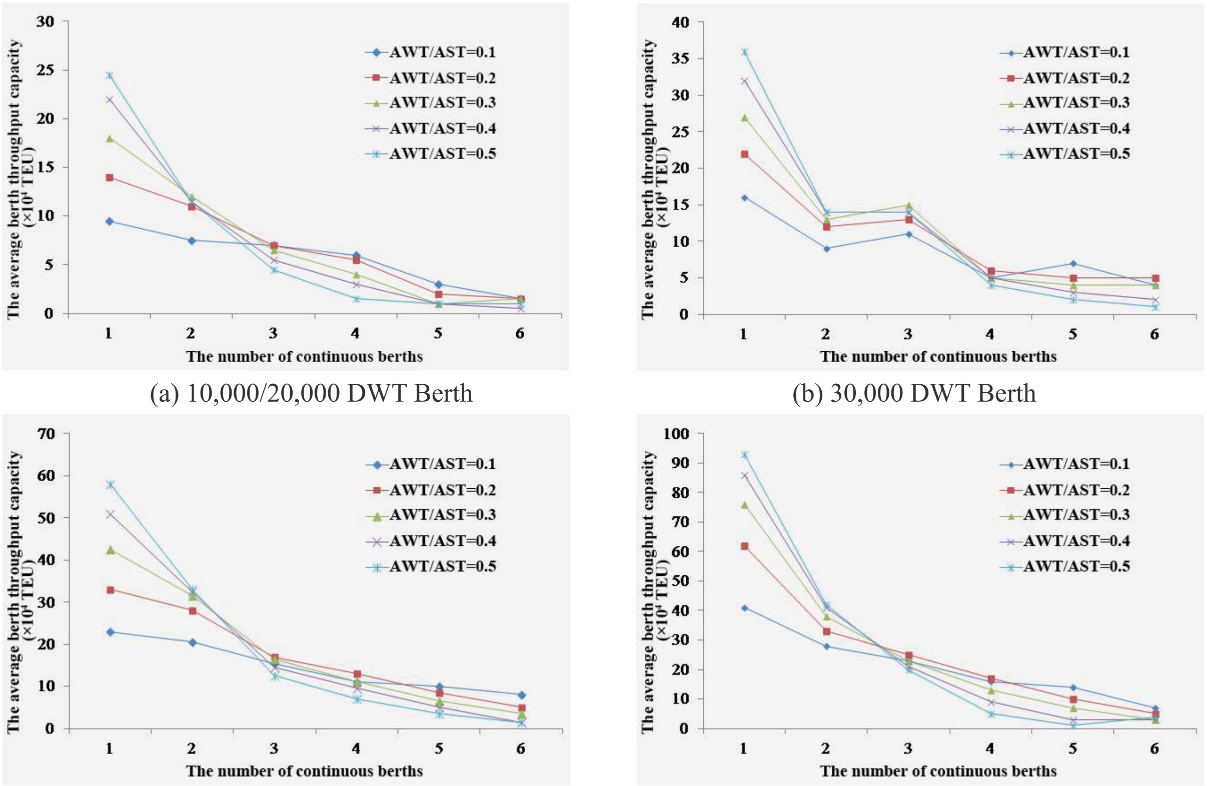
(a) 10,000/20,000 DWT Berth

(b) 30,000 DWT Berth

(c) 50,000 DWT Berth

(b) 100,000 and above DWT Berth

Figure 3: The Relationship between AWT/AST and Average Berth Throughput Capacity with Different Combination of Berths



(a) 10,000/20,000 DWT Berth

(b) 30,000 DWT Berth

(c) 50,000 DWT Berth

(b) 100,000 and above DWT Berth

Figure 4: The Relationship between Different Number of Continuous Berths and Average Throughput Capacity of Marginal Berth with Different Tonnage of Berths

Table 3: Berth Throughput Capacity for Different Number of Continuous Berths with Varying Berth Tonnage

The Number of Continuous Berths	Tonnage of Berths	AWT/AST				
		0.1	0.2	0.3	0.4	0.5
1	10,000/20,000 DWT	10	14	18	22	25
	30,000 DWT	16	22	27	32	36
	50,000/70,000 DWT	23	33	43	51	58
	100,000 DWT	41	62	76	86	93
2	10,000/20,000 DWT	17	25	30	34	36
	30,000 DWT	25	34	40	46	50
	50,000/70,000 DWT	44	61	74	84	91
	100,000 DWT	69	95	114	127	135
3	10,000/20,000 DWT	24	32	37	39	41
	30,000 DWT	36	47	55	60	64
	50,000/70,000 DWT	59	78	91	98	104
	100,000 DWT	92	120	137	148	155
4	10,000/20,000 DWT	30	38	41	42	42
	30,000 DWT	41	53	60	65	68
	50,000/70,000 DWT	70	91	102	108	111
	100,000 DWT	108	137	150	157	160
5	10,000/20,000 DWT	33	40	42	43	43
	30,000 DWT	48	58	64	68	70
	50,000/70,000 DWT	80	100	108	113	114
	100,000 DWT	122	147	157	160	161
6	10,000/20,000 DWT	35	41	43	44	44
	30,000 DWT	52	63	68	70	71
	50,000/70,000 DWT	88	105	112	114	116
	100,000 DWT	129	152	160	163	165

5. CONCLUSIONS

In this paper, we have established a simulation model of the container terminal operation system to obtain the relationship between the port service level and the throughput capacity of container terminals. According to the results of this simulation model, we can draw some conclusions.

(1) The average berth throughput capacity of container terminals increases exponentially with the value of the port service level (AWT/AST) increasing given the same number of berths.

(2) The continuous berth number *n* has great influence on the average berth capacity of container terminals. And the continuous arrangement of 3~5 berths is relatively economical and reasonable.

(3) The recommended terminal throughput capacity in terms of AWT/AST are deduced, which are used to evaluate the terminal capacity and determine the number of new berths.

The simulation model provides technical support for the further systematical research on the throughput capacity of container terminals, and it can also be used as a guide for the planning and operation of container terminals.

ACKNOWLEDGEMENTS

This paper is a partial result of the project supported by the National Natural Science Foundation, China (No. 51579035) and Support High-Level Talents Innovation and Entrepreneurship Projects, Dalian, China (No. 2016RQ024).

REFERENCES

Arena: Arena http:// www.arenasimulation.com /. (Accessed from January 2017)

Azab A.E, Amr B, Eltawil. 2016. “A Simulation Based Study of the Effect of Truck Arrival Patterns on Truck Turn Time in Container Terminals”. *In Proceedings of the 30th European Conference on Modelling and Simulation*, 80-86.

Demirci, E. 2003 Simulation modeling and analysis of a port investment, Simulation, 79, 94-105.

Ding YZ 2010. “Throughput Capacity of a Container Terminal Considering the Combination Patterns of the Types of Arriving Vessels”. *Journal of Shanghai Jiaotong University (English Edition)*, Vol 15, No.1, 124-128.

Imai, a., Nishimura, E. and Papadimitrou, S. (2001) The dynamic berth allocation problem for a container port. Transportation Research Part B, 35: 401-417.

Imai, a., Sun, X., Nishimura, E., Papadimitrou, S. (2005) Berth allocation in a container port: using a continuous location space approach. Transportation Research part B, 39: 199-221.

Lee B.K. et al. 2014. “Analysis on Container Port Capacity: A Markovian Modeling Approach.” *OR Spectrum*, Vol 36, No.2, 425–54.

Liu F. 2009. “Study on the Berth Capacity and Related Indexes of Container Terminal”. Dalian Maritime University, Dalian, China.

Ministry of Transport of the People’s Republic of China (MTPRC). (2014) Design code of general layout of sea ports, JTS 165-2013. Beijing: China Communication Press.

Quy, N.M., Vrijling, J.K., and Van Gelder, P.H.A.J.M. (2008). "Risk-and simulation-based optimization of channel depths, entrance channel of Cam Pha Coal Port." Simulation, 84, 41-55.

Tang, G., Guo, Z., Yu, X, Song, X, Du P. (2014) “SPAC to improve port performance for seaports with very long one-way entrance channels.” Journal of Waterway, Port, Coastal, and Ocean Engineering, 140(040140114).

Wang WY. et al. 2008 “System Simulation of Capacity for Container Terminal Based on Stochastic Petri Net”. *In Proceedings of 2008 International Conference on Automation and Logistics*, 2889-2892.

Wang ZM. 2004. “On the Reasonable Throughput Capacity of Container Terminals”. *Port and Waterway Engineering*, No.3, 16-20.

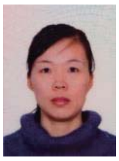
Wanke, P. (2011). Ship-berth link and demurrage costs: evaluating different allocation policies and queue priorities via simulation. Pesquisa Operacional, 31(1), 113-134.

Wu, YZ. Peng, C. 2013. “An Analysis of Capacity and Service level of the Container Terminals of Hong Kong”. *In Proceedings of the 10th International Conference on Service Systems and Service Management,* 404-409

Xie CX. 2008. “Study on Some Problems about the Calculation of the Berth Capacity of Container Terminal”. Dalian Maritime University, Dalian, China.

Zhang LN. 2013. “The Influence of Container Arrival Distribution on the Throughout Capacity”. Dalian Maritime University, Dalian, China.

AUTHOR BIOGRAPHIES



NINGNING LI studied in Shandong University from 2001 to 2005, and got the Bachelor’s Degree. Then she went to Dalian University of Technology and obtained the Master’s Degree in 2008. And now, she as an associate professor is working in Dalian Neusoft University of Information, focusing on data mining and mobile application.

JINGJING YU was born in Chaoyang City, Liaoning Province, China, and went to Dalian University of Technology, where she majored in port and waterway engineering and obtained the Bachelor’s Degree in 2015. Now, she is studying for a Doctor’s Degree in the field of simulation for port and waterway engineering analysis. Her e-mail address is: yigaoyujingjing@mail.dlut.edu.cn, and her Webpage can be found at http://port .dlut.edu.cn.

DA LI was born in Weifang City, Shandong Province, China, and went to Harbin Engineering University to study port and waterway engineering. In 2016, he obtained the Bachelor’s Degree. Now, he is studying for a Master’s Degree in the field of simulation for port and waterway engineering analysis in Dalian University of Technology. His e-mail address is: dlutllida@mail.dlut.edu.cn, and his Webpage can be found at http://port.dlut.edu.cn.

YONG ZHANG was born in Zaozhuang City, Shandong province, China, and went to Dalian University of Technology. Where he majored in port and waterway engineering and obtained the Bachelor’s Degree in 2016. Now he is studying for a master‘s Degree in the field of simulation for port and waterway engineering analysis. His e-mail address is horizon@mail.dlut.edu.cn and his Webpage can be found at http://port.dlut.edu.cn

GUOLEI TANG went to Dalian University of Technology, where he obtained the Doctor’s Degree in Hydrology and Water Resources in 2009. He worked for a couple of years for the simulation modeling in engineering, and now, he is leading a large research group in the field of simulation for port and waterway engineering. His e-mail address is: tangguolei@dlut.edu.cn, and his Webpage can be found at http://port.dlut.edu.cn.

HYBRID FLOW SHOP SCHEDULING OF AUTOMOTIVE PARTS

Tuanjai Somboonwiwat
Chatkaew Ratcharak
Department of Production Engineering
King Mongkut's University of
Technology Thonburi (KMUTT)
Thailand
E-mail: tuanjai.som@kmutt.ac.th

Tuangyot Supeekit
Department of Industrial Engineering
Mahidol University
Thailand
Email: tuangyot.sup@mahidol.edu

KEYWORDS

Automotive parts, Hybrid flow shop scheduling, Optimal production schedule

ABSTRACT

Flow shop scheduling problem is a type of scheduling dealing with sequencing jobs on a set of machines in compliance with predetermined processing orders. Each production stage to be scheduled in typical flow shop scheduling contains only one machine. However, in automotive part industry, many parts are produced in sequential flow shop containing more than one machines in each production stage. This circumstance cannot apply the existing method of flow shop scheduling. The objective of this research is to schedule the production process of automotive parts. The feature production is hybrid flow shop which consists of two-stages. In each stage, there are several manufacturing machines and each machine can produce more than one product. Thus, production scheduling is a complex problem. This paper, therefore, develops mathematical model to solve the hybrid flow shop production scheduling under different constraints of each machine. The setup time and production time of each machine can be different for each part. The solution for the experimental data sets from an automotive part manufacturer reveals that the process time can be reduced by 34.29%.

INTRODUCTION

Automotive industry is a very important sector for the country's economy since it generates trade and financial inflows to the country. Automotive part (auto part) manufacturers play an important role in the industry to supply parts for vehicle manufacturers. The response time of auto part manufacturers, which is the total amount of time the manufacturers takes to respond to the orders of auto parts, greatly affects the vehicle production. Responding to vehicle manufacturer demands, then is the goal of automotive parts manufacturers. They have to plan their productions and schedule the machine operations to ensure the shortest total completion time for all orders. Typically, the production type of the automotive parts manufacturers is flow shop where the processes are in a predetermined processing order; one process must be completed before another. The machine scheduling in the auto part manufacturer requires flow shop scheduling.

In flow shop scheduling the jobs must be produced through the first, second and the following stages. This scheduling problem is considered easy if there is only one machine for each production stage. The typical objective of flow shop scheduling is to minimize the makespan, i.e. to find the minimum total time needed to finish all of the production orders. Hence, the sequencing is decided for the scheduling problem. Typically, the flow shop scheduling deals with scheduling a number of jobs on different stages which contain only one machine on each stage. If each stage consists of many machines working in parallel, this problem is called hybrid flow shop scheduling (Choi et al. 2009). The scheduling problem becomes complicated which assigning and sequencing are required. The previous studies regarding the hybrid flow shop scheduling employ heuristics approaches to schedule the production. For example, Vignier et al. (1996) applies a branch and bound based algorithm to schedule jobs in multi-stages flow shop to minimize the makespan. Watanakich (2001) studies a two-stage hybrid flow shop scheduling with machine setup time, and solved the problem using a heuristic. He presents a two phase heuristic approach; constructing a schedule and assigning jobs with setup time consideration. This represents a difference between regular and hybrid flow shop scheduling. Wong et al. (2001) propose a genetic algorithm to schedule cutting and sewing operations in a manufacturer. Mallikarjuna et al. (2013) apply tabu search algorithm to complete flow shop scheduling. Puck-In (2014) tries to solve the scheduling problem by applying genetic algorithm and hybrid local search to minimize the makespan. It can be seen that most of previous studies apply heuristic algorithm to solve the hybrid flow shop scheduling in order that the makespan are minimized. However, the heuristic approaches do not typically guarantee the optimal solution for the problem.

This paper intends to present a mathematical formulation to solve a two-stage hybrid flow shop scheduling with the job and time constraints in order to achieve minimum makespan of all customer orders. Then the mathematical formulation is validated by applying the formulation to solve the hybrid scheduling in a case manufacturer.

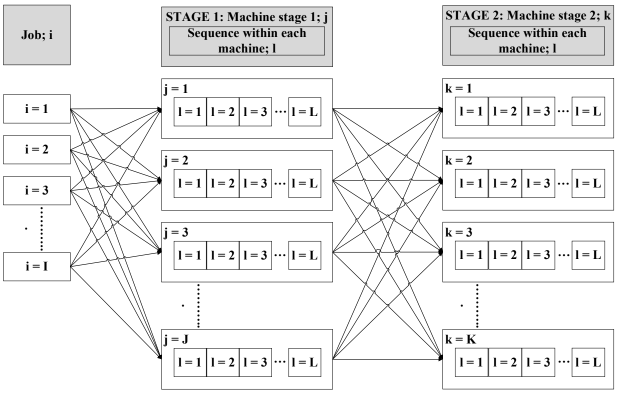
Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

The organization of this paper is as the following. The next section describes the hybrid flow shop scheduling problem. Then the generic mathematical formulation for hybrid flow shop scheduling is presented as a binary integer programming. After that, a numerical example of a case auto part manufacturer is presented to illustrate an application of the formulation to assign jobs to facilities and sequence the jobs. Finally, the conclusion and future research are presented.

PROBLEM DESCRIPTION

This scheduling problem is a two-stage hybrid flow shop. In this flow processes, the jobs can be different types but they must be produced through the first and second stages. There are several non-identical parallel machines in each stage which some jobs cannot be produced at some machines. Also, the processing time of each job at each stage can be different when it is produced at different machines.

This problem studies the assigning and sequencing of jobs for each stage of two flow processes. The job must be accomplished and produced at a particular machines and specific sequence. The job *i* must be produced through the first stage using machine *j* in the sequence *l* and the second stage using machine *k* in the sequence *l* as shown in Figure 1.



Job; i
STAGE 1: Machine stage 1; j
Sequence within each machine; l
STAGE 2: Machine stage 2; k
Sequence within each machine; l
i = 1
i = 2
i = 3
i = I
j = 1
j = 2
j = 3
j = J
k = 1
k = 2
k = 3
k = K
l = 1
l = 2
l = 3
l = L

Figure 1: Hybrid Flow Shop Scheduling

MATHEMATICAL FORMULATION

The section describes the mathematical model formulation for this hybrid flow shop with various constraints**.** The objective is to minimize the makespan in order to find the optimal scheduling consisting of assigning and sequencing**.**

Indices

i job : $i = 1,2,3,..., I$

j machine at j stage 1: j = 1,2,3,....J

k machine k at stage 2: k = 1,2,3,....K

l sequence of job: 1 = 1,2,3,....L

Parameters

P_{ij} processing time of job *i* processed at machine *j* in stage 1

P_{ik} processing time of job *i* processed at machine *k* in stage 2

S_{ij} setup time of job *i* processed at machine *j* in stage 1

S_{ik} setup time of job *i* processed at machine *k* in stage 2

ST_{ijl} starting time of job *i* processed at machine *j* on sequence *l*

ST_{ikl} starting time of job *i* processed at machine *k* on sequence *l*

ET_{ijl} completion time of job *i* processed at machine *j* on sequence *l*

ET_{ikl} completion time of job *i* processed at machine *k* on sequence *l*

C_{ijl} total time of job *i* processed at machine *j* on sequence *l*

C_{ikl} total time of job *i* processed at machine *k* on sequence *l*

Decision Variables

X_{ijl} = 1 if job *i* is assigned at machine *j* on sequence *l*; and 0 otherwise

Y_{ikl} = 1 if job *i* is assigned at machine *k* on sequence *l*; and 0 otherwise

Dependent Variables

$(C_{max})_{ik}$ total time of job *i* processed at the last sequence of machine *k*

Objective Function

The objective function for the scheduling is to minimize the makespan for all jobs.

$$Minimize\ Z = C_{max} \qquad (1)$$

where $C_{max} = Max\{ E_{ikl} \} \quad ; \forall i,\ \forall k$

Constraints

1. Constraints regarding the job assigned:

For each job, there should be only one job to be processed at machine *j* in sequence *l.*

$$\sum_{i=1}^{I} X_{ijl} = 1;\ \forall j, \forall l \qquad (2)$$

Similarly, for each job, there should be only one job to be processed at machine *k* in sequence *l.*

$$\sum_{i=1}^{I} Y_{ikl} = 1;\ \forall k, \forall l \qquad (3)$$

All the decision variables are binary.

$$X_{ijl}, Y_{ikl} \in \{0,1\}; \forall i, \forall j, \forall k, \forall l \qquad (4)$$

If there is any job that cannot be processed at a particular machine, that decision variable equals to 0. For example, if the jobs number 1 to 5 cannot be processed at machine 5 or stage 1, the decision variable X_{i5l} equals to 0:

$$X_{ijl} = 0; \ \forall l, i = 1, 2, 3, 4, 5 \qquad (5)$$

2. Constraints related to time

The starting time of job *i* processed at machine *j* on the first sequence in stage 1 equals to 0.

$$ST_{ij1} = 0; \ \forall i, \forall j \qquad (6)$$

The starting time of job *i* processed at machine *j* in sequence *l* equals to the completion time of its immediate predecessor job *i* processed at machine *j*.

$$ST_{ijl} = ET_{ij(l-1)} \ \ ; \ \forall i, \forall j, l = 2, .., L \qquad (7)$$

The starting time of job *i* processed at machine *k* in sequence *l* must greater than or equal to the completion time of job *i* processed at machine *j* in sequence *l*.

$$Y_{ikl} ST_{ikl} \geq E_{ijl} X_{ijl} \ \ ; \ \forall i, \forall k, \forall l \qquad (8)$$

The completion time of job *i* processed at machine *j* in sequence *l* equals to the summation of starting time of job *i*, set up time and processing time at machine *j*.

$$E_{ijl} = \left(ST_{ijl} X_{ijl}\right) + \left(S_{ij} + P_{ij}\right) X_{ijl} \, ; \ \forall i, \forall j, \forall l \qquad (9)$$

The completion time of job *i* processed at machine *k* on sequence *l* equals to the summation of starting time of job *i*, set up time and processing time at machine *k*.

$$E_{ikl} = \left(ST_{ikl} X_{ikl}\right) + \left(S_{ik} + P_{ik}\right) X_{ikl}; \ \forall i, \forall k, \forall l \qquad (10)$$

NUMERICAL EXAMPLE

The aforementioned mathematical formulation can be applied to the case of an automotive part manufacturer to solve the scheduling problems in the factory. The main processes in the case factory are metal cutting processes which machine automotive parts as per customer orders including washer and washer 5th gear thrust. The production of the two parts into the production flow shop is currently scheduled by assigning the job to the idle machines without scheduling plan. This results in long makespan and tardy jobs. The parts of washer and washer 5th gear thrust are often tardy. This needs to be change by planning the scheduling in advance.

The major machining processes for washers and washer 5th gear thrusts to be studied consists of 2 stages; Cutting and Turning processes. Cutting and Turning contain 5 and 3 machines, respectively. The problem can be depicted in Figure 2.

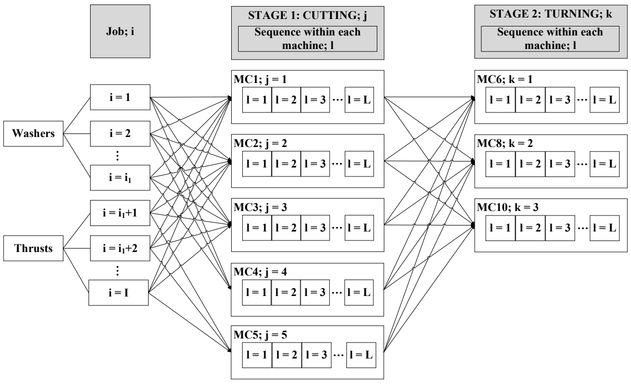


Figure 2: The case scheduling problem

The machines in each stage are interchangeable. There are only little exception regarding the selection of machines as presented in Table 1.

Table 1: Exceptions of machine selection

Stage	Machine no.	Product A Washer	Product B Thrust
Stage 1 Cutting	1	✓	✓
	2	✓	✓
	3	✓	✓
	4	✓	
	5		✓
Stage 2 Turning	6	✓	✓
	8	✓	✓
	10	✓	✓

In order to schedule these jobs, the job orders for the products must be grouped and assigned the job numbers. The example of the case contains 14 jobs; 7 jobs for washers and the rest for washer 5th gear thrust. Jobs of washers are assigned the job numbers 1 to 7, while jobs of thrusts are assigned numbers 8 to 14. The information regarding orders, number of pieces and processing time of each order on each machine are presented in Table 2. The setup time for each machine is 1 hour per a changeover.

Table 2: Jobs and their processing times

Product	Job (i)	Pieces	Processing time (hours)							
			Cutting					Turning		
			1	2	3	4	5	6	8	10
Washer	1	450	7	3	6	6	0	5	3	1
	2	552	3	8	7	4	0	6	3	2
	3	487	7	2	6	3	0	5	3	5
	4	650	10	10	9	2	0	7	4	3
	5	500	8	8	7	2	0	3	5	5
	6	480	7	7	6	2	0	3	5	5
	7	378	6	6	5	1	0	4	2	4
Washer 5th gear thrust	8	442	7	5	2	0	6	5	2	4
	9	398	6	5	5	0	4	5	2	4
	10	375	3	5	3	0	5	2	4	3
	11	426	7	4	2	0	6	2	4	3
	12	500	8	6	3	0	7	2	5	4
	13	415	6	5	2	0	6	4	2	3
	14	387	6	4	2	0	5	4	1	3

Though the machines for each stage are interchangeable, the processing time on different machines are different. For example, Job 1 of 450 washers can be processed on machine no. 1, 2, 3, and 4. It takes 7 hours to complete 450 washers on machine no.1, while it takes only 3, 6, and 6 hours on machine 2, 3, and 4, respectively. Therefore, the selection of machines affects the makespan for all jobs. And it eventually affects the utilization of machines.

The previous scheduling technique used in this case factory yielded 35 hours makespan of scheduling for 14 jobs. The gantt chart to present the makespan of previous scheduling technique can be depicted in Figure 3.

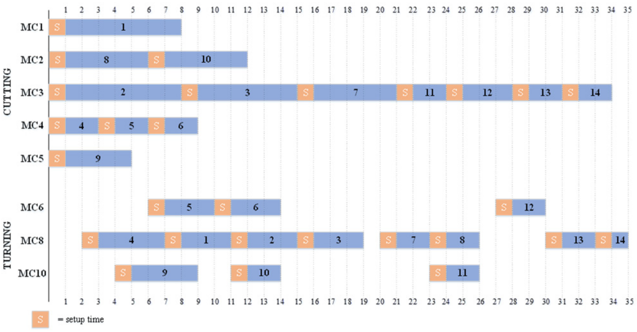


Figure 3: Scheduling applying previous scheduling technique

The aforementioned formulation can be used to shorten the makespan for all the 14 jobs.

Results

The mathematical formulation of the case is applied to the case to make a decision for the scheduling of 7 orders of washers and 7 orders of washer 5th gear thrusts over one-week period (Table 2). The formulation is then solved by the Premium Excel Solver. Using the data of processing time in Table 2, the suitable machine for each job can be selected. The selection of machines yields the total makespan of 23 hours which is the minimum numbers of makespan for the 14 jobs. The result of machine selection can be presented in Table 3.

Table 3: Jobs and selections of machines

Decision variables Job (i)		Processing time (hours)							
		Cutting					Turning		
		1	2	3	4	5	6	8	10
X_{ij}	1	0	*1*	0	0	0	0	0	0
	2	*1*	0	0	0	0	0	0	0
	3	0	*1*	0	0	0	0	0	0
	4	0	0	0	*1*	0	0	0	0
	5	0	0	0	*1*	0	0	0	0
	6	0	0	0	*1*	0	0	0	0
	7	0	0	0	*1*	0	0	0	0
	8	0	0	*1*	0	0	0	0	0
	9	0	0	0	0	*1*	0	0	0
	10	0	0	*1*	0	0	0	0	0
	11	0	0	*1*	0	0	0	0	0
	12	0	0	*1*	0	0	0	0	0
	13	0	0	*1*	0	0	0	0	0
	14	0	0	*1*	0	0	0	0	0

Table 3: Jobs and selections of machines (cont.)

Decision variables Job (i)		Processing time (hours)							
		Cutting					Turning		
		1	2	3	4	5	6	8	10
Y_{ik}	1	0	0	0	0	0	0	0	*1*
	2	0	0	0	0	0	0	0	*1*
	3	0	0	0	0	0	0	*1*	0
	4	0	0	0	0	0	0	0	*1*
	5	0	0	0	0	0	*1*	0	0
	6	0	0	0	0	0	*1*	0	0
	7	0	0	0	0	0	0	*1*	0
	8	0	0	0	0	0	0	*1*	0
	9	0	0	0	0	0	0	*1*	0
	10	0	0	0	0	0	*1*	0	0
	11	0	0	0	0	0	*1*	0	0
	12	0	0	0	0	0	*1*	0	0
	13	0	0	0	0	0	0	0	*1*
	14	0	0	0	0	0	0	*1*	0

1	= Selected machine for the particular job
▨ 0	= Cannot be considered

From Table 3, it can be seen that the mathematical formulation can be used to select proper machines that yield the minimum makespan. Job 1 is to be cut on machine 2 and turned on machine 10; Job 2 is to be cut on machine 1 and turned on machine 10; and so on. Then, the jobs that need to be processed on the same machine need to be to be sequenced. The sequence of jobs in all cutting and turning machines are presented in Table 4.

Table 4: Sequence of jobs in each machine

Job (i)		MC No.	Sequence						
			1	2	3	4	5	6	7
Cutting X_{ij}	1	2	*1*	0	0	0	0	0	0
	2	1	*1*	0	0	0	0	0	0
	3	2	0	*1*	0	0	0	0	0
	4	4	*1*	0	0	0	0	0	0
	5	4	0	*1*	0	0	0	0	0
	6	4	0	0	*1*	0	0	0	0
	7	4	0	0	0	*1*	0	0	0
	8	3	*1*	0	0	0	0	0	0
	9	5	*1*	0	0	0	1	0	0
	10	3	0	*1*	0	0	0	0	0
	11	3	0	0	*1*	0	0	0	0
	12	3	0	0	0	*1*	0	0	0
	13	3	0	0	0	0	*1*	0	0
	14	3	0	0	0	0	0	*1*	0
Turning Y_{ik}	1	10	0	*1*	0	0	0	0	0
	2	10	0	0	*1*	0	0	0	0
	3	8	0	0	0	*1*	0	0	0
	4	10	0	0	0	*1*	0	0	0
	5	6	0	0	0	0	*1*	0	0
	6	6	0	0	0	0	0	*1*	0
	7	8	0	0	0	0	*1*	0	0
	8	8	0	*1*	0	0	0	0	0
	9	8	0	0	*1*	0	0	0	0
	10	6	0	0	*1*	0	0	0	0
	11	6	0	0	0	*1*	0	0	0
	12	6	0	0	0	0	0	0	*1*
	13	10	0	0	0	0	0	*1*	0
	14	8	0	0	0	0	0	0	*1*

Following the sequence of jobs processed at each machine presented in Table 4, the Gantt Chart to present the makespan for all 14 jobs can be depicted in Figure 4.

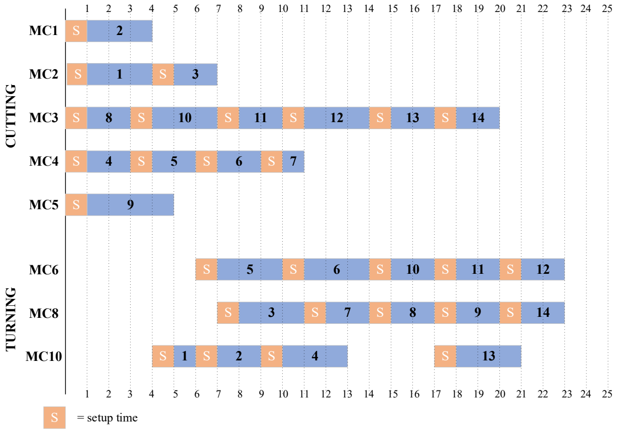


Figure 4: New scheduling applying created mathematical formuation

The result shows the optimal schedule of all the jobs in the two-stage hybrid flow shop by minimizing the makespan. The total makespan of all the 14 jobs is 23 hours. The scheduling in Figure 4 also informs starting, ending, and processing times for each job on each machine. The makespan of scheduling applying the mathematical formulation initiated in this paper is shorter than the 35 hours makespan of previous scheduling method (Figure 3) or 34.29% reduction in the makespan, regardless of longer total time in some jobs.

CONCLUSION

This paper attempts to formulate the scheduling technique for hybrid flow shop production. It can be seen from the case that this mathematical model has a potential to create the optimal schedule for the two-stage flow shop that contains multiple machines in each process. This scheduling methodology concurrently considers the processes of 2 stages to ensure the minimum makespan of all the jobs. It is simply because the final makespan of all the jobs depends on the completion time of the last job on stage 2, whereas the starting time of stage 2 relies on the completion time of stage1. The 2 stages must be considered simultaneously in order to obtain the optimal solution for the scheduling. This mathematical model can be considered useful since it can assign jobs to proper machines and sequence the jobs for each machine to create the optimal production schedule. For future research, this scheduling technique can be expanded to incorporate 3 or more stage production which is typical cases in many manufacturing industry, especially the automotive part industry.

REFERENCES

Choi, H.S., Kim, H.W., Lee, D.H., Yoon, J., Yun, C.Y. and K.B. Chae. 2008. “Scheduling algorithms for two-stage reentrant hybrid flow shops: minimizing makespan under the maximum allowable due dates,” *International Journal Advanced Manufacturing Technology*, Vol.42, pp. 963–973.

Mallikarjuna, K., Rajashekarpatil, Naween Kumar, H.M, and Hanumanthappa, S. 2013. “Performance heuristic method over traditional method for solving the job sequencing problem”, International Journal of Engineering Research & Technology, Vol. 3, pp.3095 – 3099.

Puck-In, A. 2014. “Solving sequence of job scheduling problem by Genetic Algorithm with local search”, *Industrial Technology Lampang University Journal*, Vol. 7, pp. 111 – 126.

Vignier, A., Dardilhac, D., and Proust, C. 1996. “A branch and bound approach to minimize the total completion time in a k-stage hybrid flowshop”, *Proceedings od 1996 IEEE Conference on Emerging Technologies and Factory Automation*, Vol. 1, pp. 215 – 220.

Watanakich, P., 2001, Scheduling for a Two-Stage Hybrid Flow Shop with Machine Setup Time, *Master of Engineering Thesis, Industrial Engineering*, Faculty of Engineering, Kasetsart University, pp. 1-55.

Wong, W.K., Chan, C.K. and Ip, W.H. 2001. “A hybrid flowshop scheduling model for apparel manufacturer, *International of Clothing Science and Technology*, Vol. 13, pp. 115 – 131.

AUTHOR BIOGRAPHIES

TUANJAI SOMBOONWIWAT is an Associate Professor in the Industrial Management section, Department of Production Engineering Faculty of Engineering, King Mongkut’ s University of Technology Thonburi, Thailand. She received her M. Eng. in Industrial Engineering from Chulalongkhorn University, Thailand and Ph. D. in Industrial Engineering from Corvallis, Oregon State University, USA. Her research interests include green supply chain and logistics, business process and applications of operations research. She can be reached at her e-mail address: tuanjai.som@kmutt.ac.th.

TUANGYOT SUPEEKIT is an Assistant Professor in the Department of Industrial Engineering, Faulty of Engineering, Mahidol University in Thailand. He received his Master of Engineering Management from University of Technology, Sydney. His research interests include business process improvement, performance measurement and logistics and supply chain management. He can be reached at his email address: tuangyot.sup@mahidol.edu.

CHATKAEW RATCHARAK was a graduate student at the Department of Production Engineering, King Mongkut’ s University of Technology Thonburi, Thailand. Her research interests are in supply chain management and applied operations management. Her e-mail address is: ratcharak_chatkaew@hotmail.com.

INTEGRATED MODELLING OF COMPLEX PROCESSES ON BASIS OF BPMN

Semyon A. Potryasaev
Russian Academy of Science,
Saint Petersburg Institute of Informatics and Automation
39, 14 Linia, VO, St.Petersburg, 199178, Russia
E-mail: semp@mail.ru

KEYWORDS
integrated modeling, business-process modeling, logic-dynamic models, pro-active control, integrated modeling automation.

ABSTRACT

The description of designed complex of analytical logic-dynamic models supported by corresponding complex of dynamic simulation models developed on the basis of BPMN is offered in this article.

INTRODUCTION

Modern enterprises in such high-technology industries as shipbuilding, aerospace sector, fuel and energy system and the like, represent complex objects (CO) functioning in dynamically changing environments. Specified complexity is caused by the increase in number of composing subsystems and objects, and, respectively, rapid growth in number of internal links that reveals itself in such aspects as structural and functional complexities, the complexity of the choice of behavior, the complexity of decision making, the complexity of development and the complexity of modeling (Sokolov et al. 2006).

INTEGRATED MODELING OF COMPLEX OBJECTS FUNCTIONING

Studying of CO mentioned above requires joint use of diverse models and combined methods, and in some cases methodological and systematic basic concepts, multiple theories and scientific disciplines and conducting relevant interdisciplinary research. In this case to increase the level of relevance and reliability of the prognoses for the development of existing and projected CO it is necessary to carry out preemptive modeling and multi-variant forecasting of different scenarios for the implementation of the life cycles of the objects under review based on the concept of integrated modeling (IM).

Hereafter by IM of CO of any nature (a particular case of which is high-technology enterprises) we will mean the methodology and technologies of multiple-model description of the specified objects and combined use of methods, algorithms and techniques of multi-criteria analysis, synthesis and choice of the most preferable managerial decisions in connection with creation, use and development of the considered objects under various conditions of dynamically changing external and internal environments (Sokolov et al. 2006). The combined use of diverse models, methods and algorithms allows both compensating their actual drawbacks and limitations and, simultaneously, strengthening their advantages.

Moreover, IM of manufacturing processes (MP) of an enterprise is a step to its pro-active control (PaC). Unlike traditionally used in actual practice reaction control focused on rapid response and subsequent prevention of incidents, pro-active control involves prevention of their occurrence by creating in the relevant control system fundamentally new predictive and preemptive capabilities (such as parametric and structural model adaptation for past, present and future events) while forming and implementing control activities based on counteracting not consequences but reasons causing possible abnormal, emergency and critical situations.

Alongside with the set of the described advantages provided by IM of MP there appears a number of problems associated with its use. Thus, the first and, perhaps, the main distinctive feature of IM of MP is the necessity to effect coordination (conditioning) in the modeling process at the concept, model and algorithm, information and program levels of the models, methods and algorithms used. An emergent effect from IM can only be achieved while carrying out profound and reasonable conditioning of specific models based on the principles of coordination of decomposed models and multiple-model complexes (Trotsky and Gorodetsky 2009).

The second problem is the analysis of fulfillment of manufacturing program that is estimating the possibility of achieving preset MP quality indices considering existing space-temporal, technical, technological and resource constraints. The third problem involves the necessity to use widely modern automation technology for modeling at all stages of IM implementation. Otherwise IM will not be possible because of considerable consumption of time, funds and other resources that are to be allocated when unified automation technology is not available. Also, the third problem involves the stage of basic data input that remains exclusively labor-consuming even if automation technology is available.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Moreover, while solving problems of structural and functional synthesis of CO of different classes in the framework of IM one may face a new challenge (Sokolov et al. 2006):

- large dimensionality and non-linearity of models describing the structure and variants of functioning of elements and subsystems of complex objects;
- necessity for constructive consideration in the models of uncertainty factors caused by the influence of external environment on a complex object;
- necessity of performing of multi-criteria optimization on a multiple-model complex.

OVERVIEW OF MODERN MEANS OF DESCRIPTION AND AUTOMATION PROCESSES FOR INTEGRATED MODELING OF COMPLEX OBJECTS

To overcome the listed difficulties to the present moment there were developed numerous instruments and automation environments of simulation modeling, such as GPSS, AnyLogic, BPsim, PowerSim, Simplex, Modul Vision, Triad.Net, CERT, ESimL, Simulab, NetStar, Pilgrim, MOST, KOGNITRON, etc. (Trotsky and Gorodetsky 2009). Recently the above mentioned automation environments were supplemented with intellectual information technologies (neural networks, multi-agent systems, fuzzy logic, technologies of evolutionary modeling, etc.). The absence of generally accepted mechanisms of conditioning of the models used both at technical and semantic levels blocks the joint application of diverse set of these instruments in the framework of IM.

Brief mention should be made on the capabilities of modern means of description and automation of the processes of IM of CO. It is known that at the initial stage of application of analytical-simulation modeling (ASM) for CO it is reasonable to make its description with the use of certain specification.

Currently there exists a dozen of the most popular languages for process description.

Petri net model was one of the first formal models designed for specification of process models. Weak expressive power and means for operational semantics of this resulted in the fact that in practice this model is mainly used as a basis for other languages (Laue and Müller 2016).

The group of standards of IDEF (Integrated DEFinition) amounts to 15 separate directions but only IDEF0 (functional modeling), IDEF1 (modeling of information flows) and IDEF3 (documenting of technological processes) came to be widespread.

The most common UML diagrams are focused more on the description of software architecture and support of object-oriented approach than on the description of technological and logistic processes.

The eEPC standard (Extended Event Driven Process Chain) well suits for the description of resource flows and flows of events but is not appropriate for the description of technological and logistic processes that use a great amount of different resources and means. Along with eEPC, BPMN (Business Process Model and Notation) is assigned for the description of the diagrams of business-processes familiar to both technical specialists and business users but it is of more interest in the context of integrated modeling of manufacturing processes. Among the advantages of BPMN the following ones should be mentioned: set of the applied primitives combines the advantages of other notations and allows to represent the models of distributed processes; provides a wide range of capabilities for formal representation of components of complex processes (Trotsky and Gorodetsky 2009).

The processes described in BPMN can be used to carry out both analytical and simulation modeling with application of corresponding software environments. It is referred to the extension of notation - BPSim standard (Business Process Simulation Interchange Standard). Unfortunately, currently there are no programming solutions that fully support this standard (Laue and Müller 2016). The application of IM of business processes assisted by BPsim may possibly increase the relevance of this standard.

COMPLEX OF ANALYTICAL LOGIC-DYNAMIC MODELS OF COMPLEX OBJECTS FUNCTIONING

Referring to analytical modeling one should take into consideration that parameters and structures of a complex object are constantly changing at different stages of its life cycle due to various reasons: objective and subjective, internal and external, etc. Mentioned in article (Sokolov et al. 2006) peculiarity is defined as structure dynamics of complex technical objects (CTO). In the same work it was shown that in order to maintain, increase or restore the level of working efficiency and capacities of the system it is necessary to control their inherent complexity. In particular, it is required to maintain control over their structures.

Taking into account above mentioned, it is offered to choose a dynamic alternative system-related graph with controllable structure as a basic mathematic structure, with the help of which it is possible to describe multiple-model structure dynamics of CO. Analysis of the possible options for creation of analytical models of control over CO structure dynamics showed that during the process of their creation, it is worthwhile to focus on the class of logic-dynamic models (LDM).

A significant theoretical and practical experience is accumulated in this area, and the results are provided in a number of works (Potriasaev 2006, Potriasaev et al. 2008).

Within the framework of the developed multiple-model complex, the following basic content of LDM is suggested: logic-dynamic models of control over operations, flows, resources, operation parameters, and structures.

Formally, the designed generalized model of enterprise structure dynamics control (SDC) represents finite-dimensional non-stationary non-linear differential dynamic system with variable area of acceptable control actions with partially fixed boundary conditions at the initial and finite timepoints. For the purpose of this article, the objective of the enterprise SDC may be formulated as the task of searching optimal controls over this specified generalized dynamic model.

On the basis of above mentioned particular dynamic models a generalized LDM of enterprise functioning processes was formed:

$$M = \left\{ \begin{array}{l} \vec{u}(t) \mid \dot{\vec{x}} = \vec{f}(\vec{x}, \vec{u}, t); \\ \vec{h}_0(\vec{x}(t_0)) \le \vec{O}; \vec{h}_1(\vec{x}(t_f)) \le \vec{O}; \\ \vec{q}^{(1)}(\vec{x}, \vec{u}) = \vec{O}; \vec{q}^{(2)}(\vec{x}, \vec{u}) \le \vec{O}; \end{array} \right\} \qquad (1)$$

where $\vec{x} = \left\| \vec{x}^{(o)^T} \vec{x}^{(r)^T} \vec{x}^{(s)^T} \right\|^T$; $\vec{u} = \left\| \vec{u}^{(o)^T} \vec{u}^{(r)^T} \vec{u}^{(s)^T} \right\|^T$ – generalized vectors of status and control over the enterprise manufacturing processes (index *o* refers to the model of the basic control over operations, *r* – to the model of control over resources, *s* – to the model of control over flows); $\vec{h}_0, \vec{h}_1$ are common vector-functions that set boundary conditions for vector $\vec{x}$ at the timepoints $t = t_0$ and $t = t_f$; $\vec{q}^{(1)}, \vec{q}^{(2)}$ are vector functions that set basic space-temporal, technical and technological constraints imposed on the process of the enterprise functioning.

Also, vector quality index of planning quality metrics with such components as vectors of particular indices of quality of programmed control over operations, resources and structures is offered:

$$\vec{J}_{gen} = \left\| \vec{J}^{(o)^T} \vec{J}^{(r)^T} \vec{J}^{(s)^T} \right\|^T . \qquad (2)$$

Problem and formal description of above mentioned task as well as the methods of its solution are specified in works (Sokolov 1992, Potriasaev et al. 2008).

For brevity sake a simplified variant of such formalization is given in this article. In this case the main technological and technical constraints defining the priority of serial-parallel business operation execution within the framework of the proposed complex of models (1) can be represented in the following way:

$$\Delta = \Bigg\{ \mathbf{u} \mid \dot{x}_i = \sum_{j=1}^{m} u_{ij}; \sum_{i=1}^{n} u_{ij}(t) \le 1; \sum_{j=1}^{m} u_{ij} \le 1; u_{ij}(t) \in \{0,1\};$$

$$t \in (t_0, t_f] = T; \quad x_i(t_0) = 0; \; x_i(t_f) = a_i;$$

$$\sum_{j=1}^{m} u_{ij} \left[\sum_{\alpha \in \Gamma_{1i}^-} (a_\alpha - x_\alpha(t)) + \prod_{\beta \in \Gamma_{i2}^-} (a_\beta - x_\beta(t)) \right] = 0;$$

$$i = 1,\ldots, n; \; j = 1,\ldots, m \Bigg\}, \qquad (3)$$

where *xᵢ(t)* is a variable characterizing operation completion status at the timepoint *t*; a_i is the set value of specified operation completion; $u_{ij}(t)$ is a control action taking on value 1, if operation D_i is completed using enterprise resource B_j, 0 – otherwise; $\alpha \in \Gamma_{1i}^-$, $\beta \in \Gamma_{2i}^-$ is a set of numbers of operations, directly precedent and technologically connected with operation D_i with the help of logical operations "AND", "OR" (alternative "OR"), *T* is time interval during which enterprise functioning is examined; t_0, t_f are the initial and finite timepoints. It is necessary to emphasize that the very recording of these constraints allows to refer the designed model (1) to the class of logic-dynamic models.

The most distinctive feature of the offered multiple-model complex is unification of control and flow modeling at the constructive level. Thus, for example, the model of programmed control over operations M_o affects on the model of programmed control over resources M_r with the use of control $\vec{u}^{(o)^T}$. In its turn, programmed control $\vec{u}^{(o)^T}$ has an impact on the model of flows control M_s through corresponding constraints. In its turn, the flow model M_s via boundary conditions determines initial timepoints when to start operation execution.

Due to the use of listed properties of the offered logic-dynamic model with its capabilities, the above-mentioned problems inherent to IM of complex object functioning can generally be solved at the constructive level.

The originality and the main advantages of the developed complex of analytical LDM complemented with the corresponding complex of dynamic simulation models consist of below mentioned points. Firstly, unlike earlier offered approaches to formal description of the considered class of logic-dynamic models of complex object control (Zimin and Ivanilov 1971), all basic space-temporal, technical and technological constraints having absolutely non-linear character are taken into account not while setting differential equations describing the dynamics of the relevant processes but while forming an area of acceptable control actions values. In addition offered dynamic interpretation of the complex of carried out operations allows to substantially reduce the dimensionality of the current tasks of optimization defined by the number of independent ways in the generalized graph of fulfilled works that form existing front of operations ready for execution. Secondly, constructive recording of nonstationarity of complex objects functioning (in this particular case manufacturing enterprise, for example, a shipbuilding yard) is carried out in the designed model on the basis of introduction of multi-dimensional dynamic matrix functions, such as "contact potential" and "potential of availability" (Potriasaev 2006). Thirdly, consideration of factors of uncertainty in the framework of the considered class of LDM describing ctructure dynamic control of control objects makes provision for adaptation of parameters and models structures, algorithms of structural dynamics control of CO with relation to previous, current and possible future conditions of control objects on the basis of multi-variant scenario prognosis and complex preemptive analytical-and-simulation modeling.

It is necessary to emphasize once again that with the application of the unified language the use of the offered variant of formalization of logic-dynamic control models of CO allows to describe both the processes of application planning of CO and the processes of plan completion, the processes of multi-variant prognosis of implementation of different scenarios of proactive control of CO.

Finally, in the framework of the offered formalization there were developed several approaches to the solution of the problem of multi-criteria optimization of SDC of CO based both on orthogonal projection of target set on the extensible set of attainability of dynamic model (1) (Sokolov et al. 2006) and on methodology of creation and use of integral index of quality and efficiency of CO functioning based on combined use of mathematic tools of fuzzy logic and experimental design theory (Adler et al. 1976).

CORRELATION OF THE ELEMENTS OF ANALYTICAL MODEL AND CONCEPTS OF BPMN

The considered complex of analytical logic-dynamic models relies on the corresponding conceptual model that includes the following basic notions: “operation”, “resource”, “objective”/”task”, “flow”, “structure”. Their detailed description was given earlier in work (Sokolov et al. 2015). Using the concepts listed above, one can set different classes of relations that in their turn are defined by those space-temporal, technical, technological, material, informational, and energy constraints, etc. being typical for specific subject area.

Detailed consideration of BPMN 2.0.2 and particularly the specific section “BPMN Process Execution Conformance” allows to make a conclusion about the possibility of using this notation with the purpose of formation of the set of basic data for the analytical model described above (see Table 1).

Table 1: Correlation of Elements of Analytical Model (AM) and BPMN

Concept of AM	Concept of BPMN	Data Available in BPMN	Extensible data for AM
Operation	Task	Identifier, name, resources used	Target volume of operation, interruption feasibility
Resource	Resource	Identifier, name, supply, cost of single use, cost of use per minute	Performance
Goal	–	–	Status variable values at finite timepoint
Flow	Sequence Flows / Message Flows	Identifier, name, input source, output source	Maximum Flow Rate
Structure	Pool	Identifier, name, resource scope, scheduled availability	Total output
Operation	Task	Identifier, name, resources used	Target volume of operation, interruption feasibility
Resource	Resource	Identifier, name, supply, cost of single use, cost of use per minute	Performance

As follows from Table 1, in the basis of BPMN there are not enough declared attributes to carry out analytical modeling. At the same time, BPMN originally has been created as an extensible language that allows to freely supplement the model description with necessary attributes without losing backward compatibility with its runtime environment.

Recorded in BPMN complex process with all necessary additional attributes can be executed in the earlier-developed environment of analytical modeling based on dynamic interpretation of the processes of carrying out the operations and distributing the resources of complex objects.

Thus, when using the model of complex process described in the extensible BPMN it is possible to simultaneously perform simulation and analytical modeling that allows to speak about conditioning of the model at the conceptual, model-algorithmic, informational and program levels.

Moreover, application of ASM allows to analyze more profoundly the models of complex processes described in the extensible BPMN, i.e. the application of the categories of control theory to the analysis of actual manufacturing tasks.

The advantage of extension of BPMN application area consists in considerable reductionof labor intensity of basic data input while conducting analytical modeling of actual manufacturing systems. For example, this refers to dozen thousands of variables and thousands of constraints when considering mathematical models for manufacturing processes in shipbuilding industry. While speaking on specified advantages, it is necessary to mention, firstly, that BPMN is focused on the simplification of data input and their visualization due to availability of graphic representation and limited number of concepts. Secondly, many enterprises already have manufacturing processes described in this notation; and consequently preparation of basic data for analytical modeling is limited to introduction of some additional concept attributes. Thirdly, the area of automatic creation of diagrams described in the form of a text form is being developed (Deeptimahanti and Sanyal 2009). For example, a number of works informing on successful implementation of the method of creation of BPMN on the basis of sequence of actions description in the form of text are known (Fabian et al. 2011, Henrik at al. 2012). In addition, there appears a possibility to apply modern technologies of modeling automation at all stages of complex modeling implementation.

SAMPLE OF PRACTICAL IMPLEMENTATION OF SHIPBUILDING ENTERPRISE INTEGRATED MODELING

The proposed in this article approach was used while carrying out the research work devoted to the investigation and selection of methods and algorithms of solving tasks of integrated and simulation modeling as well as multi-criteria analysis of the manufacturing systems in shipbuilding industry. BPMN was used to perform IM of MP including technological and auxiliary manufacturing processes. In Figure 1 one can find an extract of specified processes description.

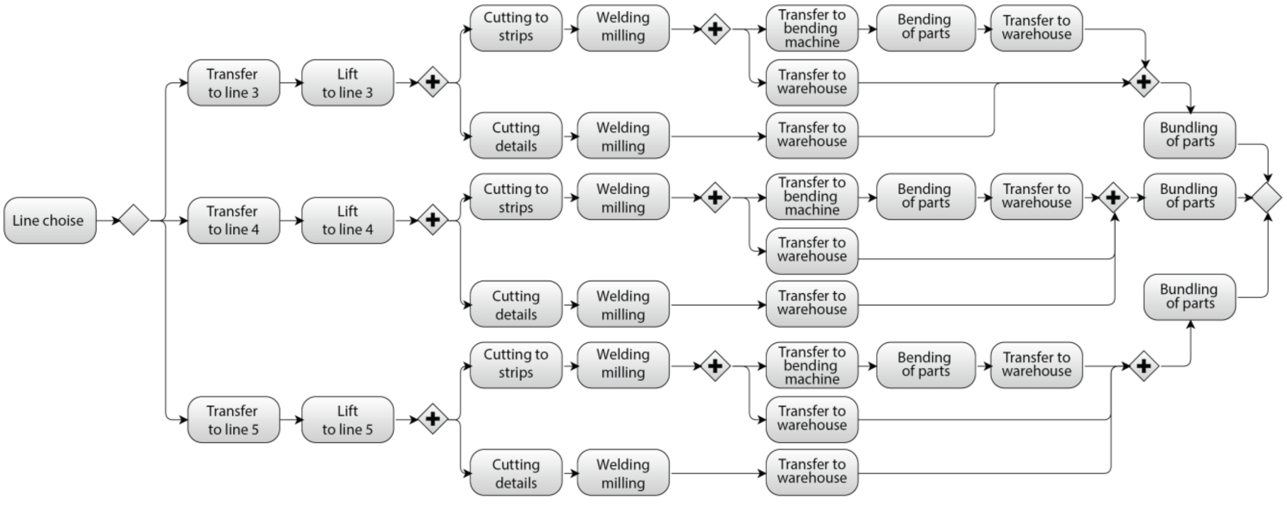


Figure 1: Fragment of Manufacturing Process in BPMN

Agreed use of simulation and analytical logic-dynamic model on the basis of BPMN application allowed to extend the set of calculated indices of shipbuilding enterprise functioning and to make computation, multi-criteria evaluation and analysis of structure dynamics of a shipbuilding enterprise under different variants of input effect.

It is important to emphasize once again that designed special software of IM of CO using BPMN represents unified modern automation tool for modeling built on service-oriented architecture and web-technologies.

CONCLUSION

Finally, it may be concluded that considering the problems of IM of MP of an enterprise in overall context of SDC allows, firstly, to directly connect those common goals for achieving of which the functioning of the enterprise is oriented with the goals that are executed in the course of manufacturing processes control. Secondly, to reasonably define and choose relevant sequence of the problems to be solved and operations to be executed; in other words, to synthesize technology of control over enterprise manufacturing processes. And, thirdly, to consciously find compromise solutions while distributing limited resources of an enterprise.

ACKNOWLEDGMENT

The research described in this paper is partially supported by the Russian Foundation for Basic Research (grants 15-07-08391, 15-08-08459, 16-07-00779, 16-08-00510, 16-08-01277, 16-29-09482-ofi-i, 17-08-00797, 17-06-00108, 17-01-00139, 17-20-01214), grant 074-U01 (ITMO University), project 6.1.1 (Peter the Great St.Petersburg Politechnic University) supported by Government of Russian Federation, Program STC of Union State “Monitoring-SG” (project 1.4.1-1), state order of the Ministry of Education and Science of the Russian Federation №2.3135.2017/K, state research 0073–2014–0009, 0073–2015–0007.

REFERENCES

Adler, Yu.P., Markova, E.V., Granovsky Yu.V. 1976. “Experiment Planning while Searching for Optimal Conditions”. Moscow, Nauka. 280p. (in Russian)

Deeptimahanti D. K., Sanyal R. 2009. “An Innovative Approach for Generating Static UML Models from Natural Language Requirements”. Advances in Software Engineering. Communications in Computer and Information Science. Vol. 30. Pp. 147-163.

Fabian Friedrich, Jan Mendling, Frank Puhlmann. 2011. “Process model generation from natural language text”. Proceedings of the 23rd international conference on Advanced information systems engineering, June 20-24, 2011, London, UK.

Henrik Leopold, Jan Mendling, Artem Polyvyanyy. 2012. “Generating Natural Language Texts from Business Process Models”. Proceedings of the 24th international conference on Advanced Information Systems Engineering, June 25-29, 2012, Gdańsk, Poland.

Laue, R. & Müller, C. 2016 “The Business Process Simulation Standard (BPSIM): Chances and Limits”. Proceedings of

30th European Conference on Modelling and Simulation, ECMS. Pp. 413.

Potriasaev, S.A. 2006. “Integrated Planning of Reconfiguration in Disaster-Resistant Systems”. Logbook of Information of Higher Schools. Instrument Making. Vol.49, N.11. Pp. 54 –59. (in Russian)

Potriasaev, S.A. 2006. “Statement and Solutions of Problem of Planning of Fault-Tolerant Computer Systems’ Reconfiguration”. SPIIRAS Proceedings. S.-Petersburg. Institute of Computing and Automation. Under the general editorship of R.M. Yusupov. Iss. 3, Vol.2. 406p. ISBN 5-02-025122-4. (in Russian)

Potriasaev, S.A., Petrova, I.A., Ikonnikova, A.V., Sokolov, B.V. 2008. “Dynamic Model Of Complex Planning of Information System Modernization and Functionaing”. Instrument Engineering. N 11. (in Russian)

Sokolov, B.V. 1992. “Complex Operations Scheduling and Structure Control in Automation Control Systems of Active Mobile Objects”. RF Department of Defense. (in Russian)

Sokolov, B.V., Okhtilev, M.Yu., Yusupov, R.M. 2006. “Intellectual Technologies for Monitoring and Control of Structure-Dynamics of Complex Technical Objects”. Moscow, Nauka, 410 p. (in Russian)

Sokolov, B.V., Zelentsov, V.A., Brovkina, O., Mochalov, V.F., Potryasaev, S.A. 2015. “Models Adaptation of Complex Objects Structure Dynamics Control”. Advances in Intelligent Systems and Computing 348. Pp. 21. (in Russian)

Trotsky, D.V., Gorodetsky, V.I. 2009. “Scenario-based Knowledge Model and Language for Situation Assessment and Prediction”. SPIIRAS Proceedings. Iss 8. Pp. 94-127. (in Russian)

Zimin, I.N., Ivanilov, Yu.P. 1971. “Solving of Network Planning Problems via a Reduction to Optimal Control Problems”. Journal of Calculus Mathematics and Mathematical Physics, Vol. 11, N 3, pp.632-631. (in Russian)

AUTHOR BIOGRAPHIES

Semyon A. Potryasaev graduated from the Baltic State Technical University “VOENMEH” with a degree of control systems engineer and Moscow Institute of International Economic Relations as an economist in finance and credit. Successfully defended his PhD thesis in 2004 at Saint Petersburg Institute of Informatics and Automation of the Russian Academy of Science (SPIIRAS).

Currently works as a senior Researcher at Saint Petersburg Institute of Informatics and Automation of the Russian Academy of Science (SPIIRAS). Previously worked in commercial educational centres as trainer and consultant on information security and web technologies. Research interests: applied research in mathematical modelling, optimal control theory, mathematical models and methods of support and decision making in complex organization-technical systems under uncertainties and multicriteria.

E-mail: spotryasaev@gmail.com

MODELLING AND SIMULATION OF PUBLIC TRANSPORT SAFETY AND SCHEDULING ALGORITHM

Anna Beinarovica
Industrial Electronics and
Electrical Engineering Institute
Riga Technical University
Riga, Latvia
anja19892@inbox.lv

Mikhail Gorobetz
Industrial Electronics and
Electrical Engineering Institute
Riga Technical University
Riga, Latvia
mihails.gorobecs@rtu.lv

Anatoly Levchenkov
Industrial Electronics and
Electrical Engineering Institute
Riga Technical University
Riga, Latvia
anatolijs.levcenkovs@rtu.lv

KEYWORDS
Transport Safety, Scheduling, Algorithm, Energy Saving, Optimization, Database, Intelligent Systems, Logic, Mathematical Model

ABSTRACT

The main objective of the transport operations is a safe transportation process with minimal energy consumption. There are various methods for gaining these tasks. This paper discusses one of the possible problem solving - proper planning of public transport operations. The main goal of the research is to develop the adaptive algorithms for transport control and optimization. The main task of the target function is to minimize total downtime at intermediate stations. The specific unique Web-based computer model was developed. It uses Web database for simulation data storage and processing. Simulation results shows the workability of the developed algorithm.

1. INTRODUCTION

Safety is one of the first priority tasks in transport domain. For example, in the report of European Railway Agency (European Railway Agency 2010) the most part of railway accidents and crashes caused by the human factor.

Nowadays a lot of scientists make researches about energy saving process (Staņa et al. 2014) and time planning on railway transportation process (Nikolajevs and Mezitis 2016), that proves the actuality of the chosen topic. There are some studies made separately in fields of scheduling, planning and transport systems. For example, in (Shui et al. 2012) the solution of optimization of public bus timetable is proposed using a cultural clonal selection algorithm based approach is proposed to obtain a vehicle scheduling solution. However, scheduling for public electric transport energy saving and schedule overlapping is not so well studied.

In the previous works, the system and the algorithm for anti-collision system reducing human factor has been developed and tested (Levchenkov and Gorobetz 2014) Also intelligent transport safety system was investigated and designed by paper authors (Levchenkov et al. 2012). This research is a development of the intelligent public transport safety system also, but it is based on the scheduling theory (Conway et al. 2003).

Transport safety is actual problem and the control instruction infringements and human factor may cause crashes (Matsumoto et al. 2015). Schedule planning is one of the opportunities for reaching safety and energy efficient public transport movement process (Alps 2012; Alps et al. 2016) The main advantage of the proposed system is transportation safety performance without human being. At present time special worker – dispatcher or other person makes the biggest part of public transport timetables. Proposed system saves human labor by easy planning opportunity, reduces possibility of unsafety movements, minimize energy consumption by reducing stops and downtimes at intermediate points.

2. PROBLEM FORMULATION

The main purpose of the proposed system is still to minimize amount of collisions, safe the amount of energy consumption and avoid traffic jams, but new system is designed for public transport, because public transport vehicles have timetable.

The goal of current research is to develop the algorithm for optimizing timetable, to improve safety and prevent collisions in public transport, to use fast and comfortable transportation planning process, which could save dispatcher or work planner time and labor, to reduce energy consumption.

The following tasks are defined and solved:
1) to define the structure and functions of the system.
2) to develop the mathematical model and target function for optimizing the transportation process.
3) to develop the adaptive algorithms of the system functions for optimization
4) to develop the database for the mathematical model.
5) to develop the computer model and simulate the developed algorithm to compare the results before and after optimization.

3. PROPOSED STRUCTURE AND MAIN FUNCTIONS

The general structure of the system is presented on the Figure 1.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Scheduling algorithm proposed in this study can be used in different transportation tasks. In this study is described the example of algorithm use in railway transportation process.

Proposed system structure consists of:

- Tdep – preferred departure time from the first station.
- Route - road between stop points, stations. On the one route at the same moment only one transport vehicle can run. After transport vehicle intersects stop point, previous route can be engaged by another transport vehicle. This type of running can lead to no traffic jam and equable distance between transport vehicles
- TD – train device, which consists of an algorithm, is connecting to the database on the server by using an internet and calculates optimal schedule for the train.
- Algorithm – method of analyzing the route busyness, possibility of route intersection and optimizing the transportation process by reducing routes intersections and regulating distance between the transport vehicles.
- Optimal schedule – schedule, which meets the conditions of the preferred departure time, safety transportation process, minimization of downtime and reduced energy consumption.
- Internet – needs to make the connection to the server.
- Coordinating algorithm – needs to refresh the information about the timetables in the database and to connect other intelligent devices.
- DB – database, which consists of:

- Schedule – public transport timetables.

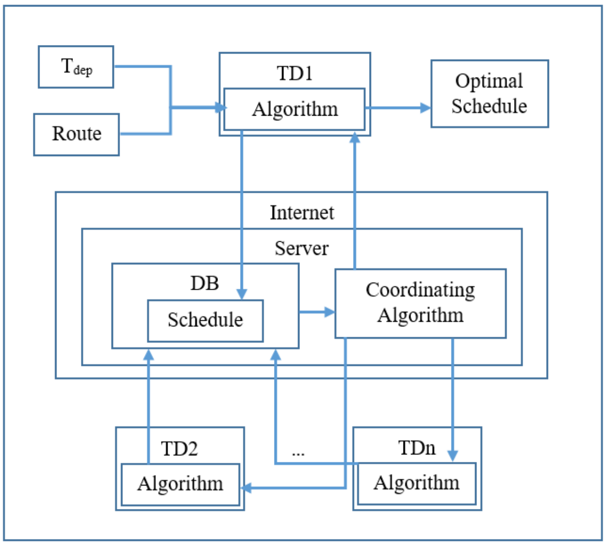


Figure 1. Structure of the system

Preferred departure time and selected route are on the entrance. Train device TD connects to the server. The nearest possible departure time is found and checked in the coordinating algorithm. Other train devices TD might be connected to the server as well and can chose their optimal schedules also. Time interval for the transportation process needs to be checked, to prevent the situation when two trains chose the same schedule. After the free schedule is chosen, the algorithm of train device TD is checking the possibility of downtime and minimizes it by finding in the database another departure time.

4. MATHEMATICAL MODEL AND TARGET FUNCTION

The mathematical model is represented with following sets:

- SP = (SP$_1$, SP$_2$, … , SP$_n$) – set of stops;
- R = (SP$_1$SP$_2$, SP$_2$SP$_3$, … , SP$_n$-1,SP$_n$) – set of routes;
- $U \subset (U^1,\dots,U^n)$ - set of transport units as a subsets of different types, where for different transport safety task it could be:
  - $U^1 = (u_1^1,\dots,u_{n_1}^1)$ - subset of railway transport units;
  - $U^2 = (u_1^2,\dots,u_{n_2}^2)$ - subset of buses;
  - $U^3 = (u_1^3,\dots,u_{n_3}^3)$ - subset of trams etc.
- Schedule A = Σ^{U}_{i} = (t$_1$, t$_2$, … t$_j$, … t$_m$) – time moments of arrival of j-th vehicle if i-th SP is included in vehicles route.
- Schedule D = Σ^{U}_{i} = (t$_1$, t$_2$, … t$_j$, … t$_m$) – time moments of departure of j-th vehicle if i-th SP is included in vehicles route.

Target function:

$$T = \sum_{i=2}^{n}(t_i^s - t_{i-1}^b) \rightarrow min \qquad (1)$$

Where,
t_i^s– time of departure from the station;
t_{i-1}^b– time of arrival to the station.

Target function can be explained by following time axes:

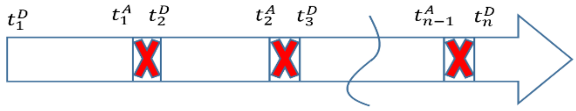


Figure 2. Explanation of the target function

Where:
t_1^D - departure from the first SP time, s;
t_1^A - arrival into second SP time, s;
t_2^D - departure from the second SP time, s;
t_2^A - arrival into third SP time, s;
t_3^D - departure from the third SP time, s;
t_{n-1}^A - arrival into n-1th SP time, s;
t_n^D - departure from the nth SP time, s.

The main task of the target function is:

$$T = (t_2^D - t_1^A) + (t_3^D - t_2^A) + \dots + (t_n^D - t_{n-1}^A) \rightarrow min \qquad (2)$$

Downtime at intermediate stops reduced to zero. This will lead to non-stop running through the intermediate stops and energy consumption will be decreased (Beinarovica et al. 2017).

5. ALGORITHM OF SAFE SCHEDULING

The steps of the algorithm are described for HMI (human-machine interface). The algorithm is presented in 33 steps.

STEP 1. Connecting to the server and database.
STEP 2. Departure point SP_1 and destination point SP_n, time of departure t^D_1, the buttons "Check" output.
STEP 3. Data checking.
STEP 4. Storing of the set parameters SP_1, t^D_1, SP_n.
STEP 5. Spans ID finding SELECT * FROM `spans`.
STEP 6. Departure point SP_1 and destination point SP_n names output. Spans ID output.
STEP 7. Number of j-th vehicle output from the table "Timetables", time moment of arrival and departure of j-th vehicle output.
STEP 8. Additional limitation "Closed" output.
STEP 9. Calculation of the remaining spans included in the route SP_1SP_n.
STEP 10. The buttons "Continue" output. It clicking check.
STEP 11. The direction of movement checking.
STEP 12. Spans ID finding.
STEP 13. Departure point SP_1 and destination point SP_n names output.
STEP 14. Unlimited array $arr create.
STEP 15. Additional limitation entering into an $arr array.
STEP 16. All the chosen vehicles entering into an $arr array.
STEP 17. Saved timetables entering into an $arr array.
STEP 18. Sorting an $arr array.
STEP 19. Setting a time interval from $date_now to $date_p.
STEP 20. Intervals entered in an array separation on the beginning of time interval 'time from' and the end of the time interval 'time to'.
STEP 21. The condition of the closing of the day, if before the end of the day is no more busy intervals.
STEP 22. The condition of the coincidence of the free interval with a noted additional limitation interval.
STEP 23. Calculation of the free intervals. And saving in $memory variable.
STEP 24. Time of departure t^D calculation and saving in $first variable.
STEP 25. Time of arrival t^A calculation and saving in $second variable.
STEP 26. Free intervals $first.'-'.$second entering into an $arr array.
STEP 27. All free intervals from an $arr array output.
STEP 28. Departure time t^D output.
STEP 29. Downtime $dd (on passing points) calculation and output.
STEP 30. From the first point recommended departure time t^D output.
STEP 31. "Cancel" and "Save" buttons output.
STEP 32. Results saving in the database.
STEP 33. Return to the STEP 1.

The developed algorithm can be also represented as it is shown on the Figure 3.

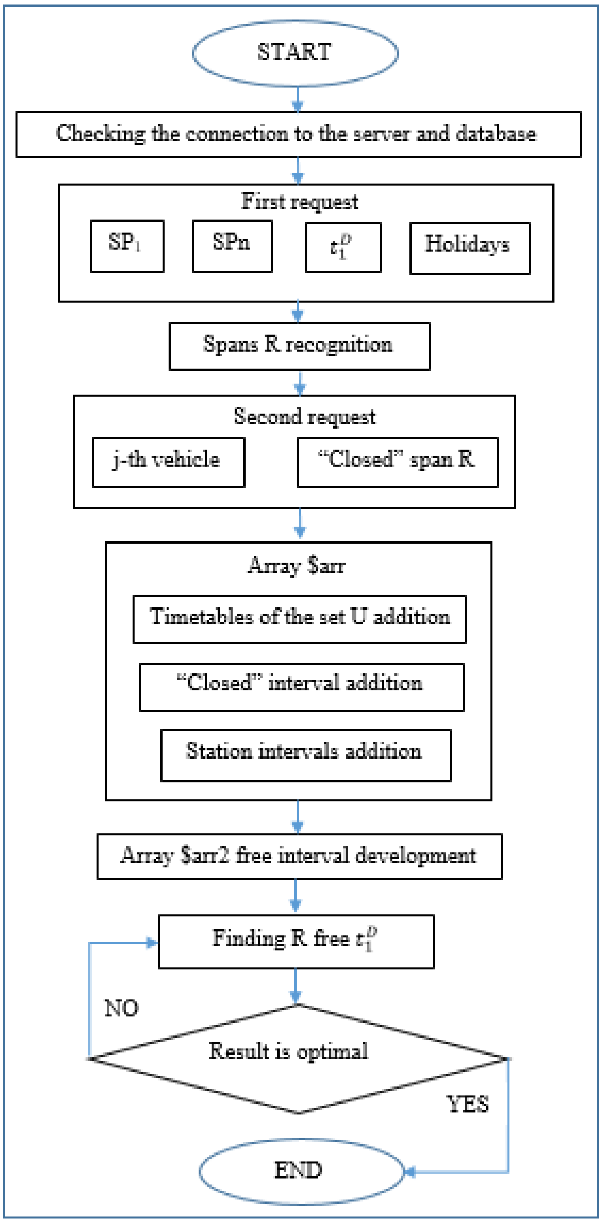
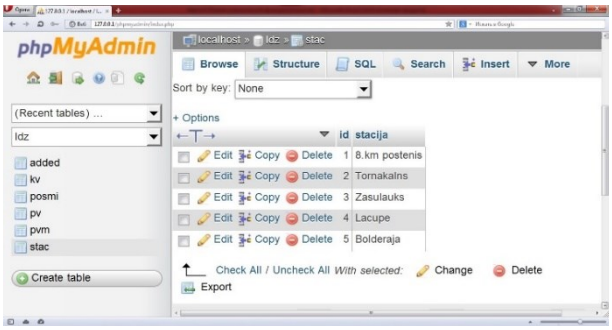

START
Checking the connection to the server and database
First request
SP₁
SPn
t_1^D
Holidays
Spans R recognition
Second request
j-th vehicle
"Closed" span R
Array $arr
Timetables of the set U addition
"Closed" interval addition
Station intervals addition
Array $arr2 free interval development
Finding R free t_1^D
NO
Result is optimal
YES
END

Figure 3. Scheduling algorithm

6. STRUCTURE OF THE DATABASE FOR THE MODEL

For the simulation model the database in MYSQL DBMS was developed for the railway transport. The usage of database improve the control of the data processing of the model. If conditions are changed, there is no need to rewrite the entire code of the model.

Proposed database contains 4 types of tables:
1) Table of the departure and arrival points (Figure 4). It can be stops, stations or other points on the route. Each post, park or station in this table is marked with an unique number ID.
Columns of the table: <Point ID, Station>



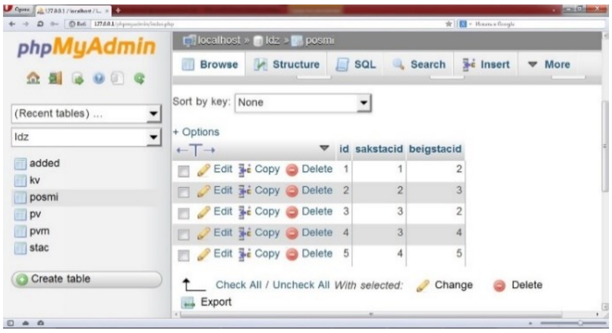
id	stacija
1	8.km postenis
2	Tornakalns
3	Zasulauks
4	Lacupe
5	Bolderaja

Figure 4. Table of departure and arrival stations

In this table ID numbers are given to all the necessary stations.

2) Table of the spans between departure and arrival points (Figure 5). Each span is marked with an unique number ID.

Columns of the table: <Span ID, Departure point ID, Arrival point ID>



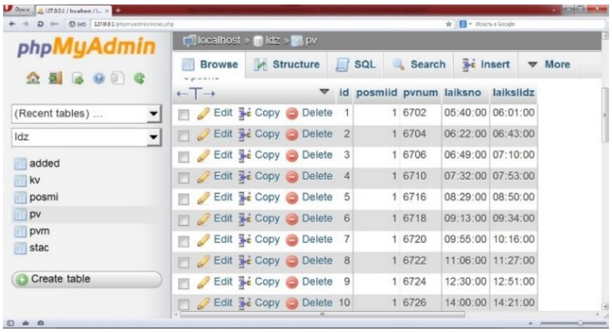
id	sakstacid	beigstacid
1	1	2
2	2	3
3	3	2
4	3	4
5	4	5

Figure 5. Table of spans

Before making the table about spans it is necessary to find out the direction of the movement.

3) Tables with transport vehicles timetable. One table for each span.

Columns of the table: <Vehicle ID, Span ID, Number of the vehicle, Departure time, Arrival time>.

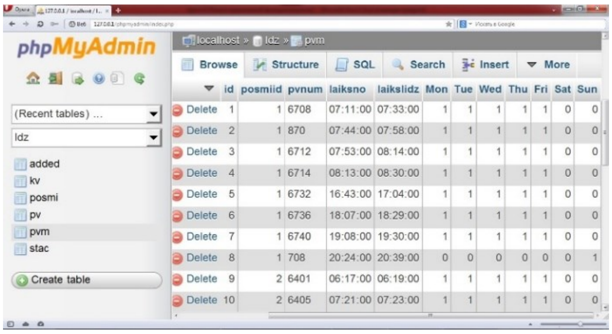


id	posmiid	pvnum	laiksno	laikslidz
1	1	6702	05:40:00	06:01:00
2	1	6704	06:22:00	06:43:00
3	1	6706	06:49:00	07:10:00
4	1	6710	07:32:00	07:53:00
5	1	6716	08:29:00	08:50:00
6	1	6718	09:13:00	09:34:00
7	1	6720	09:55:00	10:16:00
8	1	6722	11:06:00	11:27:00
9	1	6724	12:30:00	12:51:00
10	1	6726	14:00:00	14:21:00

Figure 6. Table of passenger trains timetable

For the experiment, two tables with timetables for the passenger trains were made. One table for the passenger trains running every day (Figure 6).

Another one table was made for passenger trains, which run depending on the weekday or weekend (Figure 7). Uniqueness of this table is availability of train finding by using clear logic. Depending on either under the weekday is “1” – train is running today, either under the weekday is “0” – train is not running today.



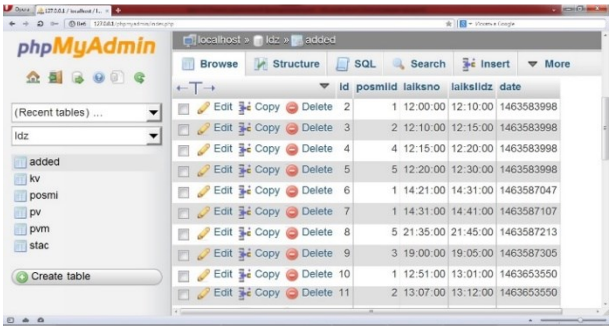
id	posmiid	pvnum	laiksno	laikslidz	Mon	Tue	Wed	Thu	Fri	Sat	Sun
1	1	6708	07:11:00	07:33:00	1	1	1	1	1	0	0
2	1	870	07:44:00	07:58:00	1	1	1	1	1	0	0
3	1	6712	07:53:00	08:14:00	1	1	1	1	1	0	0
4	1	6714	08:13:00	08:30:00	1	1	1	1	1	0	0
5	1	6732	16:43:00	17:04:00	1	1	1	1	1	0	0
6	1	6736	18:07:00	18:29:00	1	1	1	1	1	0	0
7	1	6740	19:08:00	19:30:00	1	1	1	1	1	0	0
8	1	708	20:24:00	20:39:00	0	0	0	0	0	0	1
9	2	6401	06:17:00	06:19:00	1	1	1	1	1	0	0
10	2	6405	07:21:00	07:23:00	1	1	1	1	1	0	0

Figure 7. Table of passenger trains, which runs depending on the weekday, timetable

Table with timetable of the freight trains was made as well. The principle is same as for the tables with timetables for the passenger trains.

4) Tables with saved results were made (Figure 8).

Columns of the table: <Schedule ID, Span ID, Departure time, Arrival time, Date>



id	posmiid	laiksno	laikslidz	date
2	1	12:00:00	12:10:00	1463583998
3	2	12:10:00	12:15:00	1463583998
4	4	12:15:00	12:20:00	1463583998
5	5	12:20:00	12:30:00	1463583998
6	1	14:21:00	14:31:00	1463587047
7	1	14:31:00	14:41:00	1463587107
8	5	21:35:00	21:45:00	1463587213
9	3	19:00:00	19:05:00	1463587305
10	1	12:51:00	13:01:00	1463653550
11	2	13:07:00	13:12:00	1463653550

Figure 8. Table of saved results

After the timetable for transport vehicle is chosen, it is saved in the database, and is considered that saved in the database time interval is busy, and no other vehicles can use the same transportation time on the same route within one day.

7. DEVELOPED SIMULATION MODEL AND EXPERIMENTS

The specially developed software is offered as the system of planning safe transportation process, which makes a selection from the database, makes calculations and records the obtained results in a database.

At first movement directions for the experiment were chosen (Figure 9). For the experiment, real part of the Latvian railway was taken.

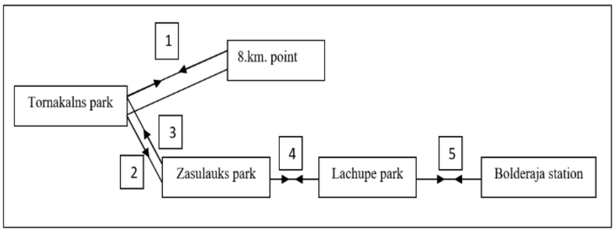
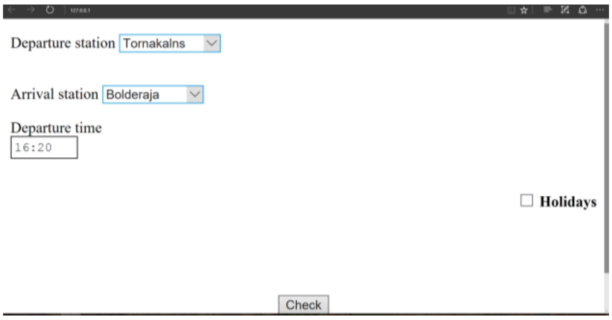


Figure 9. Movement directions

The first request was entered to the developed system (Figure 10).


Departure station Tornakalns

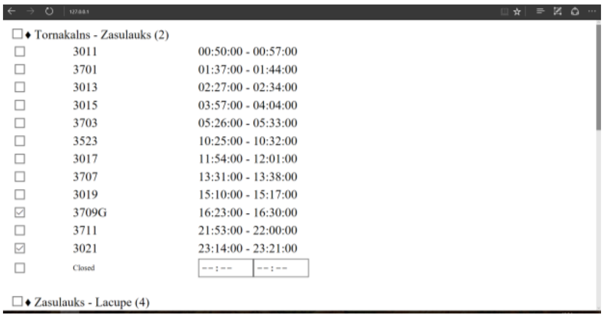
Arrival station Bolderaja

Departure time
16:20

□ Holidays

Figure 10. First request

Where,
Departure station – Tornakalns;
Arrival station – Bolderaja;
Departure time – 16:20;
Not a holiday.


□ ♦ Tornakalns - Zasulauks (2)

□	3011	00:50:00 - 00:57:00
□	3701	01:37:00 - 01:44:00
□	3013	02:27:00 - 02:34:00
□	3015	03:57:00 - 04:04:00
□	3703	05:26:00 - 05:33:00
□	3523	10:25:00 - 10:32:00
□	3017	11:54:00 - 12:01:00
□	3707	13:31:00 - 13:38:00
□	3019	15:10:00 - 15:17:00
☑	3709G	16:23:00 - 16:30:00
□	3711	21:53:00 - 22:00:00
☑	3021	23:14:00 - 23:21:00
□	Closed	--:-- --:--

□ ♦ Zasulauks - Lacupe (4)

Figure 11. Second request

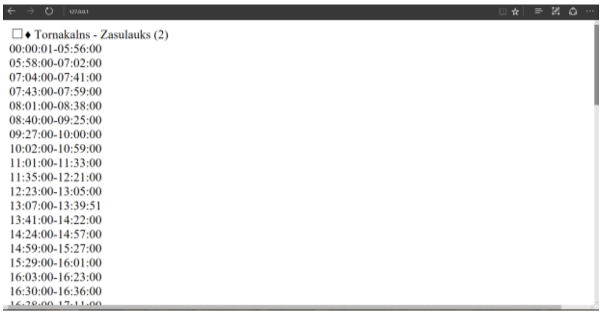
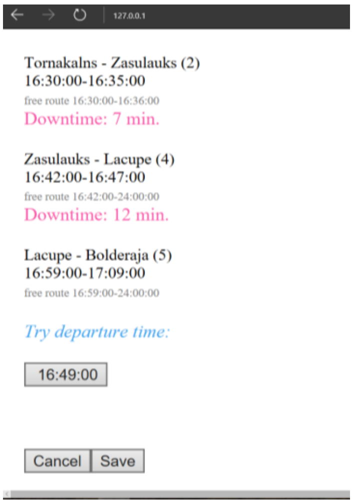

□ ♦ Tornakalns - Zasulauks (2)
00:00:01-05:56:00
05:58:00-07:02:00
07:04:00-07:41:00
07:43:00-07:59:00
08:01:00-08:38:00
08:40:00-09:25:00
09:27:00-10:00:00
10:02:00-10:59:00
11:01:00-11:33:00
11:35:00-12:21:00
12:23:00-13:05:00
13:07:00-13:39:51
13:41:00-14:22:00
14:24:00-14:57:00
14:59:00-15:27:00
15:29:00-16:01:00
16:03:00-16:23:00
16:30:00-16:36:00

Figure 12. All available intervals

After all the request was entered, the button "Check" was pushed and the second request was received (Figure 11). All the necessary freight trains were marked.

After both request were entered, "Check" button was pushed and the result was received. The result is divided into two parts. The first part - display of all available intervals on the running line during the day (Figure 12).

The second part - calculated time on the running line before the optimization (Figure 13) and after the optimization (Figure 14).


Tornakalns - Zasulauks (2)
16:30:00-16:35:00
free route 16:30:00-16:36:00
Downtime: 7 min.

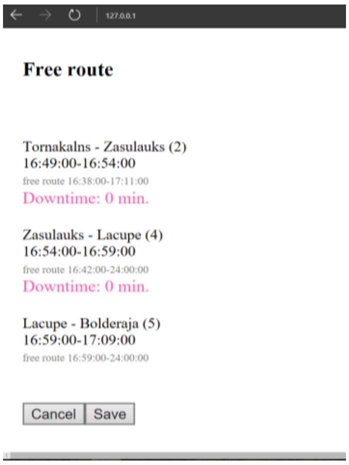
Zasulauks - Lacupe (4)
16:42:00-16:47:00
free route 16:42:00-24:00:00
Downtime: 12 min.

Lacupe - Bolderaja (5)
16:59:00-17:09:00
free route 16:59:00-24:00:00

Try departure time:

16:49:00

Figure 13. Results before optimization


Free route

Tornakalns - Zasulauks (2)
16:49:00-16:54:00
free route 16:38:00-17:11:00
Downtime: 0 min.

Zasulauks - Lacupe (4)
16:54:00-16:59:00
free route 16:42:00-24:00:00
Downtime: 0 min.

Lacupe - Bolderaja (5)
16:59:00-17:09:00
free route 16:59:00-24:00:00

Figure 14. Results after optimization

The comparison of two results, that were received in experiment, is displayed on the Table 1.

The preferred departure time was 16:20. Before optimization the system showed that the first span will be free on 16:30, and train can be sent to the next station – Zasulauks. But, second span will be busy till 16:42 and for the safety transportation process, train needs to make a stop on the Zasulauks park. The same situation is at the next station – Lacupe. Train arrives at Lacupe park at 16:47, makes a stop, and stands for the 12 minutes, because the third span is busy till 16:59. As the result, to gain a safety transportation process, train makes two stops on the intermediate stations with the common downtime 19 minutes.

After optimization - Tornakalns cargo terminal train is sent from Tornakalns park for 29 minutes later the preferred shipping time: at. 16:49. It is send through the intermediate stations without stops and downtime. And arrives on the Bolderaja station at the same time as before optimization: at. 17:09.

Table 1: Results Comparison

Operation	Park / station	Before optimization	After optimization
Departure	Tornakalns	At 16:30	At 16:49
Departure time shift from the desired time 16:20	Tornakalns	10 min.	29 min.
Arrival	Zasulauks	At 16:35	At 16:54
Downtime	Zasulauks	**7 min.**	**0 min.**
Departure	Zasulauks	At 16:42	At 16:54
Arrival	Lacupe	At 16:47	At 16:59
Downtime	Lacupe	**12 min.**	**0 min.**
Departure	Lacupe	At 16:59	At 16:59
Arrival	Bolderaja	At 17:09	At 17:09
Downtime sum		**19 min.**	**0 min.**

After the comparing the two results, was made the conclusion that in the first variant of the result, was offered the closest departure time for the safe transportation process. But this result might be suboptimal in time and energy economy. In the second variant of the result stops at intermediate stations and downtime was reduced to zero and safety task is performed as well.

8. CONCLUSIONS

Experiment shows, that system takes into account all the possible risks and advices to use safe and the most efficient time for public transport running process. Amount of stops at intermediate stations was reduced to zero, downtime was reduced to zero as well, routes were free from another vehicles during advised time, the result was received by using the information from the database, that means, that there is no need to have a special worker for planning transportation process, because it is automatized.

Proposed in this research goal was reached. In this research clear logic was used. All the conditions were clearly described. In real life human being while making a decision reasons about the possibility of the event. To make the system be artificial it is necessary to use fuzzy logic and authors started a new research based on it.

ACKNOWLEDGEMENT

The research leading to these results is performed in a frame of ARTEMIS Joint Undertaking project No. 621429.

REFERENCES

Alps I, Beinarovica A., Gorobetz M., Levchenkovs A. 2016. “Immune algorithm and intelligent devices for schedule overlap prevention in electric transport” In: Proceedings of 57th RTU International Scientific Conference. Riga, Latvia.

Alps I. 2012. “Modelling of Intelligent Electrical Transport Control Systems Scheduling Problems in the Unforseen Cases”, Doctorate work, pp.127

Beinarovica A., Gorobetz M., Levchenkovs A. 2017. “Algorithm of Energy Efficiency Improvement for Intelligent Devices in Railway Transport”. The ECCE Journal of Riga Technical University, 29-34 p., Published Online: 1.18.2017.

Conway R.W, Maxwell W.L., Miller L.W. 2003. “Theory of Scheduling” Dover, published by Addison-Wesley Publishing company, Mineola N.Y. 11501, USA

European Railway Agency. 2010. “Railway Safety Performance in the European Union 2010”, Espace International, 299 Boulevard de Leeds, Lille, France, 64 p.

Levchenkov A. and Gorobetz M. 2014. "An Evolutionary Algorithm for Reducing Railway Accidents Caused by Human Factors", in J. Pombo, (Editor), "Proceedings of the Second International Conference on Railway Technology: Research, Development and Maintenance", Civil-Comp Press, Stirlingshire, UK, Paper 230, doi:10.4203/ccp.104.230A.

Levchenkov A., Gorobetz M. and Mor-Yaroslavtsev A. 2012. “Evolutionary algorithms in embedded intelligent devices using satellite navigation for railway transport” chapter 16 in Infrastructure design, signalling and security in railway, Xavier Perpinya (Ed.), ISBN: 978-953-51-0448-3, InTech, pp. 395-420

Matsumoto A., Michitsuji Y., Tanifuji K. 2015. “Train-Overturned Derailments due to Excessive Speed - Analisys and Countermeasures”. In Proceedings of 24rd International Symposium on Vehicle Dynamics on Roads and Tracks, Graz, Austria, 2015 - 49.4, 1-6 pp.

Nikolajevs A., Mezītis M. 2016. “Level Crossing Time Prediction”. In: 57th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON) : Proceedings, pp. 199.-202. ISBN 978-1-5090-3729-2.

Shui X.; Zuo X.; Chen C. 2012. “A cultural clonal selection algorithm based fast vehicle scheduling approach”. IEEE Congress on Evolutionary Computation, 2012, Pages: 1 - 7, DOI: 10.1109/CEC.2012.6256624

Staņa, Ģ., Bražis V., Apse-Apsītis P. 2014. “Virtual Energy Simulation of Induction Traction Drive Test Bench”. In: 2014 IEEE 2nd Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE) : Proceedings of the 2nd Workshop IEEE, pp. 75.-80.. ISBN 978-1-4799-7123-7. e-ISBN 978-1-4799-7122-0. doi:10.1109/AIEEE.2014.7020330

AUTHOR BIOGRAPHIES

ANNA BEINAROVICA is a Ph.D. student of the Institute of Industrial Electronics and Electrical

Engineering of Riga Technical University. She received the degrees of B.Sc. and M.Sc. in railway transport from Riga Technical University in 2014 and 2016, respectively. In 2010, she participated in the scientific conference for young researches “Research Innovations Fundamentals 2010” in Lithuania, Klaipeda. In 2016, she participated in the scientific conference RTUCON 2016 in Riga.
Address: Āzenes iela 12–1, Riga, LV-1048, Latvia.

MIKHAIL GOROBETZ is an Assistant Professor and the leading researcher of the Institute of Industrial Electronics and Electrical Engineering at Riga Technical University. The results of his research have been published in various international scientific proceedings in the fields of adaptive control, neural networks, genetic algorithms, modelling and simulation of dynamic processes. He is a leader of various national projects and international projects. M. Gorobetz is an author of many study books and patented inventions.
Address: Āzenes iela 12–1, Riga, LV-1048, Latvia.

ANATOLIJ LEVCENKOV Dr. sc. ing., Professor at the Institute of Industrial Electronics and Electrical Engineering and the Institute of Railway Transport of Riga Technical University. He received the diploma of an engineer in electrical engineering in 1969, and the Dr. sc. ing. degree in 1978. His fields of interests are optimization theory, group decision support systems, negotiation support systems, scheduling, logistics, intelligent transport systems, evolutionary algorithms for embedded systems. A. Levcenkov has been leading various national and international projects, and he is an author of many patents, books, and publications.
Address: Āzenes iela 12–1, Riga, LV-1048, Latvia.

A DESIGN PATTERN FOR MODELLING AND SIMULATION IN HOSPITAL PHARMACY MANAGEMENT

Wirachchaya Chanpuypetch and Duangpun Kritchanchai
Department of Industrial Engineering, Faculty of Engineering
Mahidol University
Phutthamonthon 4Rd., Nakhonpathom, 73170, Thailand
E-mail: wirachchaya@gmail.com, duangpun.skr@mahidol.ac.th

KEYWORDS:

Design pattern, conceptual modelling, process redesign, modelling, simulation, hospital pharmacy management.

ABSTRACT

Nowadays, several Thai hospitals are still suffering from inefficient processes. Redesigning business processes is an effective way for improvement. However, business process mapping and analysis in healthcare environment has become a complex task. Besides, most healthcare professionals often resist change. To deal with these obstacles, a design pattern useful for modelling and simulation is suggested in this paper. For deriving patterns in this study, current problems focusing on hospital pharmacy management are initially identified through empirical investigation. Then the suitable solutions can be determined based on the literature review. The design pattern is represented through various attributes including pattern name, purpose, description, modelling structure using a modelling language. In addition, more attributes related to simulation modelling are also combined such as control variables and performance measures. Process modellers can rapidly understand the context of problems through design patterns. The proposed patterns can be reused to create a new model and conduct a successful simulation study. This approach can be a valuable tool to redesign healthcare systems.

INTRODUCTION

In Thailand, hospital pharmacy management is still suffering from inefficient processes. Many Thai hospitals urgently need to improve their business processes. However, redesign projects in healthcare systems have often been unsuccessful (Patwardhan and Patwardhan 2008). Business process analysis and improvement in this environment is a complex task and challenge (MacPhee 2007). They need a high degree of collaboration and coordination among individuals and functions. Besides, resistance to change often exists (Patwardhan and Patwardhan, 2008; Khodambashi 2013). To support these issues, simulation modelling is an effective approach to achieve success in healthcare redesign (Williams and Vanessa (2003). Changes in healthcare systems can be tested prior to real-life implementation (AbuKhousa et al. 2014). It can also lead all healthcare stakeholders to participate in a redesign project. Nevertheless, in simulation modelling, a process modeller spends the most time for understanding the context of the problems and structuring the conceptual model (Robinson 2015). It is one of the most challenging tasks in a simulation study. To accelerate these tasks, a design pattern has been suggested as an effective way. This approach represents business process flows and offers some best practices for applying them to the specific context (Barchetti et al. 2011). Typically, a design pattern depicts the description of solution through various attributes such as pattern name, description, purpose, as well as process workflow using modelling notation (Barchetti et al. 2011). Likewise, a non-software specific description of the computer simulation may be included into the pattern such as inputs (experimental factors), outputs (responses), content (scope and level of detail), assumptions, and simplifications of the model (Robinson 2015; Tolk et al. 2013). These elements can facilitate a specification of a simulation project.

A design pattern allows users to adopt in practical context. Reusing process patterns can reduce error-prone task and time in modelling. It has the potential to offer a well-defined practice to healthcare professionals in healthcare management. Besides, this way enables all stakeholders to communicate more effectively (Gschwind et al. 2008), which is one of requirements to achieving successful business process improvement in healthcare system. Therefore, in this paper, a design pattern for modelling and simulation is proposed focusing on crucial problems in hospital pharmacy management.

The paper is organised as follows. Next, current problems of hospital pharmacy management are investigated empirically. The crucial problems are described along with their solutions. Then the design pattern that integrates elements of modelling and simulation is proposed as an exemplification. Finally, the paper is concluded.

AN EMPIRICAL INVESTIGATION OF CURRENT PROBLEMS IN HOSPITAL PHARMACY MANAGEMENT

In this section, the methodology used to develop the design patterns is proposed. An empirical investigation was conducted to understand the context of hospital pharmacy management through case studies at the beginning stage. Then crucial problems and their suitable solutions can be identified. Eventually, the design patterns

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

useful for modelling and simulation in hospital pharmacy management can be established. The methodology is structured in four main stages, as illustrated in Figure 1.

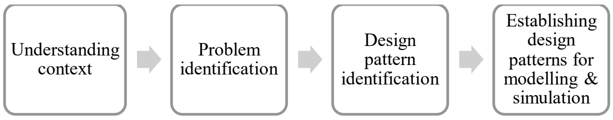

Understanding context
Problem identification
Design pattern identification
Establishing design patterns for modelling & simulation

Figure 1: Stages for Developing a Design Pattern

According to the proposed methodology, fifteen hospitals in Thailand where convenience to access were visited to investigate the problem situation. The selected cases of hospital are listed in Table 1.

Table 1: The Selected Cases of Hospital in Thailand

Nature of owner	Type of hospital	Level of expertise	# Stocked drug SKU	Number of cases
Public Hospital	Community	Primary care	360 – 450	3
	General	Secondary care	1,122	1
		Tertiary care	N.A.	1
	Regional	Tertiary care	1,000 – 2,000	2
	Teaching hospital	Tertiary care	800 - > 2,000	4
Public Organisation Hospital	-	Secondary care	1,700	1
Private Hospital	-	Secondary care	1,732 – 2,000	3

In data collection process, we used the core open-ended questions as the guideline for semi-structured interview. The interview pursued to understand hospital business processes and recognise current problems in term of hospital pharmacy management. The managers at senior level who could provide relevant information were interviewed. They included the head of a hospital pharmacy department, a pharmacy purchasing manager, a pharmacy warehouse manager, and a pharmacy store manager. All interviewees were asked the same questions, and interview length ranged from 1 to 3 hours. Simultaneously, direct observation was also performed.

After crucial problems are addressed, we then reviewed and synthesised the applicable practices based on the related literature. The context of problems in hospital pharmacy management is described along with their solutions can be offered in the following section.

PROBLEMS IN HOSPITAL PHARMACY MANAGEMENT AND THEIR SOLUTIONS

According to empirical evidence, five crucial problems of pharmaceutical products management that Thai hospitals currently confront can be identified. The contexts of each problem are described along with literature-based solutions as follows.

Inefficient Management of Inventory and Warehousing

In Thailand, an inefficient warehousing and inventory management is often found in small size primary care hospitals. These hospitals manage and control their inventory without a computerised material management system. Stock cards are mostly used to provide a simple stock control system in the hospital. Important data of drugs are also neglected in recording such as drug expiration dates and a manufacturing lot information. Besides, drug items which are stored in the hospital warehouse may be placed on the shelf inappropriately. Expired drugs are frequently found.

This problem can be resolved through the warehouse planning and control structure with the use of a computerised information system (IS) in the entire materials management processes (Ferretti et al. 2014; Holm et al. 2015). Several advantages can be obtained such as better space organisation, stock reduction, no more expired drugs with saving cost, safety improvement with reduction in administration errors, no more transcription errors, and relationship improvement with the providers (Ferretti et al. 2014). This solution can provide real-time information of inventory status. Drug inventory management has been improved with reducing of waste and inventory shortage. Moreover, availability of medication at the point of care can be increased through this approach (Holm et al. 2015).

Unable to Track and Trace Drug Items

Traceability is becoming an essential in many industries to improve supply chain efficiency. It is increasing interests related to the problem of theft, counterfeited drugs, and recall. However, drugs traceability cannot be currently achieved in Thai pharmaceutical supply chain. After drug receiving and placing, various hospitals neglect to record important drug data such as a batch or lot number. Moreover, most IS in Thai hospitals cannot support an entering of these data. Thus, the past and present locations of the movement of prescription drugs cannot be determined exact data of where the drug is located and where, or to whom, each drug has been sold. Likewise, it is impracticable to recall a drug from the market. These recalled drugs may cause patient harm due to defect or failure of quality.

Currently, to protect consumers from contaminated medicine and counterfeit drugs, the United States Food and Drug Administration (FDA) defines a drug pedigree for pharmaceutical track and trace. An electronic form of this document includes the basic data elements from pharmaceutical supply chain companies such as lot, potency, expiration, National Drug Code and Electronic Product Code (EPC). Besides, track and trace technology, such as 2D barcode and RFID, is also required for interchanging through the whole supply chain (Hamid and Ramish 2014).

Inventory Inaccuracy

According to the empirical data, the delivered items from distributors are manually checked and recorded for receiving to store in the hospital warehouse. Likewise, the replenished items are mostly operated with the similar activities for storing the received drugs in each pharmacy repository. However, these placed items may not be checked and recorded in suddenly. This fragmented operation flow leads to various error types of inventory inaccuracy. It is a common problem of supply chains (Kök and Shang 2014; Sarac et al. 2015).

Typically, inventory inaccuracy can be divided into three main types including shrinkage error, transaction error, and misplacement (Dai and Tseng 2012). They can be caused by thefts, shipment errors, delivery errors, scanning errors and misplacements (Sarac et al. 2015). When this problem occurs, the actual inventory levels are higher or lower than the nominal inventory. Likewise, ineffective decision making for replenishment cannot be avoided. This can lead to high out-of-stocks, backlog, and/or excess inventory (Kang and Gershwin 2005). Inventory costs may increase resulting from inaccurate information significantly (Kök and Shang 2014).

However, inventory inaccuracy cannot be observed by IS until an inspection is performed (Kök and Shang 2014). Thus, the actual inventory levels are hardly maintained (Kang and Gershwin 2005). To remedy this problem, the optimised cycle-count policy should be proposed to correct inventory record errors (Kök and Shang 2014). Although, a cycle-count program is used, this approach cannot prevent misplacement errors and shrinkage errors. Particularly, shrinkage type errors lead to the biggest impact on supply chain performance (Fleisch and Tellkamp 2005). To facilitate this alignment, automatic identification (Auto-ID) technology, such as RFID, is frequently suggested to provide inventory visibility (Zhang et al. 2011). The proposed system with RFID implementation can enhance supply chain effectiveness by minimising the inventory inaccuracy problem.

Inefficient Inventory Management Policy

According to empirical evidence, most hospitals in Thailand allow each hospital pharmacy to control and manage their inventory independently. Decentralised decisions may create a strong bias for making a requisition with high inventory levels. The demand distortion can affect the supply chain partners with inaccurate forecasts, increased inventory levels, and increased overall cost of inventory management (Kamalapurkar 2011).

To deal with this problem, information sharing with considering inventory control policies and/or collaboration strategies between supply chain partners has been studied by many researches. For this approach, actual end-customer demand data access should be available to all stakeholders in order to forecast the demand, instead of on the requisition data from each stage. Mostly, the well-known collaboration strategies have also been suggested to integrate supply chain and improve inventory control, namely; Vendor Managed Inventory (VMI) and Collaborative Planning, Forecasting and Replenishment (CPFR). Groznik and Maslaric (2009) reengineered business processes of oil/retail petrol supply chain by providing the improved integration of whole parts of the supply chain and centralised distribution process management. The renewed business models with information sharing by considering VMI strategy can reduce inventory holding cost and the bullwhip effect. In healthcare sector, Kim (2005) implemented the VMI system for improving pharmaceutical products management in hospital. The study developed the online procurement system, which provides real time information sharing functionalities to achieve information integration. After the system deployment, it enables hospitals to eliminate errors, decrease administration tasks, and increase reliability of information flow. Average inventory amounts and total inventory costs of drugs has been decreased significantly.

Another such supply chain collaboration practice is known as CPFR. It is the latest strategy in the evolution of supply chain collaboration that extends the idea of VMI to include joint planning process (Kamalapur and Lyth 2014; Alftan et al. 2015). Most studies show that the benefits gained from CPFR are always higher than VMI (Kamalapurkar 2011). In healthcare industry, Li (2010) evaluated the inventory collaborations through a system dynamics study. As a result, the total average inventory reduces almost 20%, the amount of backlogs decreases over 60% by VMI and fully eliminates by CPFR. Recently, the benefits of CPFR in the healthcare sector were examined by Lin and Ho (2014). They concluded that this approach could enhance medical service quality and eliminate inefficient purchase and waste of valuable medical resources.

As described above, information sharing among supply chain members can significantly deal with inefficiency inventory management practice. Inventory control policies and collaboration strategies such as VMI, CPFR are frequently combined in the proposed information sharing models. However, these approaches must rely on communication mechanisms such as Electronic Data Interchange (EDI), XML, etc. (Liu and Kumar 2003). This approach helps to reduce the bullwhip effect and improve service level.

Prescribing Problems

Prescription medication is the most important aspect of patient treatment in a hospital. However, in drug prescribing, physicians have made a prescription using a prescription form until now. This handwritten prescription leads to various types of medication error such as prescribing error, transcribing error, pre-dispensing and dispensing error. These errors may be harmful to patients who encounter wrong medication. Besides, physicians may select the drug item that is not stored in the hospital for a prescription. This event brings about to create rework activities that are non-value added.

To cope with this problem, a CPOE system is offered to facilitate medication errors related to manual drug order writing with paper (Al-Rowibah et al. 2013). A deployment of a CPOE can decrease adverse drug events (ADEs) and medication errors related to handwritten prescriptions. Physician orders can be transferred to hospital pharmacy through a secure way. However, physicians often resist an electronic prescribing. Hence, the strategy to promote an adoption of CPOE is needful to overcome this top barrier (Charles et al. 2014).

As described above, five literature-based approaches for dealing with the crucial problems of hospital pharmacy management can be summarised in Table 2. Subsequently, these approaches are considered to create the design patterns for hospital pharmacy management.

Table 2: Summary of Problems in Hospital Pharmacy Management and Their Solutions

Problems	Solutions	Benefits
Inefficient management of inventory and warehousing	Warehousing and inventory management with implementing of a computerised material management system	• Reduce stock and waste with saving cost • Reduce inventory shortage • Improve safety with reduction in administration errors and medication errors
Unable to track and trace drug items	Traceability architecture and a deployment track and trace technology–2D barcode, RFID	• Able to track and trace each movement of the drug • Increase correctness and timeliness
Inventory inaccuracy	Counting items in the pharmacy inventory and using auto-ID technology such as RFID	• Eliminate inventory inaccuracy errors such as misplaced type error and theft type error • Provide better replenishment decisions
Inefficient inventory management policy	Information sharing and providing a collaboration strategy such as VMI, CPFR	• Improve system efficiency and service level • Reduce inventory amounts and total inventory costs
Prescribing problems	A deployment of a CPOE system including the strategy to promote an adoption of an electronic prescribing	• Provide a secure way of transferring a prescription to hospital pharmacy • Prevent an occurrence of non-value added activities

DESIGN PATTERNS FOR MODELLING AND SIMULATION

As the suitable solutions presented in previous section, the details of the design patterns can be provided. Typically, the proposed pattern represents the description of solution through various attributes including pattern name, pattern description, purpose, as well as structure of the pattern using a modelling language technique (Barchetti et al. 2011). However, these attributes have not been exactly determined. To support a simulation study, which is a useful approach for a redesign project in healthcare context, necessary elements related to simulation modelling are also encompassed such as independent variables (control variables) and response variables (performance measures). Thus, the proposed design pattern consists of the following six attributes:

- *Pattern name* – A pattern identifier.
- *Purpose* – Description of the purpose or motivations needed to identify the patterns.
- *Pattern description* – Description of the pattern.
- *Modelling structure* – Representation of the pre-defined activities assigned to specific stakeholders using modelling language technique.
- *Independent variables (Control variables)* – Factors that affect the system.
- *Dependent variables (Response variables)* – Performance measures.

Here, the authors exemplify *the design pattern of information sharing and providing a collaborative strategy for hospital pharmacy inventory management.* The pattern supports a hospital that confronts with inefficient inventory management practice and policy. This problem has been mentioned for bias forecasting as the crucial problem from all hospital cases. Effective literature-based solutions related information sharing and providing a collaborative strategy have been suggested such as VMI and CPFR. Based on the process reference model for hospital pharmacy management developed by Chanpuypetch and Kritchanchai (2015), sequences of events under each strategies are represented using Business Process Modelling Notation (BPMN). The proposed pattern leads to improving system efficiency and service level. Likewise, inventory amounts and total inventory costs can be reduced. Important parameters that impact on supply chain performance are also included into the design pattern for investigating their impact for both the hospital warehouse and the hospital pharmacy in different collaboration strategies. An example of the design pattern is shown in Table 3.

CONCLUSION

Currently, redesigning business process is much needed to improve inefficient processes for hospital pharmacy management in several Thai hospitals. However, business process mapping and analysis in healthcare has become a complex task. To ease of deriving a process model and achieve success a redesign project in healthcare, a design pattern is suggested in this paper.

The design patterns proposed in this study are created based on the crucial problems focusing on hospital pharmacy management. They are inefficient management of inventory and warehouse, inefficiency management policy, unable to track and trace drug items, inventory inaccuracy, and prescribing problems. According to these problems, the suitable practices, technologies, or strategies can be then determined to overcome these problems. Each pattern is represented through various attributes useful for modelling and simulation including pattern name, purpose, description, modelling structure, control variables, and performance measures. Process modellers and business users can rapidly understand the context of problems through the patterns. A design pattern also allows users to reuse workflow patterns and conduct a successful simulation study. It can be a valuable modelling tool to redesign healthcare systems.

REFERENCES

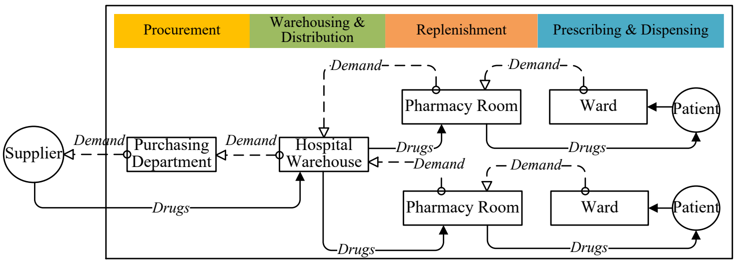
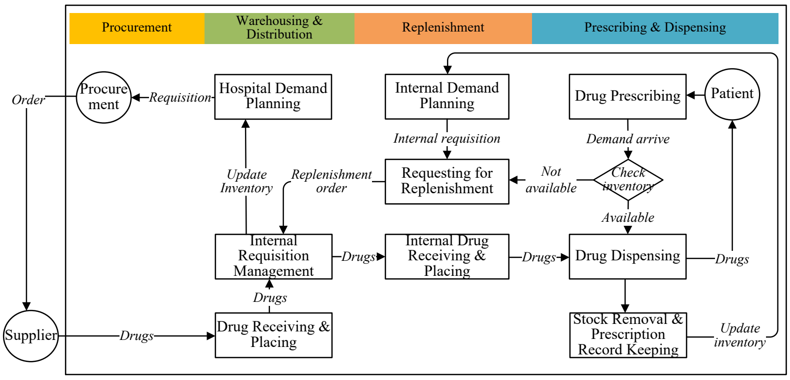
AbuKhousa, E., Al-Jaroodi, J., Lazarova-Molnar, S., and Mohamed, N. 2014. "Simulation and Modeling Efforts to Support Decision Making in Healthcare Supply Chain Management". *The Scientific World Journal,* 2014, 16.

Alftan, A., Kaipia, R., Loikkanen, L., and Spens, K. 2015. "Centralised grocery supply chain planning: Improved exception management". *International Journal of Physical Distribution & Logistics Management,* Vol.45(3), 237-259.

Al-Rowibah, F.A., Younis, M.Z., and Parkash, J. 2013. "The impact of computerized physician order entry on medication errors and adverse drug events". *Journal of Health Care Finance*, Vol.40(1), 93-102.

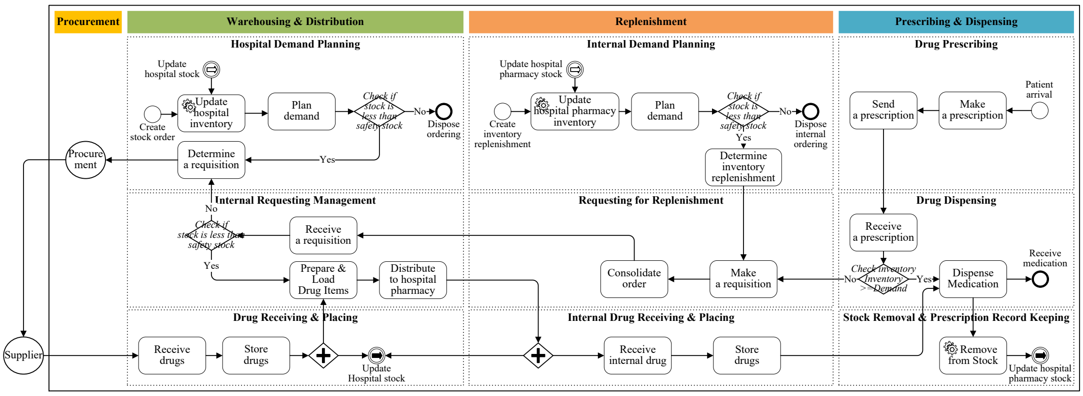
Barchetti, U., Bucciero, A., De Blasi, M., Guido, A.L., Mainetti, L., and Patrono, L. 2010. "Impact of RFID, EPC and B2B on traceability management of the pharmaceutical supply chain". In *Proceeding of the 5th International Conference on Computer Sciences and Convergence Information Technology*, (Seoul, Korea, Nov.30-Dec.2), 58-63.

Table 3: Design Pattern – Information Sharing and Providing a Collaborative Strategy for Hospital Pharmacy Inventory Management: An Exemplification

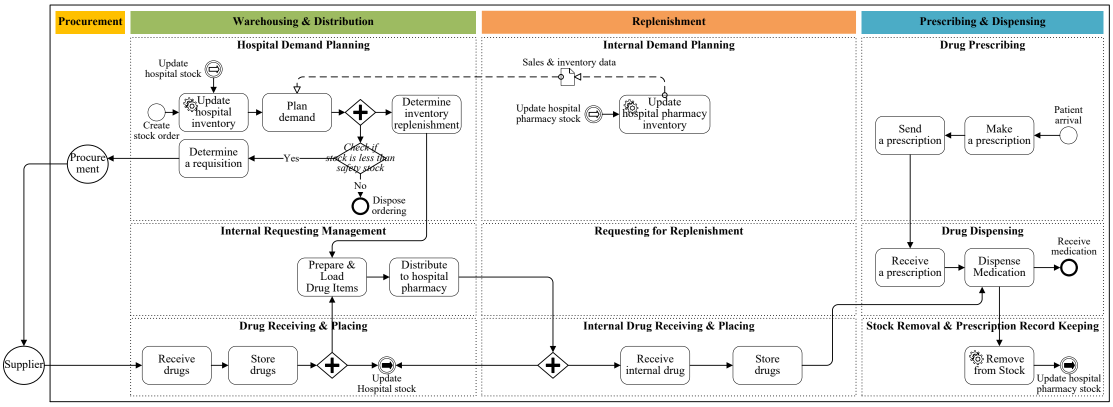
Attributes	**Description / Detail**
Name:	Information sharing and providing a collaborative strategy for hospital pharmacy inventory management
Purpose:	The design pattern represents modelling structures of hospital pharmacy inventory management process through information sharing and providing a collaborative strategy.
Description:	The pattern represents all activities related hospital pharmacy inventory management. It involves four main functions of hospital pharmacy management. They are procurement, warehousing and distribution, replenishment, and prescribing and dispensing. These functions are processed at four locations including hospital wards, hospital pharmacy, hospital warehouse, and purchasing department. A three-echelon inventory system is illustrated as the 1st hierarchical level of modelling. The current tasks and its activities are operated as follows. Tasks and activities of "Prescribing and dispensing" function: 1) *Drug prescribing* – Physicians make and send a prescription to hospital pharmacy. 2) *Drug dispensing* – Hospital pharmacy receives a prescription and dispenses medication. 3) *Stock removal and prescription record keeping* – After dispensing, the drug items are removed from stock for updating inventory level. Tasks and activities of "Replenishment" function: 1) *Internal demand planning* – If stock is less than safety stock, the internal order quantities are determined for replenishment. 2) *Requesting for replenishment* – Hospital pharmacy makes a requisition for replenishment. Internal order requisitions are then consolidated and sent to the hospital warehouse. 3) *Internal drug receiving and placing* – After receiving shipments from the hospital warehouse, hospital pharmacy inventory is available for dispensing. Tasks and activities of "Warehousing and distribution" function: 1) *Hospital demand planning* – The hospital warehouse plans their demand forecast independently and calculate their order-up-to inventory level. If stock of the hospital warehouse is less than safety stock, the order quantities are determined for purchasing. 2) *Internal requisition management* – The requirement that are requested from hospital pharmacies are prepared for replenishment. An internal delivery is operated for distributing the requested drug items to replenish pharmacy repositories. 3) *Drug receiving and placing* – The hospital warehouse receives shipments from suppliers. Hospital inventory is available for fulfilment hospital pharmacy repositories order. (see the 2nd hierarchical level of modelling and the 3rd hierarchical level of modelling: sequence of events in AS-IS) To cope with inefficient inventory management policy, two different collaboration strategies are often offered as follows. • *Vendor Managed Inventory (VMI)* – The hospital warehouse manages the inventory of hospital pharmacies. Hospital pharmacies share sales and inventory data with the hospital warehouse. The warehouse forecasts and determine inventory level of hospital pharmacies during each period (see 3rd hierarchical level of modelling – 3.2 Sequence of events under VMI). • *Collaborative Planning Forecasting and Replenishment (CPFR)* – Hospital pharmacies share historical usage, forecast and inventory level information with the hospital warehouse. The hospital warehouse does not forecast and uses this information to determine their inventory level during each period (see 3rd hierarchical level of modelling – 3.3 Sequence of events under CPRF). The pattern need inventory information sharing and/or exchanging between the hospital warehouse and hospital pharmacies via well-organised mechanisms to support their collaboration planning process. Finally, at the end of period, the hospital pharmacy cost and the hospital warehouse cost per period are calculated based on inventory level or backorder quantity.
Modelling Structure:	*1st hierarchical level of modelling* – A three-echelon inventory system  Procurement Warehousing & Distribution Replenishment Prescribing & Dispensing Demand Demand Pharmacy Room Ward Patient Supplier Demand Purchasing Department Demand Hospital Warehouse Drugs Drugs Demand Demand Drugs Pharmacy Room Ward Patient Drugs Drugs
	2nd hierarchical level of modelling  Procurement Warehousing & Distribution Replenishment Prescribing & Dispensing Order Procurement Requisition Hospital Demand Planning Internal Demand Planning Drug Prescribing Patient Internal requisition Demand arrive Update Inventory Replenishment order Requesting for Replenishment Not available Check inventory Available Internal Requisition Management Drugs Internal Drug Receiving & Placing Drugs Drug Dispensing Drugs Drugs Supplier Drugs Drug Receiving & Placing Stock Removal & Prescription Record Keeping Update inventory

3rd hierarchical level of modelling

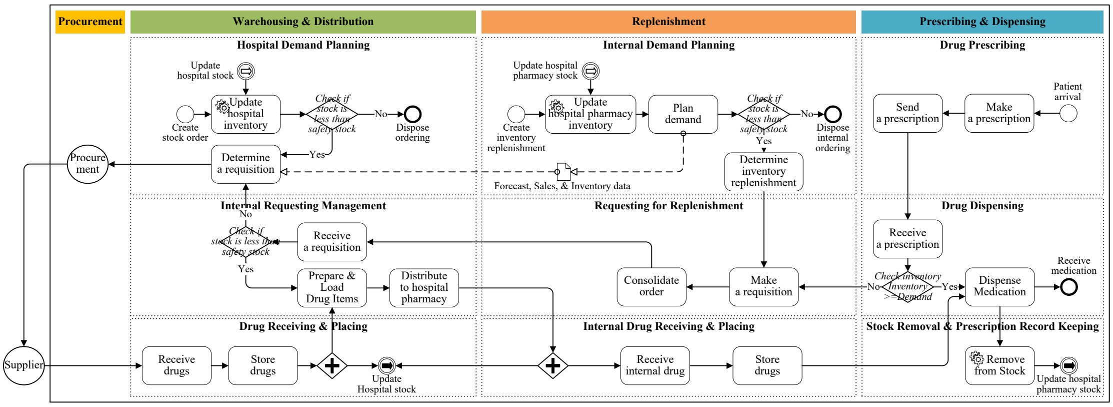
3.1) Sequence of events in AS-IS – No information is shared.


Procurement
Warehousing & Distribution
Replenishment
Prescribing & Dispensing
Hospital Demand Planning
Update hospital stock
Create stock order
Update hospital inventory
Plan demand
Check if stock is less than safety stock
No
Dispose ordering
Yes
Determine a requisition
Procurement
Internal Demand Planning
Update hospital pharmacy stock
Create inventory replenishment
Update hospital pharmacy inventory
Plan demand
Check if stock is less than safety stock
Dispose internal ordering
Determine inventory replenishment
Drug Prescribing
Send a prescription
Make a prescription
Patient arrival
Internal Requesting Management
Check if stock is less than safety stock
Receive a requisition
Prepare & Load Drug Items
Distribute to hospital pharmacy
Requesting for Replenishment
Consolidate order
Make a requisition
Drug Dispensing
Receive a prescription
Check inventory >= Demand
Dispense Medication
Receive medication
Drug Receiving & Placing
Supplier
Receive drugs
Store drugs
Update Hospital stock
Internal Drug Receiving & Placing
Receive internal drug
Store drugs
Stock Removal & Prescription Record Keeping
Remove from Stock
Update hospital pharmacy stock

3.2) Sequence of events under VMI – Sales and inventory data are shared.



3.3) Sequence of events under CPFR – Forecast, sales, and inventory are shared


Procurement
Warehousing & Distribution
Replenishment
Prescribing & Dispensing
Hospital Demand Planning
Update hospital stock
Create stock order
Update hospital inventory
Check if stock is less than safety stock
No
Dispose ordering
Yes
Determine a requisition
Procurement
Internal Demand Planning
Update hospital pharmacy stock
Create inventory replenishment
Update hospital pharmacy inventory
Plan demand
Check if stock is less than safety stock
Dispose internal ordering
Determine inventory replenishment
Forecast, Sales, & Inventory data
Drug Prescribing
Patient arrival
Send a prescription
Make a prescription
Internal Requesting Management
Check if stock is less than safety stock
Receive a requisition
Prepare & Load Drug Items
Distribute to hospital pharmacy
Requesting for Replenishment
Consolidate order
Make a requisition
Drug Dispensing
Receive a prescription
Check inventory >= Demand
Dispense Medication
Receive medication
Drug Receiving & Placing
Supplier
Receive drugs
Store drugs
Update Hospital stock
Internal Drug Receiving & Placing
Receive internal drug
Store drugs
Stock Removal & Prescription Record Keeping
Remove from Stock
Update hospital pharmacy stock

Control variables:

To investigate the impact of collaboration strategies to the hospital warehouse and hospital pharmacy in variable demand environment. Most variables are considered in the experimental design for simulation such as;

Control parameters	*1*	*2*	*3*
Supply chain strategy	TRP	VMI	CPFR
Demand variability level	Low	Medium	High
Random forecast error	Low	Medium	High
Bias forecast error	Negative	Neutral	Positive

Basic assumptions: Demand type; Demand forecast technique; Inventory replenishment policy; Review period; Information is exchanged real time without any delay in information sharing; Initial values of inventory; Capacity constrain; Delivery lead time.

Response variables:

Financial measurement	*Inventory control*	*Customer service*
⇩ Hospital pharmacy cost (inventory holding cost + Backorder cost) ($) ⇩ Hospital warehouse cost (inventory holding cost + Backorder cost) ($) ⇩ Total hospital cost (Hospital pharmacy cost + Hospital warehouse cost) ($)	⇩Inventory holding (units) ⇩Backorder quantity (units)	⇧ Fill rate (%)

Barchetti, U., Capodieci, A., Guido, A.L., and Mainetti, L. 2011. "Modelling collaboration processes through design patterns". *Computing and Informatics*, Vol.30(1), 113-135.

Chanpuypetch, W. and Kritchanchai, D. 2015. "An empirical-based construction of the multi-purpose process reference model for hospital supply chain". In M. Gen, K. J. Kim, X. Huang & Y. Hiroshi (Eds.), *Industrial engineering, management science and applications 2015, Lecture Notes in Electrical Engineering*, Vol.349, 901-912

Charles, K., Cannon, M., Hall, R., and Coustasse, A. 2014. "Can utilizing a computerized provider order entry (CPOE) system prevent hospital medical errors and adverse drug events?". *Perspectives in Health Information Management,* Vol.11(Fall), 1b.

Dai, H. and Tseng, M.M. 2012. "The impacts of RFID implementation on reducing inventory inaccuracy in a multi-stage supply chain". *International Journal of Production Economics,* Vol.139(2), 634-641.

Ferretti, M., Favalli, F., and Zangrandi, A. 2014. "Impact of a logistic improvement in an hospital pharmacy: Effects on the economics of a healthcare organization". *International Journal of Engineering, Science and Technology,* Vol.6(3), 85-95.

Fleisch, E. and Tellkamp, C. 2005. "Inventory inaccuracy and supply chain performance: A simulation study of a retail supply chain". *International Journal of Production Economics,* Vol.95(3), 373-385.

Groznik, A. and Maslaric, M. 2009. "Investigating the impact of information sharing in a two-level supply chain using business process modeling and simulations: A case study", In *Proceeding of the 23rd European Conference on Modelling and Simulation*, (Madrid, Spain, Jun.9-12), 39-45.

Gschwind, T., Koehler, J. and Wong, J. 2008, "Applying patterns during business process modeling", in Dumas, M., Reichert, M. & Shan, M.-C. (eds.), *Proceeding of the 6th Business Process Management International Conference, (Milan, Italy, Sep.2-4).* 4-19.

Hamid, Z. and Ramish, A. 2014. "Counterfeit drugs prevention in pharmaceutical industry with RFID: A framework based on literature review". *International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering,* Vol.8(4), 203-211.

Holm, M.R., Rudis, M.I., and Wilson, J.W. 2015. "Medication supply chain management through implementation of a hospital pharmacy computerized inventory program in Haiti". *Global Health Action,* Vol.8, 26546.

Kamalapur, R. and Lyth, D. 2014. "Benefits of CPFR collaboration strategy under different inventory holding and backorder penalty costs". *International Journal of Business and Management,* Vol.9(10).

Kamalapurkar, D. 2011. "*Benefits of CPFR and VMI Collaboration Strategies in a Variable Demand Environment".* (Doctoral dissertation), Western Michigan University.

Kang, Y. and Gershwin, S.B. 2005. "Information inaccuracy in inventory systems: stock loss and stockout". *IIE Transactions,* Vol.37(9), 843-859.

Khodambashi, S. 2013. "Business process re-engineering application in healthcare in a relation to health information systems". *Procedia Technology,* Vol.9, 949-957.

Kim, D. 2005. "An integrated supply chain management system: A case study in healthcare sector". In *E-Commerce and web technologies*, K. Bauknecht, B. Pröll & H. Werthner (Eds.). Springer Berlin Heidelberg, Vol.3590, 218-227.

Kök, A.G. and Shang, K.H. 2014. "Evaluation of cycle-count policies for supply chains with inventory inaccuracy and implications on RFID investments". *European Journal of Operational Research,* Vol.237(1), 91-105.

Li, Z. 2010. "Evaluating inventory collaboration in healthcare industry: A system dynamic study". (Master thesis), Tilburg University.

Lin, R.-H. and Ho, P.-Y. 2014. "The study of CPFR implementation model in medical SCM of Taiwan". *Production Planning & Control,* Vol.25(3), 260-271.

Liu, E. and Kumar, A. (2003). *"*Leveraging information sharing to increase supply chain configurability". In *Proceeding of the 24th International Conference on Information Systems*, (Seattle, Washington, Dec.15-17), Paper 44.

MacPhee, M. 2007. "Strategies and tools for managing change". *Journal of Nursing Administration,* Vol.37(9), 405-413.

Patwardhan, A. and Patwardhan, D. 2008. "Business process re-engineering–saviour or just another fad?". *International Journal of Health Care Quality Assurance,* Vol.21(3), 289-296.

Robinson, S. 2015. "*A tutorial on conceptual modeling for simulation*". In *the Proceedings of the 2015 Winter Simulation Conference*, (Huntington Beach, California).

Sarac, A., Absi, N., and Dauzere-Peres, S. 2015. "Impacts of RFID technologies on supply chains: a simulation study of a three-level supply chain subject to shrinkage and delivery errors". *European Journal of Industrial Engineering,* Vol.9(1), 27-52.

Tolk, A., Diallo, Y.S., Padilla, J.J., and Herencia-Zapana, H. 2013. "Reference modelling in support of M&S—foundations and applications". *Journal of Simulation,* Vol.7(2), 69-82.

Williams, E.J. and Vanessa, H. 2003. "Improving logistical procedure within a hospital inpatient pharmacy". In Proceeding of the International Workshop on Harbour, Maritime and Logistics Modelling & Simulation, (Riga TU, Latvia, Sep.18-20), 137-141.

Zhang, A.N., Goh, M., and Meng, F. 2011. "Conceptual modelling for supply chain inventory visibility". *International Journal of Production Economics,* Vol.133(2), 578-585.

AUTHOR BIOGRAPHIES

WIRACHCHAYA CHANPUYPETCH is a PhD candidate in Logistics and Engineering Management, Department of Industrial Engineering, Faculty of Engineering, Mahidol University, Thailand. She holds a BSc in Food Science and MSc in Technology and Information System Management. Her research interests focus on logistics and supply chain management, information technology, business process and applications of operation management. Her e-mail address is wirachchaya@gmail.com.

DUANGPUN KRITCHANCHAI was engaged as the project leader for several large-scale projects for hospital logistics and supply chains. She received her PhD in Manufacturing and Operation Management from the University of Nottingham, UK. Her research interests are in healthcare logistics and supply chain management. She is currently serving at Mahidol University, in the capacity of Associate Professor in the Department of Industrial Engineering, Faculty of Engineering and Director in the Healthcare Supply Chain Excellence Centre and Centre of Logistics Management and her e-mail address is duangpun.skr@mahidol.ac.th.

DISCRETE EVENT SIMULATION – PRODUCTION MODEL IN SIMUL8

Jakub Fousek, Martina Kuncova and Jan Fábry
Department of Econometrics
University of Economics in Prague
W.Churchill Sq. 4, 13067 Prague 3, Czech Republic
E-mail: kuba.fousek@seznam.cz; martina.kuncova@vse.cz; jan.fabry@vse.cz

KEYWORDS

Discrete Event Simulation, Production Model, Radial Fans, SIMUL8.

ABSTRACT

Computer simulation is a method for studying complex systems that are not solvable with the use of standard analytical techniques. This contribution deals with the application of simulation program SIMUL8 to the analysis of production process in company Alteko, Inc. producing radial fans. The main purpose of the computer experiments is to identify the bottleneck processes and to suggest the management the appropriate solution. Computer experiments are aimed at the possibility of the parallelization of contracts in terms of the utilization of shared resources. Acceptable requirements for resources are recommended without the necessity of hiring additional operators. Although more suitable software can be used for the analysis of manufacturing processes, the advantage of SIMUL8 consists of its simplicity and interpretability of achieved results.

INTRODUCTION

The main reason for using computer simulation in the analysis of managerial problems is the impossibility of using standard analytical tools due to complexity of real processes. Many production and logistic problems in reality are suitable for simulation approach because of their dynamic and probabilistic character (Banks, 1998). Analyzing the production process, it is usual to use discrete-event simulation. All activities, their sequence, duration and required resources must be defined. O'Kane et al. (2000) show the importance of discrete-event simulation for the decisions to increase in total production output. The automotive industry is a typical area for the application of computer simulation. Masood (2006) investigates how to reduce the cycle times and increase in the machine utilization in an automotive plant. Montevecchi et al. (2007) show the meaning of simulation experiments representing different scenarios and company strategies. In the following text we present the manufacturing problem in company Alteko, Inc. dealing with the production of radial ventilator fans. The contract of producing 50 pieces is the subject of our investigation. As the company has no experience with simulation models we were asked to help them with it. First, a conceptual model will be created. The whole production process is divided into individual activities being executed on corresponding work centers and using prescribed resources. Then, a simulation model will be developed in the environment of SIMUL8. After debugging the model, results obtained from the simulation runs should show the bottleneck parts of the system and possibilities of improvements. The experiments will be performed with the objective to suggest the management the most responsible decision.

SIMUL8

SIMUL8 is a software package designed for Discrete Event Simulation or Process simulation. It has been developed by the American firm SIMUL8 Corporation (www.simul8.com). The software has started to be used in 1994 and every year a new release has come into being with new functions and improved functionality. It allows user to create a visual model of the analyzed system by drawing objects directly on the screen of a computer. SIMUL8 belongs to the simulation software systems that are widely used in industry and available to students (Greasley 2003). This software is suitable for the discrete event simulation but usually it is not used for the simulation of a production. So the task was if it can be a suitable environment for the given situation. Contrary to similar simulation software like Witness or Plant Simulation (Greasley 2003; Bangsow 2010) that are more suited for the production modelling via 3D animation, SIMUL8 uses 2D animation only to visualize the processes. It is similar to SIMPROCESS which is also aimed at the discrete even simulation (Douhý et al. 2011) but we decided to use SIMUL8 because of the easier way of the queue modelling. On the other hand it is a challenge to create a production model in this software. SIMUL8 operates with 6 main parts out of which the model can be developed: Work Item, Work Entry Point, Storage Bin, Work Center, Work Exit Pont, Resource (Concannon et al. 2007).

Main components

Work Item: dynamic object(s) (customers, products, documents or other entities) that move through the processes and use various resources. Their main properties that can be defined are labels (attributes), image of the item (showed during the animation of the simulation on the screen) and advanced properties (multiple Work Item Types).

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Work Entry Point: object that generates Work Items into the simulation model according to the settings (distribution of the inter-arrival times). Other properties that can be used in this object are batching of the Work Items, changing of the Work Items! Label or setting of the following discipline (Routing Out).

Storage Bin: queues or buffers where the Work Items wait before next processes. It is possible to define the capacity of the queue or the shelf life as a time units for the expiration.

Work Center: main object serving for the activity description with definition of the time length (various probabilistic distributions), resources used during the activity, changing the attributes of entities (Label actions) or setting the rules for the previous or following movement of entities (Routing In / Out).

Work Exit Point: object that describes the end of the modeled system in which all the Work Items finish its movement through the model.

Resource: objects that serve for modelling of limited capacities of the workers, material or means of production that are used during the activities.

SIMUL8 uses various graphic components and 2D animation for a process representation. As the simple illustration (created in older version of SIMUL8) we show the model of petrol station that is used at seminars at the University of Economics in Prague (Dlouhý et al, 2011) - Figure 1.

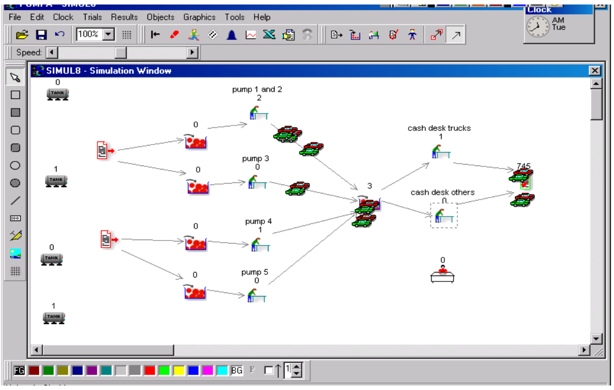

File Edit Clock Trials Results Objects Graphics Tools Help
Speed
SIMUL8 - Simulation Window
pump 1 and 2
pump 3
pump 4
pump 5
cash desk trucks
cash desk others
Clock

Figure 1: Simulation model of a petrol station in SIMUL8

All objects (except resources) are linked together by connectors that define the sequence of the activities and also the direction of movement of Work Items. The sequence is clear from the Figure 1: there are two Work Entry Points out of which the Work Items (cars) can go to four queues (Storage Bins) waiting to fuel up (four Work Centers), then all continue to one queue (Storage Bin) for two cash desks (Work Centers) and then they are leaving the system (Work Exit Point). All the resources (4 pumps and cash desk) are used during the activities but they are not linked by a line with the other objects.

After the system is modelled, simulation run follows. The animation shows the flow of items through the system and for that reason the suitability of the model can be easily assessed. When the structure of the model is verified, a number of trials can be run under different conditions. Then, the performance of the system can be analyzed statistically. Values of interest may be the average waiting times or utilization of Work Centers and Resources (Shalliker and Rickets 2002).

SIMUL8 can be used for various kinds of simulation models (Concannon et al. 2007). The case studies can be seen also on the website www.simul8.com.

Our experience shows that SIMUL8 is easy to learn (especially when the main components are used without the necessity to use Visual Logic (with different programming functions). It can serve not only for the modelling of different services (Dlouhý et al. 2011), but also for the simulation of various production processes (Ficová and Kuncová 2013).

PROBLEM DESCRIPTION

One of assembled products is low-pressure radial ventilator fan RFC (Figure 2). It is produced in several designs and sizes, in contribution RFC 200 is analyzed. Company Alteko, Inc. characterizes it as follows (2017): „Low-pressure radial fans RFC are one-side suction fans driven directly by flange-mounted motors (IP55). RFC fans are designed for air exchange in residential and industrial premises. The RFC fans may not be used for transporting the air which includes aggressive agents, abrasive additives and fibrous particles. The temperature of transported air may fluctuate between -30 °C into + 85 °C (by fans Ex - 30°C into +40°C)."

Figure 2: Radial fan RFC 200, company ALTEKO, Inc.

MODELLING OF PROCESS

The objective of the computer simulation is to observe whether it is achievable to produce 50 pieces of fans within 30 working days at present production level. The production process belongs to the longest activities in the contract. Based on the analysis using project management and PERT method (Fousek 2016), it has been found out that the production takes about 50 % of the total time to perform the contract. From this point of view, it is necessary the production duration does not exceed 15 working days. The exploration of the utilization of resources is also an integral part of the analysis. As the input entity (Work Item) used in the model we defined material that is continuously processed and assembled with other components and semi-finished products. JIT delivery is not the subject of the model, because the strategy of the company gives sufficient amount of material to the production process. Shift calendar is fixed to 8 net working hours per week (5 working days); breaks and disorders are omitted. Based on the consultation with the director of the company, availability of all workers at the contract was set to 50 % in the first level.

In production process the following resources are temporarily used: machines, work centers and various operators (see Table 1). Because some workers operate at more centers or machines, it was necessary to estimate their movement and its duration.

Table 1: List of resources for RFC 200

Machine / Work center	No. of machines	Operator (ID)
Trumatic	2	2
Cutting	1	3
Folding hub	1	3
Stamp	3	4
Balancer	1	5, 6
Spinning lathe	1	1
Folder	1	7
Lath	2	8, 9
Driller	5	3
Assembling	0	10
Spot weld	2	11
Bending	1	12
Testing room	0	13
Engine depot	0	14
Store	0	14

Assembling of fan requires the sequence of jobs with fixed durations, but after the consultation with workers they were approximated by uniform (UNI) or triangular (TRI) probability distribution. Table 2 summarizes this assignment to resources and work centers with two exceptions determined on the base of the exact calculation (Fousek 2016). These are a transport of engine from warehouse and shipping the final product to the warehouse of finished goods.

Table 2: List of resources, work centers and distribution of jobs duration

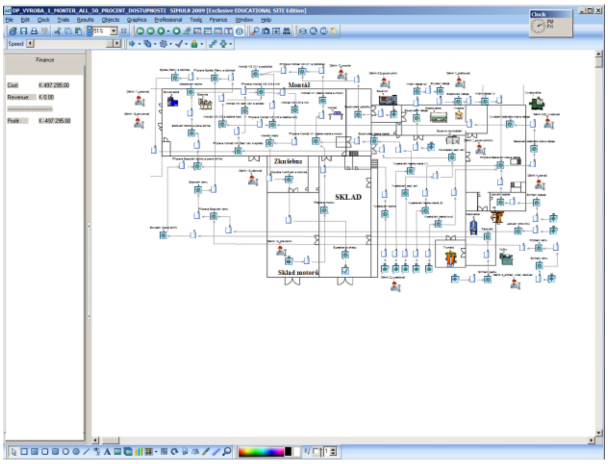
Job	Machine / Work center	Operator	Dist. (min)
Bearing plate (1) cutting out	Trumatic	2	UNI (2, 3)
Bearing plate (2) cutting out	Trumatic	2	UNI (2, 3)
Bearing plate (1) turning	Lathe	8, 9	UNI (7, 10)
Liner turning	Lathe	8, 9	UNI (2, 5)
Engine – transport from depot	Engine depot	14	2,07
(M1:) Bearing plate assembly + engine	Assembly	10	TRI (10, 12, 15)
(M2): Assembly (M1) + liner	Assembly	10	UNI (1, 2)
Turning	Lathe	8, 9	UNI (10, 15)
Drilling hub	Lathe	8, 9	UNI (5, 8)
Lathing hub	Lathe	8, 9	UNI (3, 7)
Grooving hub	Groover	7	UNI (7, 8)
Drilling hub	Driller	3	UNI (4, 5)
Cutting blades	Scissors	3	UNI (2, 3)
Stamping blades	Stamper	4	UNI (4, 8)
Dinking cover plate	Trumatic	2	UNI (2, 3)
Assembling rotor wheel and hub	Balancer	5, 6	TRI (10, 15, 21)
Balancing rotor wheel and hub	Balancer	5, 6	TRI (5, 10, 14)
(M3): Assembling (M2) + rotor wheel	Assembler	10	TRI (2, 3, 6)
Dinking pipe mouth	Trumatic	2	TRI (2, 3, 6)
Spinning pipe mouth	Spinning lathe	1	TRI (4, 5, 7)
Cutting plate	Scissors	3	UNI (5, 6)

Folding	Folder	3	UNI (3, 4)
Bending	Bending rolls	12	UNI (3, 5)
Dinking sideboards	Trumatic	2	UNI (4, 6)
Sideboards and right box spot weld	Spot welder	11	UNI (10, 14)
Sideboards spot weld finish	Spot welder	11	UNI (3, 4)
Cutting frames	Scissors	3	UNI (3, 6)
Bending frames	Stamper	4	UNI (4, 6)
Framework spot weld	Spot weld	11	UNI (4, 6)
Angle plate corner bracket pressing	Press	4	UNI (2, 6)
Sideboards and right box spot weld (“spiral”)	Spot weld	11	TRI (4, 6, 9)
(M4): Intake port assembly + “spiral“	Assembly	10	TRI (5, 6, 8)
(M5): Assembly (M3) + (M4)	Assembly	10	TRI (10, 12, 16)
Handrail assembly	Assembly	10	TRI (5, 6, 9)
Functionality testing	Testing room	13	UNI (4, 6)
Dispatch to store	Store	14	1,17

MODEL IN SIMUL8

The simulation model was developed in SIMUL8 software. During the modelling process it was necessary to define times for movements of operators between work centers (company data was explored – Fousek 2016). Figure 4 shows the scheme of the whole model.

In verification step it was necessary to check whether the ideas from conceptual model were correctly transformed into computer simulation model. We concentrated on queueing congestion due to wrong way of modelling or real foundation of lower utilization of resources. All factors seemed to be taken into account correctly. Exactly 50 units of RFC 200 entered the system at once. All resources were used in the model that is proved in Table 3. Jobs dependent on resources movement (Operator 3_ Trimmer, Folder, Driller and Operator 14_Transfers), were actually started after workers moved to the required work center.


DP_VYROBA_1_MONTER_ALL_50_PROCENT_DOSTUPNOSTI SIMUL8 2009 [Exclusive EDUCATIONAL SITE Edition]
File Edit Clock Trials Results Objects Graphics Professional Tools Finance Window Help
Speed
Clock
PM
Fri
Finance
Montáž
Zkušebna
SKLAD
Sklad motorů

Cost K 497 295.00
Revenue K 0.00
Profit K -497 295.00

Figure 4: Production Model in SIMUL8

The model is classified as the model with finite horizon, because exact simulation time is stated (14 400 minutes, i.e. 30 working days). Although we know 15 days should be the maximal time for the production we set 30 days as the total maximum to meet the order to see where are the time limits of the production. However, simulation run can be ended earlier in case all fans have been finished. To eliminate negative random events we performed 100 simulation runs with derived 95 % confident intervals for watched random variables.

COMPARISON WITH THE REAL SYSTEM

Despite correctly specified parameters, fixed availability level 50 % for all operators can differ in reality due to utilization of workers by other tasks in other contract activities. Therefore, it is the value that can fluctuate in real system according to a number of contracts and their capacity requirements. On the contrary, according to experience, total production time for producing 50 fans RFC 200 approximately corresponds the situation all operators are busy at the contract on 50 % of their working time.

Total production time is 4876.84 minutes that is 81.28 working hours, i.e. 10.16 working days. The 95 % confidence interval for total production time is <4854.60, 4899.08> minutes, i.e. 10-11 days. In this situation the limit of 15 days necessary for finishing contract would not be exceeded.

Assembly operator is the busiest one, at 94.14 % (Table 3). The assembling utilization is 94.47 % what is, according to available information, close to real experience. The utilization of other resources seems to stay in acceptable range, no one is busy over 80 %. On the contrary, the least-busy resource is obviously the drill (51 %) because there are 5 drills available and only one short job needs this type of resource. Exploring the work centers utilization in detail, we find out the analogous situation in case of jobs requiring the assembly. They are in the waiting mode in major time of the process (as the assembly operator is busy). For example, assembly M2 (M1 and bearing plate) waits 90.37 % of time for the resource (assembly operator) and 8.09 % of time for the entity and only 1.54 % of time it works. The job which has, in comparison with other jobs, the highest proportion of work performed, is assembling rotor wheel and hub (15.70 %). However, more than half of the production time it waits for the entity what concerns of most jobs with the exception of assemblies mentioned above. Assembly logically starts when all components are available. In addition, operations on this work center are most time-consuming ones. It leads to high utilization of resources and consequent generation of long queues.

Table 3: Resource Percent Utilization (Util.)

Resource	**No. of resources**	**Util. (%)**
Spot welder	2	63.79
Operator 1_Turner	1	55.45
Operator 10_Assembly	1	94.14
Operator 11_Spot welder	1	77.27
Operator 12_Bender	1	54.23
Operator 13_Tester	1	55.16
Operator 14_Transfers	1	53.09
Operator 2_Cutting	1	66.78
Operator 3_Trimmer, Folder, Driller	1	72.19
Operator 4_Presser	1	65.18
Operator 5,6_balancing	2	62.75
Operator 7_Scrubbing	1	57.62
Operator 8,9_Turner	2	68.40
Scrubbing tool	1	57.61
Folder	1	54.01
Presser	3	55.06
Assembly	1	94.47
Cutting	1	63.30
Bending	1	53.97
Lathe	2	68.41
Spinning lathe	1	55.53
Trumatic	2	58.45
Drill	5	51.00
Balancing tool	1	75.70

EXPERIMENTS WITH MODEL

According to previous analyzes, the simulation model obviously shows that the assembly is the bottleneck of the whole production process. There is only one operator that is enormously utilized. In computer experiments another operator was added to assemble products. In addition, lower availability of resources was tested (40 % a 30 %). Table 4 shows results for 100 experiment runs.

Table 4: Resource Percent Utilization When Available

Experiment	Confidence interval (hours)		Average production time (hours/days)
	Min	Max	
1 worker 50%	80.91	81.65	81.26 / 10.16
2 workers 50%	46.46	47.00	46.73 /5.84
1 worker 40%	103.03	104.10	103.56 / 12.95
2 workers 40%	58.73	59.47	59.10 / 7.39
1 worker 30%	**142.32**	**144.26**	**143.31 / 17.91**
2 workers 30%	80.32	81.60	80.96 / 10.12

From the results in Table 4 it is clear that if only one operator would be available and his availability is 30 %, the contract would not be finished in time. Therefore, the company must be very cautious in case of realizing parallel contracts to assure the availability at 40 – 50 %, or it should consider to hire an additional operator to assembly job.

CONCLUSION

The aim of the contribution was to demonstrate the applicability of SIMUL8 even in modelling of production processes, for which it is not quite eligible software (e.g. in comparison with Plant Simulation or Witness). According to our experience we can conclude that this software can be used to model the production process - with some simplifications of the reality and with a few special settings forced by SIMUL8. The model, used for the production of 50 radial fans RFC 200 in company Alteko, Inc. was developed on the basis of available information given by employees and obtained from the company's internal documentation. The simulation model should have shown whether it is realistic to perform the contract in 15 days considering the realization of parallel contracts. Although the situation was slightly simplified (e.g. fixing of the percent availability of operators), model corresponds the real situation and proves a possibility to finish the production within required 15 days in case the operator working at bottleneck work center (assembly) would not be engaged in other jobs more than 60 % of his working time. The results recommend to managers what they must concentrate on and what contracts can be parallelized with the analyzed contract. The main success we see in the fact that we have not only created the model to find the solution but we have persuaded the managers of the usability and the advantage of the simulation itself.

ACKNOWLEDGEMENTS

This work was supported by the grant No. F4/57/2017 of the Faculty of Informatics and Statistics, University of Economics, Prague.

REFERENCES

Aguirre, A., and C.A. Mendez. 2008. “Applying a Simulation-based Tool to Productivity Management in an Automotive-parts Industry”. In *Proceedings of the 2008 Winter Simulation Conference*. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, 1838-1846.

Alteko. 2016. “Radiální ventilátory s přímým pohonem RFC, RFE”. ALTEKO. [Online] [Cit: 2016-04-08] Available:http://www.alteko.cz/922_100020-radialni-ventilatory-s-primym-pohonem-rfc-r.

Bangsow S. 2010. *Manufacturing Simulation with Plant Simulation and SimTalk*. Springer-Verlag Berlin Heildelberg

Banks, J., 1998. *Handbook of Simulation*. USA, John Willey & Sons

Concannon, K. et al. 2007. *Simulation Modeling with SIMUL8*. Visual Thinking International, Canada.

Dlouhý, M., Fábry, J., Kuncová, M. and T. Hladík. 2011. *Business Process Simulation* (in Czech). Computer Press.

Ficová, P. and M. Kuncová. 2013. „Looking for the equilibrium of the shields production system via simulation model“. In *Modeling and Applied Simulation 2013*. (Athens, Sept. 25-27). Genova : DIME Universita di Genova, 50–56.

Fousek, J. 2016. „Využití simulačních modelů a programů k analýze a zlepšení chodu výrobního podniku“. Master Thesis, University of Economics, Prague.

Greasley, A. 2003. *Simulation modelling for business*. Innovative Business Textbooks, Ashgate, London.

Kuncová, M. and P. Wasserbauer. 2007. „Discrete Event Simulation – Helpdesk Model in SIMPROCESS“. In *ECMS 2007* (Prague, June 4-6). Dudweiler : Digitaldruck Pirrot, 105–109.

Law, A. 2000. *Simulation Modelling and Analysis. Boston (USA),* MC-Graw Hill

Masood, S. 2006. Line balancing and simulation of an automated production transfer line, *Assembly Automation*, Vol. 26 Iss: 1, 69 – 74.

Montevechi, J. A. B., et al. 2007. Application of design of experiments on the simulation of a process in an automotive industry. In *WSC’07 Proceedings of the 39th Conference on Winter Simulation*, IEEE Press Piscataway, NJ, USA, 1601-1609.

O’Kane, J.F. et al., 2000. Simulation as an essential tool for advanced manufacturing technology problems. *Journal of Materials Processing Technology* 107/2000, 412-424

Shalliker, J. and C. Ricketts. 2002. *An Introduction to SIMUL8, Release nine*. School of Mathematics and Statistics, University of Plymouth.

AUTHOR BIOGRAPHIES

JAKUB FOUSEK was born in Prague, Czech Republic. He studied at the University of Economics Prague, study programme Quantitative Methods in Economics, study field Econometrics and Operational Research. He graduated in 2016. Economics in Prague. Master´s degree Econometrics and Operational Research he graduated with honors. He was a member of Student club of Project Management at the University of Economics in Prague for 5 months. Since September 2016 he has been working as Media and Data Analyst in Newton Media, Ltd. in Prague. He is responsible for database functionality of data media and he analyzes media inputs according to long-term criteria. His email address is: kuba.fousek@seznam.cz

MARTINA KUNCOVÁ was born in Prague, Czech Republic. She has got her degree at the University of Economics Prague, at the branch of study Econometrics and Operational Research (1999). In 2009 she has finished her doctoral study at the University of West Bohemia in Pilsen (Economics and Management). Since the year 2000 she has been working at the Department of Econometrics, University of Economics Prague, since 2007 also at the Department of Economic Studies of the College of Polytechnics Jihlava (since 2012 as a head of the department). She is a member of the Czech Society of Operational Research, she participates in the solving of the grants of the Grant Agency of the Czech Republic, she is the co-author of four books and the author of many scientific papers and contributions at conferences. She is interested in the usage of the operational research, simulation methods and methods of multi-criteria decision-making in reality. Her email address is: martina.kuncova@vse.cz

JAN FÁBRY was born in Kladno, Czech Republic. In 1993 he was graduated in Operational Research at the University of Economics Prague (UEP) and in 2006 he received his Ph.D. in Operational Research and Econometrics at the UEP. In 2015 he successfully completed the habilitation procedure in Econometrics and Operational Research at the UEP. Since 2002 he has been working at the department of econometrics at the UEP, initially as an assistant professor, later (since 2015) as an associate professor. In 2016 he joined SKODA AUTO University (SAU) in Mlada Boleslav as a member of the Department. of Logistics, Quality and Automotive Technology. He participates in projects founded by Grant Agency of the Czech Republic, he is the author or co-author of three books and many papers and contributions at conferences. At SAU, he is a supervisor of the courses of Operations Research and of Computer Simulation in Logistic Processes. At the UEP, he is a supervisor of the Czech courses of Discrete Models Case Studies in Operations Research, and of the English courses of Combinatorial Optimization and Operations Research. He is interested in Vehicle Routing Problems and the application of mathematical methods and simulation in production and logistics. Since 2002 he has been the secretary of the Czech Society for Operations Research. His email address is: jan.fabry@vse.cz

CONTEXT-AWARE MULTI-OBJECTIVE VEHICLE ROUTING

Jānis Grabis Vineta Minkēviča

Institute of Information Technology
Riga Technical University, Kalku 1
Riga, LV-1658, LATVIA

KEYWORDS
Vehicle routing, multiple objectives, context

ABSTRACT

A vehicle routing deals with designing optimal delivery routes for a fleet of vehicles serving spatially distributed customers. There is a multitude of variants of this problem varying according to problem characteristics and assumptions. Routing is guided by multiple objectives and affected by different context factors. This paper formulates a vehicle routing optimization model allowing for incorporation of arbitrary selected context factors and objectives represented by KPI. In order to assess relative importance of the objectives and context factors, an adaptive approach is proposed. Experimental studies are conducted to illustrate application of the model and the adaptation method.

INTRODUCTION

Vehicle routing deals with finding a set of routes served by multiple vehicles that jointly traverse a number of customers (Breakers et al., 2016). Eksioglu et al. (2009) proposed the taxonomy of the vehicle routing problem. The taxonomy indicates that there are many variants of this problem. These variants differ by decision-making objectives, constraints, parameters and other factors considered. Koç et al. (2016) focus on structural variants routing models. They show derivation of different types of vehicle routing problem variants on the basis of the general model. The vehicle routing problem is inherently multi-objective problem (Current and Marsh, 1993). Optimization of costs, time and travelling distance are typical objectives (Jozefowiez et al., 2008). Additionally, environmental issues (Xiao and Konak, 2016), safety concerns (Carotenuto et al., 2007) and other factors are often mentioned as relevant. However, majority of models consider only one or two objectives. Execution of the routes obtained as a result of optimization is often affected by contextual factors such as traffic accidents (Psaraftis, 1995) and traffic intensity variations over time (Kok et al., 2010). Quality of routing results could be improved if all relevant contextual factors are taken into account already during the route optimization. If every context factor is treated individually, diversity of context factors might lead to a large number of highly specialized routing models. In order to streamline representation of multiple objectives and context, it is possible to treat these factors in a uniform manner.

The objective of the paper is to formulate an optimization model for multi-objective context-aware vehicle routing that supports usage of arbitrary selected objectives and context factors. Different decision-making objectives are represented by their measurements or Key Performance Indicators (KPI). The optimization is performed to minimize travel cost and deviations of actual values of KPI from their target values. KPI used in the objective function are selected for every specific case depending on decision-making needs. The travel cost is context dependent and is a composite of travel distance, time and context factors affecting route execution. Impact of the context factors and importance of KPI is not always known in advance, therefore that is evaluated and incorporated in the optimization model using an adaptive procedure. Application of the proposed model is illustrated using an example.

The rest of the paper is organized as follows. The next sections describe the vehicle routing problem considered. That is followed by model formulation and illustrative example. The paper completes with the concluding section.

PROBLEM STATEMENT

A company providing logistics services operates a fleet of vehicles. It receives customer service requests on the periodical basis. The customers should be visited within a specified time window. The vehicles should be routed to serve the customers at minimum cost where the cost can be expressed as a sum of multiple factors. The routes start and end at a depot. The main decision variables are vehicle allocation to customers and vehicle arrival time at the customer. The routing problem is formulated as a mathematical programming model and optimal routes are found by performing route optimization.

The company has multiple vehicles routing objectives including customer services level satisfactions, environmental impact reduction and ensuring a safe working environment. The objectives are measured by a set of KPI. Every KPI has a target value specified by management. The route optimization should be performed to take into account these specific KPI and their deviation from the target value. Actual values of KPI depend upon routing decisions made. The route

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

execution is affected by several case specific context factors such as weather, traffic accidents and calendar events. The context factors are beyond company's control.

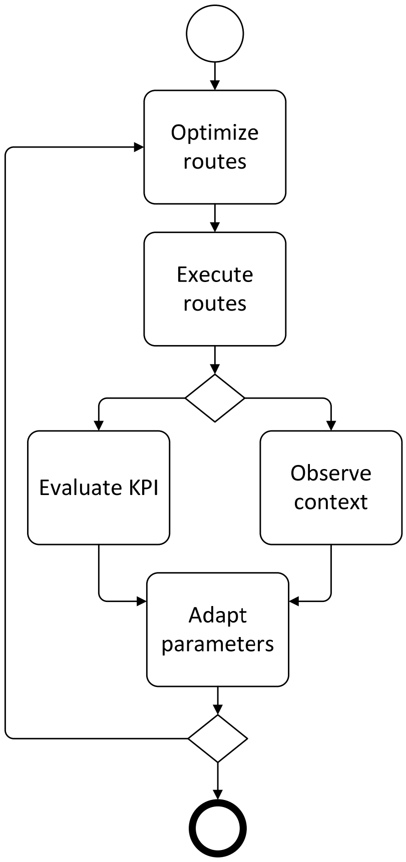


Figure 1: The performance-driven context-aware routing optimization, execution and adaptation process.

Route planning and execution occurs on regular basis. For example, a set of customer requests is received at the beginning of each day, optimal routes are found and customers are visited during the day following these routes. Performance data are accumulated and context data are observed during the route execution. These data are compared with the planned values and deviations are observed. In particular, the actual KPI values are evaluated and compared with those estimated during the route optimization. One of the reasons of potential deviations is that different KPI are mutually contradicting and the right trade-off among the objectives has not been achieved. That can be remedied by changing relative importance of KPI represented by appropriate parameters in the optimization models. The change is performed in an adaptive manner because the right balance is not known in advance.

Similarly, context values are observed and these observations can be used to evaluate relationships among them, decisions-made and performance achieved. This way one can estimate impact of context on performance and this information can be incorporated in the optimization model in an adaptive manner.

The aforementioned route optimization, execution and adaption process is shown in Figure 1.

MODEL

The vehicle routing model is formulated as a mathematical programming model. The formulation is based on traditional routing models (e.g., Kallehauge et al., 2005), which are augmented to include treatment of multiple objectives represented by KPI and to account for impact of different context factors.

Notation

i,j – client indices
k – route type index
l – vehicle index
h – KPI index
n – context element index

Decision variables

$X^l_{ijk} \in \{0,1\}$ - *l* vehicle travels from client *i* to *j* client along *k* route
T^l_i - arrival time of *l* vehicle at customer *i*
P_h - penalty for not achieving *h* KPI
$x_{i'} \in \{0,1\}$ - is path *i'* travel as a part of directions from one customer to another

Parameters

g_{ijk} - cost of route type *k* from customer *i* to *j*
τ_{ijkt} - travel time of *k* route type from customer *i* to *j*
$\left(t^s_i, t^b_i\right)$ - visiting time window for customer *i*
w_m - weights balancing importance of composite cost and other KPI
v_h - weights indicating relative contribution of *h* KPI
$\mu_{ii'}$ - route between customers *i* and *i'*
$d_{i'}$ - length of path *i′*
$ctx^n_{i'}$ - value of *n* context factor for path *i′*
$\boldsymbol{\omega}$ - weights balancing distance, time and other context factors in composite costs calculation for *k* route type
σ_n - weights indicating relative contribution of *n* context element
μ_{ij} - route between customers *i* and *j*

Objective function

The objective function (Eq. 1) minimizes a weighted sum of the total travel cost and the penalty function for failing to achieve target values of KPI (Eq. 2). The total travel cost g_{ijk} parameter represents not only directs costs for covering some distance or spending time on the trip but also impact of contextual factors. The objective function combines traditional minimization of costs and minimization of non-performance penalty what allows to capture both structural and managerial decision-making characteristics of the vehicle routing problem.

$$\min Z = w_1 \sum\nolimits_{l=1}^{L} \sum\nolimits_{i=1}^{N} \sum\nolimits_{j=1}^{N} \sum\nolimits_{k=1}^{K} g_{ijk} X_{ijk}^{l} + w_2 P \quad (1)$$

$$P = \sum\nolimits_{h=1}^{H} v_h P_h \quad (2)$$

Constraints

The objective function is minimized subject to traditional vehicle routing constraints and additional constraints to represent context-awareness and evaluation of KPI.

$$\sum\nolimits_{l=1}^{L} \sum\nolimits_{j=1}^{N} \sum\nolimits_{k=1}^{K} X_{ijk}^{l} = 1, \forall i \quad (3)$$

$$\sum\nolimits_{j=1}^{N} \sum\nolimits_{k=1}^{K} X_{0jk1}^{l} = 1, \forall l \quad (4)$$

$$\sum\nolimits_{i=1}^{N} \sum\nolimits_{k=1}^{K} X_{ii'k}^{l} - \sum\nolimits_{j=1}^{N} \sum\nolimits_{k=1}^{K} X_{i'jk}^{l} = 0, \forall l, \forall i' \quad (5)$$

$$\sum\nolimits_{i=1}^{N} \sum\nolimits_{k=1}^{K} X_{iN+1k}^{l} = 1, \forall l \quad (6)$$

$$t_i^s \le T_i^l \le t_i^b, \forall i \quad (7)$$

$$X_{ijk}^{l} \left(T_i^l + \tau_{ijk} - T_j^l \right) \le 0, \forall i, j, k, l \quad (8)$$

$$KPI_h^{Current} \le KPI_h^{Target} - P_h, \forall h \quad (9)$$

$$\min g_{ijk} = \omega_{k1} \sum_{x_{i'} \in \mu_{ij}} d_{i'} x_{i'} + \omega_{k2} \sum_{x_{i'} \in \mu_{ij}} \tau_{i't} x_{i'} + \omega_{k3} CTX_{i'} \quad (10)$$

$$CTX_{i'} = \sum\nolimits_{n=1}^{N} \sigma_n \sum_{x_{i'} \in \mu_{ij}} ctx_{i'}^{n} x_{i'} \quad (11)$$

$$\mu_{ij} = \left\{ x_{i'} \mid x_{i'} = 1 \right\} \quad (12)$$

Eq.3 specifies that every customer should be visited exactly once. Eq. 4 imposes that trips start at the depot (referred as location zero). Incoming and outgoing flows are balanced by constraint (5). All trips end at the depot (6). The customer service time windows should be observed (7). Eq. 8 relates arrival times and subsequent customer and traveling time between the customers. Eq. 9 imposes a penalty if target values of KPI are not achieved. For every route between customers i and j, k different best paths (according to their composite cost) are found (10). The different variants are obtained by exploring various combinations of weights $\boldsymbol{\omega}$. For instance, one set of weights favors the shortest path while another set of $\boldsymbol{\omega}$ favors the safest path. Additional constraints should be added for calculating estimated KPI values depending on the decision variables.

The treatment of KPI and context factors is done in a uniform manner allowing to incorporate case specific KPI and context elements in the model with relative ease.

ADAPTIVE PARAMETERS

The routing model depends on a number of weighting parameters. The initial values of these parameters are specified in a judgmental manner. Subsequently, they are continuously updated to improve routing performance.

The relative importance of KPI in Eq.2 is determined by the weight coefficients v_h. New values of these coefficients are calculated as

$$v_h^{new} = \frac{v_h + v_h^{'}}{\sum_h v_h},$$

where $v_h^{'}$ is the adjustment for the hth weight and v_h^{new} is the adapted value of the weight factor to be used in the next routing run. The adjustment is determined by maximizing the weighted total penalty (i.e., the biggest increment should be given to KPI with the largest penalty)

$$\max \Pi = \sum\nolimits_{h=1}^{H} v_h^{'} P_h$$

The sum of weights is required to be equal to one and the adjustment in a single step cannot exceed a specified threshold.

It is expected that there is a relationship between the context factors and observed values of KPI. This relationship is evaluated using a regression equation

$$KPI = \alpha_0 + \sum\nolimits_{n=1}^{N} \alpha_n \bar{ctx}_n,$$

where $\bar{ctx}_n$ is an aggregated context value and α_n are coefficients of the regression equation.

The weights characterizing relative importance of every context factor are obtain by ranking the context factors according to alpha and computing them as

$$\sigma_n = r^{-1} \sum\nolimits_{n=1}^{N} n^{-1}$$

where u=1,2,...r are rankings and $\sum\limits_{n=1}^{N} \sigma_n = 1$

It is important to note that the KPI values in the routing model are estimated values calculated during the route planning activity while adaptation is performed using the actual values evaluated during the route execution. The adaption is performed periodically once information about route execution is accumulated in the transportation planning application.

APPLICATION EXAMPLE

Application of the proposed model is demonstrated using an example. This example is aimed at illustrating context dependency of routing results and adaptation of the model parameters.

It is assumed that case specific KPI are KPI1) customer service measured as a percentage of the clients served during the specified time windows; KPI2) travel cost calculated as time spent on deliveries times hourly rate; KPI3) vehicle operating cost incurred for every vehicle used on a given day regardless of distance travelled; and KPI4) safety aimed at avoiding traversal of accident

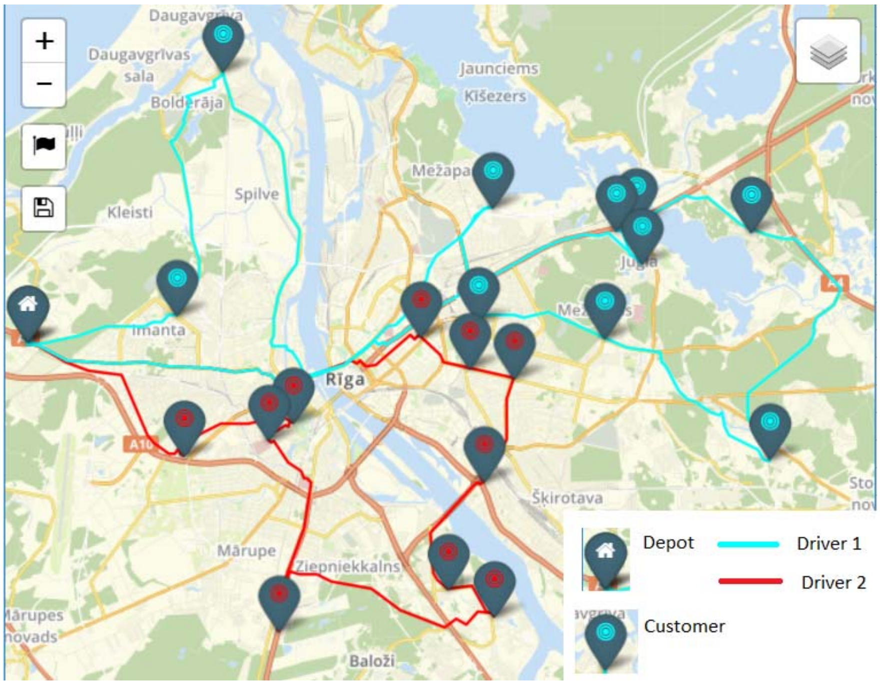


Figure 2: Routing results for EXP1.

prone routes measured by an index characterizing frequency of the accidents. Case specific context elements affecting route execution: CTX1) route variability measured as variation of driving time from day to day; and CTX2) route safety measured as a number of accidents observed for the given route.

Routing is performed for 20 client requests received for a single day. The optimization is performed using ILOG. The travel distance and time data are retrieved from OpenStreetMap (https://www.openstreetmap.org). The accident data are gathered from a web mapping service. For every pair of customers, three different paths are obtained by varying the context dependency weights in Eq. 10 (Table 1). The path type is referred as Short because the best route between two customers is found giving the most importance to the distance minimization. The path type is referred as Safe because the largest weight is given to the context factors including the safety context element CTX2.

Table 1: The weights used to find the best path between two customers.

Path type	ω_1	ω_2	ω_3
Short	0.8	0.1	0.1
Safe	0.1	0.1	0.8
Balanced	0.34	0.33	0.33

The optimization is performed by allowing to select any of the paths (EXP1), only the shortest path (EXP2), only the safe path (EXP3) and only the balanced path (EXP4). The values of KPI obtained for these four experiments are reported in Table 2. These values are reported relative to EXP1 or the optimal case but *Z*, which is the actual objective value observed and is a weighted sum of all criteria. It can be observed that EXP2 yields the best result in term of actual costs but neglects the impact of context factors (high value of the cost) and delivers weak customer service performance. Similarly, EXP3 selects safe paths and scores the best according to the safety KPI4. Selecting balanced path (EPX4) expectedly yields results close to the optimal. None of the experiments yields satisfactory customer service performance (KPI1). The actual values recorded were 65 to 75% while the KPI target was 100%. It is important to emphasize that different routes were obtained in different experiments indicating that route selection is indeed context dependent.

Table 2: Values of KPI

Experiment	Z	Cost	KPI1	KPI2	KPI3	KPI4
EXP1	0.28	1.00	1.00	1.00	1.00	1.00
EXP2	0.40	1.76	0.86	0.75	0.50	0.95
EXP3	0.72	3.41	1.00	1.85	1.50	0.38
EXP4	0.30	1.10	1.07	1.08	1.00	0.97

As mentioned before KPI1 did not yield satisfactory performance. The adaption is performed to alter balance among KPI in the objective function, i.e., by changing the weights **v**. Five adaption cycles are performed for EXP1. The same set of customer requests is used in all five cycles though different customer requests would be expected in real life situations. Figure 3 shows the adaption results. KPI values are reported relative to the target values. Values above one indicates that the KPI target value has been achieved. In the first cycle the set of weights **v** has values (0.25,0.25,0.25,0.25). Given these parameters, the target values are not achieved for KPI1 and KPI2. Adaption allows to reach the target value for KPI2 already after the third cycle with **v**=(0.32,0.16,0.36,0.16). The value of KPI1 changes from 0.65 to 0.75 though the target value cannot be achieved. The final set of weights is (0.277,0.102,0.518,0.103).

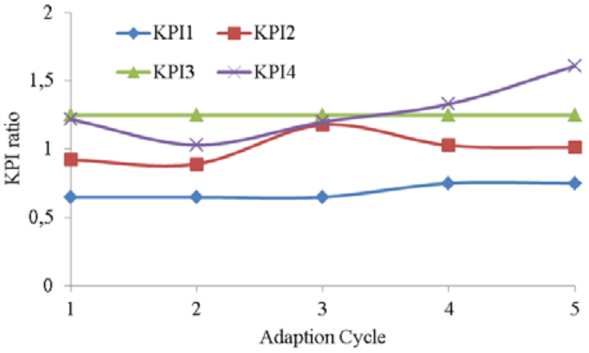


Figure 3: Change of KPI according to adaption cycle

CONCLUSION

The vehicle routing model has been elaborated. The distinctive features of this model are the uniform treatment of decision-making objectives and context elements what allows to include arbitrary or case specific objectives and context elements in the model. It is shown that routes developed are context dependent and the adaptation allows for balancing impact of different objectives.

The vehicles routing model is computationally demanding and an algorithm for improving computational efficiency should be elaborated. The adaptation procedure currently deals with the weights assigned to individual KPI. This procedure should be made more comprehensive to cover all judgmental parameters included in the optimization model. Properties and behavior of the adaption procedure also should be explored in more details.

The model is developed as a part of the industrial research project where the customer service level has been identified as a major vehicle routing concern. Simulation could be used to better evaluate relationships among context elements and the customer service level. It also could be used to analyze sensitivity and stability of the adaptation process.

REFERENCES

Braekers, K., K. Ramaekers and I. van Nieuwenhuyse. 2016. “The vehicle routing problem: State of the art classification and review.” *Computers & Industrial Engineering* 99, 300–313.

Carotenuto, P., S. Giordani, S. Ricciardelli and S. Rismondo. 2007. “A tabu search approach for scheduling hazmat shipments.” *Computers & Operations Research* 34, No.5, 1328-1350.

Current, J. R., & Marsh, M. 1993. “Multiobjective transportation network design and routing problems: Taxonomy and Annotation”. *European Journal of Operational Research* 65, 4–19.

Eksioglu, B., A.V. Vural and A. Reisman. 2009. “The vehicle routing problem: A taxonomic review.” *Computers & Industrial Engineering* 57, 1472-1483

Jozefowiez, N., F. Semet, and E.G. Talbi., 2008. “Multi-objective vehicle routing problems.” *European Journal of Operational Research* 189, No.2, 293-309.

Kallehauge, B., J. Larsen, O.B.G. Madsen and M.M. Solomon 2005. "Vehicle routing problem with time windows" in *Column Generation*, Desaulniers et al. (eds.) Springer, pp. 67-98.

Koç, C., T. Bektaş, O. Jabali and G. Laporte. 2016. "Thirty years of heterogeneous vehicle routing." *European Journal of Operational Research* 249, No. 1, 1-21.

Kok, A. L., C.M. Meyer, H. Kopferand J.M.J. Schutten. 2010. “A Dynamic Programming Heuristic for the Vehicle Routing Problem with Time Windows and European Community Social Legislation.” *Transportation Science* 44, No.4, 442–454.

Psaraftis, H. N. 1995. “Dynamic Vehicle Routing: Status and Prospects”, *Annals of Operations Research* 61, No. 1, 143–164.

Xiao, Y. and A. Konak. 2016. “The heterogeneous green vehicle routing and scheduling problem with time-varying traffic congestion.” *Transportation Research Part E* 88, 146-166.

AUTHOR BIOGRAPHIES

JĀNIS GRABIS is a Professor at the Faculty of Computer Science and Information Technology, Riga Technical University, Latvia. He obtained his PhD from the Riga Technical University in 2001 and worked as a Research Associate at the College of Engineering and Computer Science, University of Michigan-Dearborn. He has published in major academic journals including OMEGA, European Journal of Operational Management, International Journal of Production Research, Computers & Industrial Engineering, IEEE Engineering Managenet Review and others. He has been a guest-editor for two top academic journals and member of the program committee of several academic conferences. His research interests are in supply chain management, enterprise applications and project management. His email address is: `grabis@rtu.lv`.

VINETA MINKĒVIŠA is a docent at the Faculty of Computer Science and Information Technology, Riga Technical University. He holds a doctoral degree in mathematics. Her major areas of interest are operations research, queueing systems, Markov processes and project risk management.

A SIMULATION OPTIMIZATION TOOL FOR THE METAL ACCESSORY SUPPLIERS IN THE FASHION INDUSTRY: A CASE STUDY

Virginia Fani
Romeo Bandinelli
Rinaldo Rinaldi

Department of Industrial Engineering
University of Florence
Florence, Viale Morgagni 40/44, 50134
E-mail: virginia.fani@unifi.it, romeo.bandinelli@unifi.it, rinaldo.rinaldi@unifi.it

KEYWORDS
Discrete Event Simulation, Optimization, Planning, Fashion.

ABSTRACT

This paper presents a Simulation Optimization (SO), decision-support tool developed for metal accessories' suppliers in the fashion industry. The tool is composed of a discrete-event simulator and a multi-objective, integer linear-optimization scheduler, based on a commercial spreadsheet and an open-source solver, linked together through an import-export routine. The tool can be used to enable suppliers to compare scheduling algorithms in order to optimize their performances in terms of customers' due dates compliance and cost and processing time reduction. The analyzed scenario takes into account variable and uncertain production plans, represented by the aggregation of orders received from different brands. The model has been applied to a real company, where costs, delay, and advances are considered in order to define the Objective Function (OF), whilst rush orders are introduced to simulate stochastics events.

INTRODUCTION

As widely recognized in the literature, challenges in the fashion–product industry do not only deal with creativity and styles, but also supply chain (SC) management. One of the main criticalities of the fashion SC is the high uncertainty of the demand (Ait-Alla et al. 2014; d'Avolio et al. 2015; Hu et al. 2013). In recent years, the fashion-product lifecycle has become ever shorter, and the number of fashion seasons has increased. As a consequence, the fashion industry is centered on the ability to react quickly to changes in customers' desires, increasing the need to compress time-to-market. On the other hand, fashion customers ask for a higher service level, mainly in terms of quality. All of these aspects have pushed companies to increase their own pressure on the upstream SC actors. This evidence reflects the fact that these results cannot be obtained through operations at the single-company level, but rather throughout the entire SC, because outstanding quality and delays of a final product are directly linked to other components (Caniato et al. 2013). Within this SC, an important role is played by metal accessories suppliers, often SMEs, located close to the fashion brands. The optimization of the performance of various production units, identifying the optimal quantities and places to allocate items production, which affect the problems of multi-plant production planning, are widely discussed in the literature and several surveys can be found (e.g. Fujimoto 2015; Jeon and Kim 2016;). In this paper, a first step of the problem concerns the local optimization of scheduling performance in the domain of Simulation and Optimization Integration. Even if this topic has been debated in the literature and applied to various domains, from manufacturing to healthcare (Jung et al. 2004; Sowle et al. 2014), no contributions have been found related to the fashion industry, making the present work innovative as first implementation of a model combining optimization and simulation tools within this industry.

FASHION SCHEDULING OPTIMIZATION REVIEW

In the scientific literature, several different approaches to the definition of scheduling formulation can be found. Published reviewing papers on scheduling (Maravelias 2012; Méndez et al. 2006; Mula et al. 2010; Phanden et al. 2011; Ribas et al. 2010) study various problems, moving from single to parallel machines, job, or flow shop, and considering different levels of data aggregation (i.e., strategical, tactical, and operative). Focusing on contributions related to the fashion industry in the literature, there are many papers considering finite or infinite capacity, where finite capacity can be considered in terms of hours (Rahmani et al. 2013) or units per resource (Ait-Alla et al. 2014; Guo et al. 2015; Wong et al. 2014; Rahmani et al. 2013). Betrand and Van Onijen (2008) present various multi-objective OFs, both liner and not linear, including costs, time, and plant-performance optimization. At the same time, several contributions can be identified where simulation techniques are used in order to optimize production in the fashion industry, i.e. Al-Zubaidi et al. (2004), Cagliano et al. (2011) and Jung et al. (2004). Nevertheless, no contribution can be found where a combined simulation-optimization model is described and applied that considers the specific characteristic of this industry.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

PROBLEM STATEMENT

One of the consequence of the high pressure applied by brand owners on suppliers of metal accessories is their motivation to adopt optimization of process planning and scheduling. On the other hand, fashion brands, due to the variability of the demand, are at the same time reducing the suppliers' orders visibility, consequently increasing the frequency of re-scheduling their production plan. This dichotomy has caused fashion suppliers to ask for inexpensive and easy-to-use tools capable of determining optimal, or sub-optimal, scheduling techniques to be adopted for the production plan they have to accomplish. An important aspect of this optimization model deals with the fact that metal fashion suppliers have to optimize a local production plan that is a matrix of several fashion-brands production plants, each of them developed according to different criteria, and in which seasonal products and carryover are mixed together and can quickly and significantly change from one day to another. Another important boundary of the problem is the use of sub-suppliers with unknown production capacity. These sub-suppliers are used to perform mechanical work, except for the machine shop, which is internal and considered to possess finite capacity in the model.

Model Description

Starting from the boundaries described above, the simulation-optimization model proposed in the paper has the objective to define the optimal supplier- production plan, according to a set of KPI mixed with various weights, based on the companies Critical Success Factors (CSFs) and on the uncertainty due to internal factors (machine failures, reworks, employees unavailability) and external stochastics events (rush orders). The importance of including rush orders is due to the uncertainty and high variability of the brands' production orders. Unexpected orders can represent a high proportion of the value of the production, up to the 20% of the total capacity. The variables considered by the model are derived from time measurements (tardiness, lateness), costs (machine costs, labor costs, overtime costs), and production (wip, leadtime).

The model describes, in a stochastic way, the behavior of metal suppliers' production cycle and that of one of their sub-suppliers, while the scheduler, according to an Objective Function (OF) defined as a mix of parameters (costs, delays and minimization of processing time) chosen by the company, defines the optimal scheduling solution for a single phase of the production process.

In the specific, the solver model has been developed with the following function:
OF: $Min\{\sum_{i \in II} (cw_i * C_i + dw_i * D_i + aw_i * A_i + ptw_i * PT_i\}$ where cw_i, dw_i, aw_i, ptw_i, and weights of the various objectives, according to Guo et al. (2008), and C_i, D_i, A_i, and PT_i are respectively $\sum_{i \in II}$ *Costs* (C_i), $\sum_{i \in II}$ *Delays* (D_i), $\sum_{i \in II}$ *Advances* (A_i), and $\sum_{i \in II}$ *Processing Time* (PT_i).

More information on the model and how the objectives are evaluated can be found in Fani et al. (2016).

Model Architecture

The model is composed of a Java discrete-event simulator, AnyLogic® (www.anylogic.com), and an open solver optimization tool, OpenSolver (www.opensolver.org). The version of AnyLogic that has been used is the 7.3, and the version of the Solver is the 2.8.5. The solver has been used integrated on Microsoft Excel®. The architecture of the model, assuming a comparison between three different OFs weights distribution, is represented in Fig. 1.

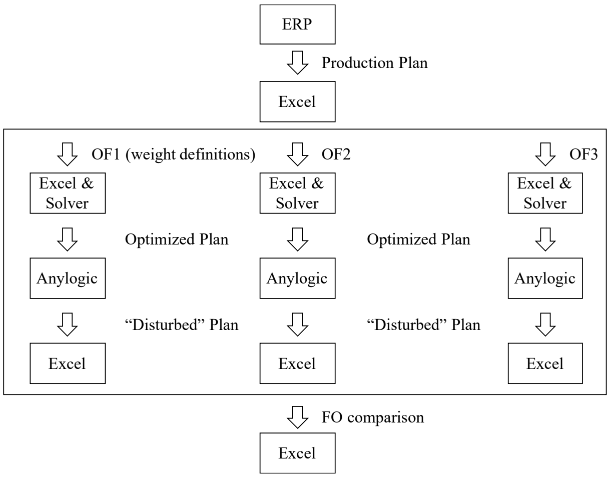


Figure 1: Model Architecture

The procedure to define the optimal scheduling solution, according to a specific production plan of the company, the stochastics events estimated for that company and the defined KPIs, is described in Fig. 2. This procedure has to be done for each of the OFs that has to be compared.

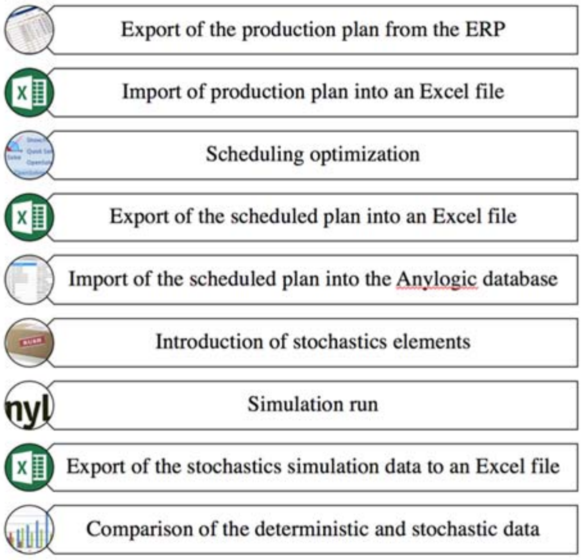

Export of the production plan from the ERP
Import of production plan into an Excel file
Scheduling optimization
Export of the scheduled plan into an Excel file
Import of the scheduled plan into the Anylogic database
Introduction of stochastics elements
Simulation run
Export of the stochastics simulation data to an Excel file
Comparison of the deterministic and stochastic data

Fig. 2 Scheduling simulation optimization procedure

CASE STUDY

Implementation of the model

The model described in the previous section has been validated using a data set acquired from the real experience of a supplier of fashion metal accessories, where the production starts with the realization of a semifinished item and proceeds through shaving removal, followed by some mechanical operations (e.g., vibration, vibratory finishing, drilling, cutting). Then, items have to be covered by one or more precious metals, such as gold, palladium, and ruthenium, through an electroplating process. In the final step, the items have to pass the quality control, be packaged, and be delivered in order to be applied on the final product.

In the case study, the phase characterized by finite capacity is the machine shop, where the shape-removal is done by Computer Numerical Control (CNC) machines. All the steps after this phase have been considered working at infinite capacity and have been modeled with a generic processor. The model has been implemented considering a production plan of 40 days to be processed by a production plant operating 24 hours per day, 7 days per week.

The optimization tool has been parametrized with an OF that has been defined combining costs, delays, and advances, whilst processing time has not been considered in this first implementation. According to this, the two combinations of weights (i.e. one for each OF) chosen for each one of the three parameters have been decided by the analysed company, guaranteeing coherence with its specific CSFs.

The aim of this work is twofold: on the one hand, the developed simulation model has been validated comparing the output with that one of the optimization model; on the other hand, the deterministic optimized scheduled plan has been compared with the one simulated taken into account stochastic elements for evaluating their impact on KPIs (deterministic vs. stochastic). The stochastic elements that have been considered are rush orders, created according to a uniform distribution that represents the unexpected orders received by the company in the last year.

The structure of the production plan exported from the company ERP that represents the input of the optimization model is summarized in Table 1.

Table 1: Production plan exported from the ERP

Code	Description
db_id	Order ID
db_key	Item code
db_order	Order code
db_order_line	Order line position
db_machine	CNC type (turn, mill)
db_qt_prod	Number of items to be produced
db_delivery_date	Item due date

The production order considered in the case study includes 11 articles with different lot sizing, (i.e from 35 up to 10,000 items for item code) and customer delivery dates between 15th and 23th of February 2017. Data related to the production plan exported from the ERP have been integrated with that ones that characterize the items' working cycle, collecting data from historical information. This prarameters are the processing time per item on each companies' CNC machines, the lead time per item for the subsequent production steps and the processing costs per item for each machines.

Processing times can be different for different item codes (i.e longer moving from simpler to more complex) but even for the same one (i.e. working the same article through a newer machine require a shorter processing time than an old one). Processing times for the scheduled items are between 10 and 135 seconds, while costs are equally evaluated for the CNC machines.

Finally, processing time and cost per machine can be recorded as null, because not every machine can be used for producing a specific item code.

All the described information represent the input for the optimization model run using the OpenSolver tool on Excel®.

The OFs of the two analyzed scenarios are the following:

a) OF_1: $Min\{\sum_{i \in II} (cw_i * C_i + dw_i * D_i + aw_i * A_i)\}$
where $cw_i = 1$, $dw_i = 1$, $aw_i = 1$;

b) OF_2: $Min\{\sum_{i \in II} (cw_i * C_i + dw_i * D_i + aw_i * A_i)\}$
where $cw_i = 1$, $dw_i = 1$, $aw_i = 0$.

The reason these two scenarios have been chosen is related to the management request to develop a tool able to show what are different impact on the scheduling performance if advances are included (i.e OF_1) or not (i.e OF_2), enabling users to analyze the output evidences value of delays and advances in terms absolute terms or related to a specific working machines or subset of items. The amount of delays and advances are calculated as sum of all the quantity per item respectively not produced or produced in advance if compared to the requested delivery date, considering both final and intermediate process steps. This comparison has been made per single day during the analyzed time slot.

On the other hand, costs value has been calculated multiplying assigned quantity per item with the unitary working cost per machine mapped as input data on the Excel® (i.e. exported from the company's ERP) and as agents' parameter on the simulator. According to the management request, at the first stage of model implementation costs are considered equal for every item processed by every machine. Moreover, no difference as been made according to the work schedules actually used (i.e. 24 hours per day, 7 days per week), that push the company to not considering overtime and related extra-costs.

The format of the optimal solution is listed in Table 2.

Table 2: Scheduled optimization plan output

Code	Description
db_key	Item code
n-t1	Items assigned on the day n to the turn 1
n-t2	Items assigned on the day n to the turn 2
n-t3	Items assigned on the day n to the turn 3
n-m1	Items assigned on the day n to the mill 1
n-m2	Items assigned on the day n to the mill 2

In the case study, the optimal scheduled production plan shows that, according to both the OFs and the optimization model constraints (i.e. CNC machines' capacity and demand fulfillment), some items will be produced in advance and some others with a delay respect to the delivery dates specified in the production plan (see the "Results" section).

Considering the schedulation time slot, the optimization model results cover 15 days in the first scenario and 24 days in the other.

Once the optimization plan has been recorded, the assigned items have been imported into the AnyLogic® simulator using an sql script and represent the input for the simulation. In fact, the agents (i.e. the list of items included in the company's production plan) are generated according to the parameters within the Excel® file containing the optimization model results.

Moreover, the output of the optimization model in terms of end processing date, processed quantity and assigned machine per single item has defined the rules for developing the simulation model (i.e. the way the agents are generated and the path that they have to follow along the process flow).

Deeply analyzing the simulated model, it is composed by two resources type (i.e. turns and mills) with several machines for each one (in particular, three turns and two mills, as the company's layout already mapped on the optimization model on the Excel®). According to the model overview, each machine processes only the items that have been assigned to itself by the optimization model, considering as processing time the one that is reported on the Excel® file and recording it as an agent parameter. According to the production plan parameters, the total number of agents generated are 14,482.

As mentioned before, all the process activities that follow turning and milling ones are modeled as a unique processing block, called "postpocessing", that covers a specific processing time for each item, extracted from the Excel® file as agent's parameter (in the same way of turning and milling processing times).

The simulation model described until now represents the same deterministic scenario of that one modeled on Excel® and has been successfully used to validate the optimization model in terms of resulting key performance indicators (KPIs), such as delays and advances days for each item.

In the case study, stochastic events have been included considering a deviation from the deterministic plan due to the presence of rush orders. These orders are generated with a uniform statistical distribution U(40,50) considering an arrival rate generated according to a normal distribution with average and variance equals to 1. The percentage of rush orders generated by the simulator during the run is almost the 10% of the total production quantity, according to the historical data collected in the analyzed fashion company.

The modeled rush orders have been generated as a set of items included into the original production plan, inheriting production cycle, processing and post-processing times. Moreover, rush orders have priority over the scheduled items that, on the other hand, move on the simulated production process following a FIFO queue.

Two simulation campains have been conducted: the first one generates the input items from the source block of the AnyLogic® simulator according to the optimized plan that minimize the OF_1, while the second one processes items following the scheduled production referred to the OF_2.

The simulation time slot covers four months, in order to complete the scheduled orders considering the presence of priority rush orders.

In order to compare the different simulation campains with the scheduled deterministic optimizated production plan, the key performances indicators (KPIs) reported in Table 3 have been defined.

Table 3: Simulation model's KPIs

KPI	Formula
gap delivery date1	deathdate – customerRequestedDate
gap delivery date2	customerAssignedDate – customerRequestedDate
gap delivery date3	deathdate – customerAssignedDate
gap processing delivery date1	stopDate – requestedDate
gap processing delivery date2	assignedDate – requestedDate
gap processing delivery date3	stopDate – assignedDate

The "gap delivery date 1" shows the gap between the simulated end processing date for the final product and the one requested by the customer. In other words, it shows the lateness, as calculated by the simulation model. The "gap delivery date 2" is the lateness but referred to the optimization model's output (i.e. Excel® file). Finally, the "gap delivery date 3" compares the simulator and the optimization models' outputs, again in terms of delays or advances per item related to the final product production. This KPI represents the deviation between the optimized lateness and the one evaluated by the simulation.

KPIs "gap processing delivery date 1", "gap processing delivery date 2" and "gap processing delivery date 3" are defined as the previous ones but refers to the semifinished products (i.e. outputs of turning and milling machines) instead of final ones (i.e. outputs of post-processing block).

Results

As first result of the present work, the simulation model has been successfully validated comparing the resulting outputs to that ones calculated through the optimization model on the Excel®. In particular, for each run simulation campaign the gap, in terms of days, between real and requested delivery date per each item calculated through the two models (i.e. Opensolver and AnyLogic®) has been compared, considering both final and intermediate steps. This comparison results in a punctual alignment between the two models' outputs and it has been evaluated considering both the OFs (i.e. OF_1 and OF_2).

In particular, the first scenario, that takes into account all the three parameters (costs, delays and advances), results in an OF_1 value equals 88,019, composed by $\sum_{i \in II} (cw_i * C_i) = 14{,}482$, $\sum_{i \in II} (dw_i * D_i) = 72{,}056$ and $\sum_{i \in II} (aw_i * A_i) = 1{,}481$. The second one, that differs from the prievious scenario for not considering advances as one of the OF's parameters, results in an OF_2 value equals to 19,296, composed by $\sum_{i \in II} (cw_i * C_i) = 14{,}482$ and $\sum_{i \in II} (dw_i * D_i) = 4{,}814$; advances equals to $\sum_{i \in II} (aw_i * A_i) = 110{,}805$.

As expected, the value of delays is lower for OF_2 in comparison with OF_1. In fact, the constraints in terms of CNC machines' capacity and demand fulfillment force to not always respect the requested delivery date in both scenarios, but while in the first one delays and advances are equally weighted, in the other one just delays are taken into account, pushing the optimized schedulation of the items towards anticipate their production respect to the delivery date. In the same way, the advances related to the second scenario largely overcome that ones of the first one because their amount is not included into the OF.

The second results of this work is related to the comparison between the schedulation plan simulated considering unexpected orders to be priority processed and the optimal solution that considers just pre-scheduled orders as input.

This gap analysis has been conducted considering both the OFs, and the compared KPIs are shown for OF_1 and OF_2 respectively in Table 4 and 5.

KPIs related to the output of the models, in terms of number of worked items, refers both to the scheduled and rush orders in the analysis on the simulation model, while the others related to delays and advances are related just the scheduled orders. The reason why we have chosen to consider only these orders is that the aim is to assess the impact of rush orders on the previous schedulation, modeled on the Excel®. Moreover, due to the fact that rush orders are priority by definition, they report null delays and advances.

Table 4: Comparison between models' KPIs (OF_1)

KPI	Optimized plan	Stochastic simulation	Δ %
Output quantity per turn 1 $^{(a)}$	10,020	12,185	21.61%
Output quantity per turn 2 $^{(a)}$	44	485	1002.27%
Output quantity per mill 1 $^{(a)}$	1,060	2,639	148.96%
Output quantity per mill 2 $^{(a)}$	3,358	5,383	60.30%
Delays per turn 1 $^{(a)}$	0	387,285	---%
Delays post-processing $^{(a)}$	72,056	491,236	581.74%
Advances per turn 1 $^{(a)}$	16,237	16,237	0%
Advances per turn 2 $^{(a)}$	88	88	0%
Advances per mill 1 $^{(a)}$	2,119	2,119	0%
Advances per mill 2 $^{(a)}$	6,715	6,715	0%
Advances post-processing $^{(a)}$	1,481	37,099	2405.00%
Max delay per turn 1 $^{(b)}$	0	367	---%
Average delay per turn 1 $^{(b)}$	0	178.88	---%
Max delay post-processing $^{(b)}$	21	141	571.43%
Average delay post-processing $^{(b)}$	6.46	76.74	1087.93%
Max advance per turn 1 $^{(b)}$	3	3	0%
Average advance per turn 1 $^{(b)}$	1.62	2	23.46%
Max advance per turn 2 $^{(b)}$	2	2	0%
Average advance per turn 2 $^{(b)}$	2	2	0%
Max advance per mill 1 $^{(b)}$	2	2	0%
Average advance per mill 1 $^{(b)}$	2	2	0%
Max advance per mill 2 $^{(b)}$	2	2	0%
Average advance per mill 2 $^{(b)}$	2	2	0%
Max advance post-processing $^{(b)}$	7	7	0%
Average advance post-processing $^{(b)}$	3.16	4.59	45.25%

* Units of measurement: (a) number of items; (b) days.

** Output quantity per turn 3 $^{(a)}$; Delays per turn 2 $^{(a)}$; Delays per turn 3 $^{(a)}$; Delays per mill 1 $^{(a)}$; Delays per mill 2 $^{(a)}$; Advances per turn 3 $^{(a)}$; Average and maximum delay per turn 2, turn 3, mill 1 and mill 2 $^{(b)}$; Max advance per turn 3 $^{(b)}$; Average advance per turn 3 $^{(b)}$;

Table 5: Comparison between models' KPIs (OF_2)

KPI	Optimized plan	Stochastic simulation	Δ %
Output quantity per turn 1 $^{(a)}$	10,000	11,726	17.26%
Output quantity per turn 2 $^{(a)}$	44	444	909.09%
Output quantity per turn 3 $^{(a)}$	20	433	2065.00%
Output quantity per mill 1 $^{(a)}$	3,067	5,129	67.23%
Output quantity per mill 2 $^{(a)}$	1,351	3,042	125.17%
Delays per turn 2 $^{(a)}$	132	132	0%
Delays per turn 3 $^{(a)}$	340	340	0%
Delays per mill 1 $^{(a)}$	3,770	3,770	0%

Delays per mill 2 (a)	572	572	0%
Delays post-processing (a)	4,814	699,982	14,440.55%
Advances per turn 1 (a)	81,100	81,100	0%
Advances per mill 1 (a)	19,330	19,330	0%
Advances per mill 2 (a)	10,375	10,375	0%
Advances post-processing (a)	110,805	81,875	-26.11%
Max delay per turn 2 (b)	3	3	0%
Average delay per turn 2 (b)	3	3	0%
Max delay per turn 3 (b)	17	17	0%
Average delay per turn 3 (b)	17	17	0%
Max delay per mill 1 (b)	9	9	0%
Average delay per mill 1 (b)	8.38	8.38	0%
Max delay per mill 2 (b)	13	13	0%
Average delay per mill 2 (b)	13	13	0%
Max delay post-processing (b)	17	125	635.29%
Average delay post-processing (b)	8.63	109	1,163.04%
Max advance per turn 1 (b)	10	10	0%
Average advance per turn 1 (b)	8.11	8.11	0%
Max advance per mill 1 (b)	9	9	0%
Average advance per mill 1 (b)	7.39	7.39	0%
Max advance per mill 2 (b)	9	9	0%
Average advance per mill 2 (b)	7.94	7.94	0%
Max advance post-processing (b)	10	13	30.00%
Average advance post-processing (b)	7.96	10.18	27.89%

* Units of measurement: (a) number of items; (b) days.

** Delays per turn 1 (a); Advances per turn 2 (a); Advances per turn 3 (a); Max delay per turn 1 (b); Average delay per turn 1 (b); Max advance per turn 2 (b); Average advance per turn 2 (b); Max advance per turn 3 (b) values are zero both for optimized than stochastic simulation.

As shown in Table 4, the number of processed items grown from 14,482 to 20,692 if rush orders are considered (+42.88%). Delays for items worked by the turn 1, that are null for the scheduled plan, grown up to 387, and the same KPI related to the post-processing increases in a more than proportional way in comparison to the total number of items (rush orders included). This is due to the fact that, in the simulation run, most of the rush orders have been processed by the first machine.

At the same time, as shown in Table 5, the number of items to be processed, considering rush orders, increased by 43.45% (i.e. from 14,482 to 20,774). Referring to the OF_2, a relevant gap in terms of delays on the delivery date considering rush orders can be registered for the post-processing phase, aligned to the fact that the production flow of all the processed items converges on the same working station, being more stressed by the extra-work. On the other hand, KPIs related to the single CNCs do not worsen their value. This is justified but the fact that OF_2 does not consider the advance as a damage. Consequently, the optimized plan is anticipated in comparison to the customer requested date, and a production of unexpected items can be done without having to change the planned scheduling. From an industrial point of view, it is important to remark that OF_2 could not be feasible at all as production scheduling strategy. In fact, fashion orders, in terms of quantity and delivery date, are usually confirmed quite close to last date available for processing them on-time, making advances in production risky.

It is important to highlight that the negative effect of rush orders is amplified in most industries, included the fashion one, because of the fact that orders can be delivered to the client (i.e. the brand owner) only when the lot is completed. Analyzing the OF_2, it is possible to see that the effect that rush orders have in terms of delay quite overcomes the increasing value of products in input in the simulated model (see Table 5), and even worse is the scenario considering the OF_1, when the delay value arrives up to 141 days (see Table 4). In fact, for an incremented quantity of items to be produced around the 45%, the maximum value of delay registered in the post-processing is up to 2,405% (i.e. 125 days) for OF_1 and up to 635.29% for OF_2.

CONCLUSION

The paper describes an SO decision-support tool developed for the fashion metal accessories' suppliers. The tool has been developed using an open-source solver (OpenSolver) and a commercial simulator (AnyLogic®), in order to be usable by these companies, and validated using a real production plan. The production plan of the company has been optimized using the solver with different OFs. The discrete-event simulator is used to validate and compare various scheduled production plans produced by the optimization tool, introducing internal and external stochastics elements.

First of all, the simulation model has been successfully validated comparing the resulting outputs (i.e. end processing dates and processed quantities, considering both final and intermediate steps) to that ones calculated through the optimization model on the Excel®.

Moreover, the impact of unexpected orders to be processed on the KPIs have been analysed using the simulation model, allowing to measure the gap between schedulation outputs considering just the minimization of costs and delays (OF_2) or also advances (OF_1) and results have been reported.

Future development of this work deals with firms' ERP integration in order to automatize the process of import of the production plan and the visualization of the scheduling results. Moreover, an in-depth analysis of the post-processing activities, modeled on the simulator as a unique block, will be conducted, and other business objectives, such as reducing processing time and leveling machines utilization, will be included in the OF.

REFERENCES

Ait-Alla, A., Teucke, M., Lütjen, M., Beheshti-Kashi, S. and Karimi, H. R. (2014). Robust production planning in fashion apparel industry under demand uncertainty via conditional value at risk. Mathematical Problems in Engineering, 2014(2014), 1–10.

Al-Zubaidi, H. and Tyler, D. (2004). A simulation model of quick response replenishment of seasonal clothing. International Journal of Retail & Distribution Management, 32(6), 320–327.

Bertrand, J.W.M., Van Ooijen, H.P.G. (2008). Optimal work order release for make-to-order job shops with customer order lead-time costs, tardiness costs and work-in-process costs. International Journal of Production Economics, 116(2), 233–241.

Cagliano C., A., DeMarco, A., Rafele, C. and Volpe, S. (2011). Using system dynamics in warehouse management: A fast-fashion case study. Journal of Manufacturing Technology Management, 22(2), 171–188.

Caniato, F., Caridi, M. and Moretto, A. (2013). Dynamic capabilities for fashion-luxury supply chain innovation. International Journal of Retail & Distribution Management, 41(11/12), 940–960.

d'Avolio, E., Bandinelli, R., Rinaldi, R. (2015). Improving new product development in the fashion industry through product lifecycle management: A descriptive analysis. International Journal of Fashion Design, Technology and Education, 8(2), 108–121.

Fani, V., Bandinelli, R., Rinaldi, R. (2016). Toward a scheduling model for the metal accessories' suppliers for the fashion industry. Proceedings of the Summer School Francesco Turco, 13-15-September-2016, pp. 166-170.

Fujimoto, R. (2015). Parallel and distributed simulation. Proceedings of the 2015 Winter Simulation Conference, pp. 45-59. IEEE Press.

Guo, Z. X., Ngai, E.W.T., Can, Y., Xuedong, L. (2015). An RFID-based intelligent decision support system architecture for production monitoring and scheduling in a distributed manufacturing environment. International Journal of Production Economics, 159, 16–28.

Guo, Z.X., Wong, W.K., Leung, S.Y.S., Fan, J.T., Chan, S.F. (2008). A genetic-algorithm-based optimization model for solving the flexible assembly line balancing problem with work sharing and workstation revisiting. IEEE Transactions on Systems, Man and Cybernetics Part C - Applications and Reviews, 38(2), 218–228.

Hu, Z.-H., Zhao, Y., Choi, T.-M. (2013). Vehicle routing problem for fashion supply chain with cross-docking. Mathematical Problems in Engineering, 2013(2013), 1–10.

Jeon, S. M. and Kim, G. (2016). A survey of simulation modeling techniques in production planning and control (PPC). Production Planning & Control, 27(5), 360–377.

Jung J. Y., Blau G., Pekny J. F., Reklaitis G. V., Eversdyk D. (2004). A simulation based optimization approach to supply chain management under demand uncertainty. Computers & Chemical Engineering, 28(10), 2087–2106.

Maravelias, C. T. (2012). General framework and modeling approach classification for chemical production scheduling. AIChE Journal, 58(6), 1812–1828.

May, G., Stahl, B., Taisch, M., Prabhu, V. (2015). Multi-objective genetic algorithm for energy-efficient job shop scheduling. International Journal of Production Research, 2015, 1–19.

Méndez, C. A., Cerdá, J., Grossmann, I. E., Harjunkoski, I., & Fahl, M. (2006). State-of-the-art review of optimization methods for short-term scheduling of batch processes. Computers & Chemical Engineering, 30(2006), 913–946.

Mula, J., Peidro, D., Díaz-Madroñero, M., Vicens, E. (2010). Mathematical programming models for supply chain production and transport planning. European Journal of Operational Research, 240(3), 377–390.

Phanden, R. K., Jain, A., Verma, R. (2011). Integration of process planning and scheduling: a state-of-the-art review. International Journal of Computer Integrated Manufacturing, 24(6), 517–534.

Pinedo, M., Chao, X. (1999). Operations scheduling with applications in manufacturing and services. Boston: Irwin/McGraw-Hill, ISBN 0-07-289779-1.

Rahmani, D., Ramezanian, R., Fattahi, P., Heydari, M. (2013). A robust optimization model for multi-product two-stage capacitated production planning under uncertainty. Applied Mathematical Modelling, 37(2013), 8957–8971.

Ribas, I., Leisten, R., Framinan, J. M. (2010). Review and classification of hybrid flow shop scheduling problems from a production system and a solutions procedure perspective. Computers & Operations Research, 37(8), 1439–1454.

Rose, M. D., Shier, R. D. (2007). Cut scheduling in the apparel industry. Computers & Operations Research 34(11), 3209–3228.

Sowle, T., Gardini, N., Vazquez, F. V. A., Pérez, E., Jimenez, J. A., De Pagter, L. (2014). A simulation-IP based tool for patient admission services in a multi-specialty outpatient clinic. Proceedings of the 2014 Winter Simulation Conference, 1186–1197. IEEE Press.

Wong, W. K., Guo, Z. X., Leung, S. Y. S. (2014). Intelligent multi-objective decision-making model with RFID technology for production planning. International Journal of Production Economics, 147(2014), 647–658.

Wu, T., Shi, L., Geunes, J., Akartunali, K. (2011). An optimization framework for solving capacitated multi-level lot-sizing problems with backlogging. European Journal of Operational Research, 214(2), 428–441.

Wu, Z., Liu, X., Ni, Z., Yuan, D., Yang, Y. (2013). A market-oriented hierarchical scheduling strategy in cloud workflow systems. The Journal of Supercomputing, 63(1), 256–293.

AUTHOR BIOGRAPHIES

VIRGINIA FANI is a PhD student at the Industrial Engineering Department of the University of Florence. The issues she is dealing with are related to production processes optimization along the supply chain, with particular focus on the peculiarities that the fashion companies have to face.

ROMEO BANDINELLI is a reseacher at the University of Florence, Department of Industrial Engineering. He is member of the Observatory GE.CO. of Politecnico di Milano, program chair of the IT4Fashion conference and member of the IFIP 5.1 "Global Product development for the whole lifecycle".

RINALDO RINALDI is associate professor at the School of Engineering of the University of Florence. He teaches Supply Chain Management and Operations Management. His areas of research deal with the design of logistics systems, programming and production control, optimization of production processes and supply chain, RFId for logistic optimization.

AN OPTIMIZATION OF SPRAY COATING PROCESS TO MINIMIZE COATING MATERIAL CONSUMPTION

Nitchakan Somboonwiwat
Suksan Prombanpong
Department of Production Engineering,
King Mongkut's University of Technology Thonburi (KMUTT), Bangkok 10140 Thailand
E-mail: nonpang24@hotmail.com, suksan.pro@gmail.com

KEYWORDS
Material consumption, Optimization, Spray coating process

ABSTRACT

In cookware industry, interior spray coating process is an important process to protect the cookware product from the corrosion. In this process, the coating material "TEFLON" is used to spray to cover all the part's surface. If a pot or pan is not too deep, one spray gun is enough to spray to cover all the interior surface. However, for the high side-wall pot, two spray guns must be used to spray at two different areas. The first area is at the bottom and its corner. The second area is around the top rim of the pot or pan. Note that the spray pattern is of fan-type. Thus, the sprayed area will be covered by the mentioned two spray guns. Consequently, the large amount of coating material is consumed to meet the dry film thickness (DFT) requirement. In this paper, the optimization of spray coating process is studied aiming to minimize the material usage. The experimental design technique is applied to determine the optimal spray coating parameters. The parameters used in this study include angle, spray time, fan pattern, and air pressure. The relationship models of coating material volume and DFT are presented in this research. The optimal parameters of the two spray guns are presented.

INTRODUCTION

The cookware production comprises many processes starting from blanking to packing. The coating process is an important step before assembly. It makes the products strong, long lasting, beautiful and protects the surfaces. In the coating process, the spray gun is normally used to spray the coating material to a target surface. Basically, the coating process consist of triple-layer coatings: primer, middle, and top layer. The primer is the first layer of coating that is applied to the substrate for interfacing between the surface of the part and the middle layer. The main purpose of the primer layer is a preparation of the coating surface to ensure smoothness and good adherent of coating material to the part surface. The middle layer is the layer on the primer layer. The coating material of the middle layer is an actual coating material whose function is to increase corrosive resistant of a part. Finally, the top layer will be applied make a part look shinny and increase the efficiency of the coating material in the middle layer. Considering all three layers of the coating process, the middle layer is the most important which affects the durability of the cookware product. As a consequence, it increases the service life of products, while the primer is the substrate of coating and the top layer is the decoration purpose. Therefore, this research will focus on the middle layer coating.

In the spray coating process, a coating material waste due to an overspray and bounch of the coating material are the main issue. Thus, the objective of this research is to determine optimal spray parameters to minimize the coating material consumption due to DFT specification.
(Winnicki et al. 2014 and From et al. 2011) study the optimal parameters of spray coating process to meet DFT but not the coating material consumption. There are few researchers studied the optimal parameters affecting both the coating material consumption and DFT research (Song et al. 2008 and S. Hong et al. 2014). (Luangkularb and Prombanpong 2014) present the optimal parameters concerning both of the coating material consumption and DFT for the single spray gun. However, this paper demonstrates a determination of optimal parameters considering both of material consumption and DFT for the two spray guns. The 26 cm in diameter of a pan is the specimen in this study. The spraying area is divided two positions i.e. bottom and corner, and top rim of a pan. Thus, the first gun sprays to cover the bottom area of the part whereas the second spray gun aims to cover the side wall and rim of the pan.

METHODOLOGY

The experimental design technique is applied to obtain the optimal spray parameters to minimize the coating material consumption and to attain DFT. The concerned parameters include gun angle, spray time, fan pattern and air pressure. The two levels used for the experiment is presented in Table 1. Thus, at each spray gun, the

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

experiment is designed with regard to the 2^4 factorial design with two replications. Thus, a total of 16 experiments will be performed for each spray gun. The material consumption and DFT are measured as the response and the gun angle, spray time, fan pattern and air pressure are recorded as the independent variables. The data obtained from the experiment will be then analyzed using the analysis of variance (ANOVA) technique to determine the effect of these four independent variables on material consumption and DFT. The required DFT is in a range of 7.5-12.5 µm. The predictive relationship model for these two responses is then constructed to find the optimal parameters for the spray coating process.

Table 1: Data Used in the Experiment

Factor	First spray gun		Second spray gun	
	min	max	min	Max
Gun Angle (degree)	60	70	35	40
Spray time (sec)	1.2	1.4	1.2	1.4
Fan pattern (rev.)	315	360	315	360
Air pressure (bar)	2.5	3.0	2.5	3.0

RESULTS AND DISCUSSION

The result obtained from the experiment will be analyzed using MINITAB software. The normality, constant variance and randomization tests are performed. The analysis of variance (ANOVA) is subsequently conducted to determine the effect of variables on the coating material consumption (MC) and average dry film thickness (DFT). In addition, the process optimization is performed to minimize material consumption and also yield the DFT ranging in the specification. The results of statistical analysis of first spray gun and second one are as follows.

Coating Process of First Spray Gun

- **Statistical Analysis**

The model adequacy checking of the material consumption data shows that the p-value for the normality probability test, equal variance and randomization test equal to 0.722 0.487 and 0.153 respectively. For the DFT data, the p-value of the normality probability test, equal variance, and randomization test are equal to 0.437, 0.637 and 0.249 respectively (Table 2). The result indicates that the MC and DFT data are in the normal distribution and there is no deviation in a variance of each test. ANOVA is conducted and the results are obtained for MC and DFT responses as shown in Table 3 and 4. The p-value which is less than 0.05 is used as the criterion to determine the significance of the response.

Table 2: P-value of Model Adequacy

Model adequacy checking	P-Value	
	Material consumption	Dry film thickness
Normal distribution	0.722	0.437
Equal variance	0.487	0.637
Randomization	0.153	0.249

Table 3: Result of ANOVA Material Consumption for First Spray Gun

Term	P-Value
Gun angle	0.002
Spray time	0.000
Fan pattern	0.000
Air pressure	0.000
Gun angle*Fan pattern	0.004
Spray time* Fan pattern	0.020
Spray time*Air pressure	0.000
Fan pattern*Air pressure	0.003
S = 0.547875 PRESS = 12.8071	
R-Sq = 86.84% R-Sq(pred) = 76.61% R-Sq(adj) = 83.00%	

Table 4: Result of ANOVA Dry Film Thickness for First Spray Gun

Term	P-Value
Gun angle	0.033
Spray time	0.000
Fan pattern	0.001
Air pressure	0.000
Gun angle*Fan pattern	0.007
Spray time*Air pressure	0.023
Fan pattern*Air pressure	0.000
S = 0.547875 PRESS = 12.8071	
R-Sq = 86.84% R-Sq(pred) = 76.61% R-Sq(adj) = 83.00%	

- **Effect of Spray Condition Analysis**

The effects of gun angle, spray time, fan pattern and air pressure on the material consumption and dry film thickness are shown in Figure 1 and 2. The results indicate that the increment of gun angle, fan pattern and air pressure will decrease the material consumption and dry film thickness. However, the decrement in the spray time will decrease the material consumption and dry film thickness.

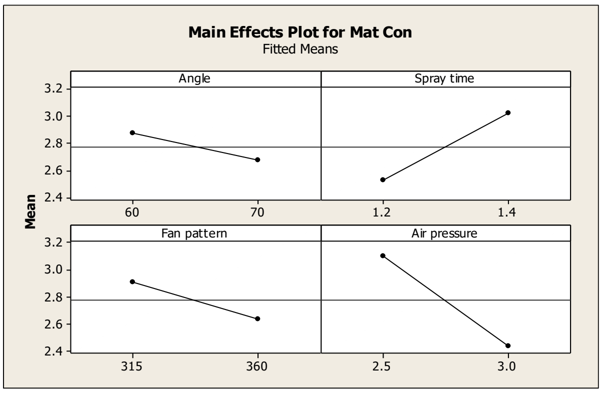


Figure 1: Main Effect Plot of Variable on Average Material Consumption for First Spray Gun

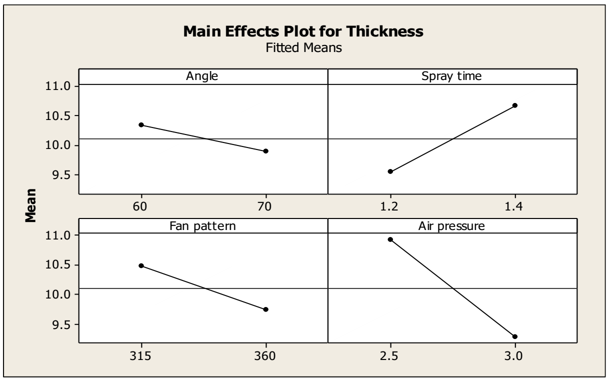


Figure 2: Main Effect Plots of Variables on Average Dry Film Thickness for First Spray Gun

- **Process Optimization and Modeling**

The regression model is constructed to present the relationship between response, coating material consumption and DFT, with the variables, angle, spray, fan pattern and air pressure as shown in (1) and (2). The optimal condition is the gun angle at 60 degree, spray time at 1.2 sec., fan pattern at 315 revolution (rev.) and air pressure at 2.7 bar. The material consumption and DFT equal to 2.8 g. and 10.45 µm which is in the specification. Table 5 shows the optimal results.

Parameters (Indepentdent Variables):

X_1 angle (degree)
X_2 spray time (sec)
X_3 fan pattern (rev)
X_4 air pressure (bar)

Responses (Dependent Variables):

MC_1 material consumption of inside area for first spray gun (gram)
DFT_1 dry film thickness (µm.)

The Relationship Models:

$$MC_1 = -25.8725-0.300469X_1+29.2672X_2+0.0296319X_3 +11.8631X_4+0.000830556X_1X_3-0.0328472X_2X_3 -5.71250X_2X_4-0.0171111X_3X_4 \quad (1)$$

$$DFT_1 = -26.0625-0.89325X_1+31.5X_2+0.0211944X_3 +33.6625X_4+0.00251667X_1X_3-9.425X_2X_4 -0.07322X_3X_4 \quad (2)$$

Table 5: Optimal Conditions, MC_1 and DFT_1

Variable	Optimal value
Angle (degree)	60
Spray time (sec.)	1.2
Fan pattern (rev.)	315
Air pressure (bar)	2.7
MC_1(g.)	2.8
DFT_1(µm.)	10.45

Coating Process of Second Spray Gun

- **Statistical Analysis**

The model adequacy checking of the material consumption data shows that the p-value for the normality probability test, equal variance, and randomization test equal to 0.699 0.427 and 0.198 respectively. For the DFT, the p-value of the normality probability test, equal variance, and randomization test are equal to 0.522, 0.571 and 0.369 respectively as shown in Table 6. The result indicates that the MC and DFT data are in the normal distribution and there is no deviation in variance of each test. The analysis of variance is conducted and the result is summarized in Table 7 and 8. The p-value which is less than 0.05 is used as the criterion to determine the significance of the response.

Table 6: P-value of Model Adequacy

Model Adequacy Checking	P-Value	
	Material Consumption	Dry film thickness
Normal Distribution	0.699	0.522
Equal variance	0.427	0.571
Randomization	0.198	0.369

Table 7: Result of ANOVA Material Consumption for Second Spray Gun

Term	P-Value
Angle	0.000
Spray time	0.005
Fan pattern	0.000
Air pressure	0.000
Angle*Fan pattern	0.003
Spray time*Fan pattern	0.029
Spray time*Air pressure	0.000
Fan pattern*Air pressure	0.005
S = 0.308862 PRESS = 4.44022	
R-Sq = 96.87% R-Sq(pred) = 93.37% R-Sq(adj) = 95.59%	

Table 8: Result of ANOVA Dry Film Thickness for Second Spray Gun

Term	P-Value
Angle	0.000
Spray time	0.000
Fan pattern	0.000
Air pressure	0.000
Angle*Spray time	0.000
Angle*Fan pattern	0.003
Spray time*Fan pattern	0.005
Spray time*Air pressure	0.000
Fan pattern*Air pressure	0.000
S = 0.547875 PRESS = 12.8071	
R-Sq = 86.84% R-Sq(pred) = 76.61% R-Sq(adj) = 83.00%	

- **Effect of Spray Condition Analysis**

The effects of gun angle, spray time, fan pattern and air pressure on the material consumption and dry film thickness are shown in Figure 3 and 4. The results indicate that the increment of gun angle, fan pattern, and air pressure decreases the material consumption and dry film thickness. However, the idecrement in the spray time decreases the material consumption and dry film thickness.

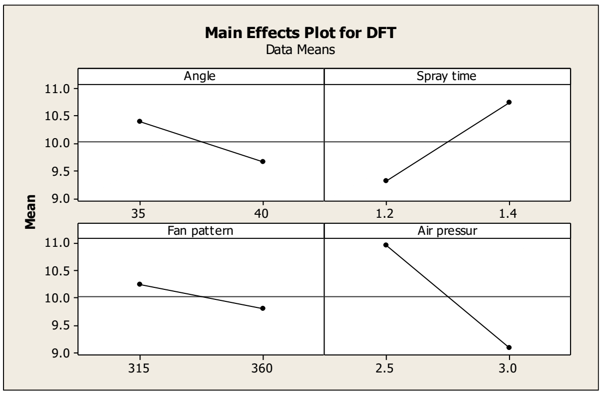


Figure 3: Main Effect Plot of Variable on Average Material Consumption for Second Spray Gun

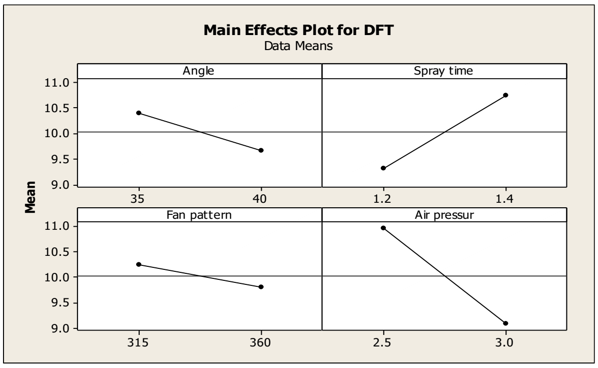


Figure 4: Main Effect Plots of Variables on Average Dry Film Thickness for Second Spray Gun

- **Process Optimization and Modeling**

The optimization of the spray coating process is performed to minimize the material consumption which the dry film thickness is in the range between 7.5-12.5 µm. The regression model is constructed to present the relationship between response, coating material consumption, and DFT as shown in (3) and (4). The optimal condition is the gun angle at 60 degrees, spray time at 1.2 sec., fan pattern at 315 rev. and air pressure at 2.7 bar. The material consumption and DFT is equal to 2.8 g. and 10.45 µm which is in the specification. Table 9 shows the optimal results.

Parameters (Indepentdent Variables):

X_1 angle (degree)
X_2 spray time (sec.)
X_3 fan pattern (rev.)
X_4 air pressure (bar)

Responses (Dependent Variables):

MC_2 material consumption of outside area for second spray gun (gram)
DFT_2 dry film thickness of outer radius for second spray gun (µm.)

The Relationship Models:

$$MC_2 = -21.6531-0.655187X_1+30.3516X_2+0.0113264X_3 +12.0244X_4+0.00183333X_1X_3-0.0327083X_2X_3 -6.10000X_2X_4-0.0159444X_3X_4 \quad (3)$$

$$DFT_2 = -86.5812-2.12275X_1+63.2062X_2+0.279889X_3 +38.6075X_4+0.745X_1X_2+0.00298889X_1X_3 -0.146389X_2X_3-12.60X_2X_4-0.077X_3X_4 \quad (4)$$

Table 9: Optimal Conditions, MC_2 and DFT_2

Variable	Optimal value
Angle (degree)	35
Spray time (sec)	1.23
Fan pattern (rev)	360
Air pressure (bar)	2.5
MC_2 (g.)	2.3
DFT_2 (µm.)	10.85

CONCLUSION

This study attempts to find the optimal solution for two spray gun process in order to minimize the coating material consumption while meeting the dry film thickness specification of 7.5-12.5 µm. The results also show that the significant parameters suchas gun angle, spray time, fan pattern, and air pressure to material consumption and DFT are in the same manner.

REFERENCES

Winnicki, M., Małachowska, A ,Ambroziak, A :Taguchi Optimization of the Thickness of a Coating Deposited by LPCS. Archives of Civil and Mechanical Engineering. 14 (2014) 561-568. [5]O. Poonkwan, V. Tangwarodomnukun and S. Prombanpong: Optimization of Teflon Spraying Process for Non-Stick Coating Application. Industrial Engineering. (2015), p. 833-839

From PJ, Gunnar J, Gravdahl JT. Optimal paint gun orientation in spray paint applications-experimental results. Proc. IEEE Transaction on Automation Science and Engineering (2011) p. 438-442.

Luangkularb, S.,Prombanpong, S.: Material Consumption and Dry Film Thickness in Spray Coating Process. Proceedings of the 47th CIRP Conference on Manufacturing Systems. 17 (2014) 789-794.

Song, E.P., Ahn, J., Lee, S., Kim, N.J.: Effects of Critical Plasma Spray Parameter and Spray Distance on Wear Resistance of Al2O3–8 wt.%TiO2 Coatings Plasma-Sprayed with Nanopowders. Proceeding of Surface and Coatings Technology. Vol.202 (2008), p.3625-3632

S. Hong, Y. Wu, B. Wang, Y. Zheng, W. Gao and G. Li: High-Velocity Oxygen-Fuel Spray Parameter Optimization of Nanostructured WC–10Co–4Cr Coatings and Sliding wear Behavior of the Optimized Coating. Materials and Design. Vol. 55 (2014), p. 286-291

AUTHOR BIOGRAPHIES

NITCHAKAN SOMBOONWIWAT is a graduate student in Industrial and Manufacturing System Engineering program at the Department of Production Engineering, King Mongkut's University of Technology Thonburi, Thailand. She received her bachelor's degree in Tool and Material Engineering from the same university. Her research interests are applied statistics and operations management. Her e-mail address is : nonpang24@hotmail.com.

Prof. Dr. SUKSAN PROMBANPONG is an Associaate Professor in the Industrial Management section, Department of Production Engineering Faculty of Engineering, King Mongkut's University of Technology Thonburi, (KMUTT) Thailand. He received Fulbright scholarship to pursue his master and doctoral degree at The Ohio State University (OSU) U.S.A. After graduation he worked at National University of Singapore for three years. Now he is a lecturer in the department of Industrial Engineering at KMUTT. He also holds a position of DEPUTY DIRECTOR for the Institute for Scientific and Technological Research and Services and a head of Continuing Education Center at KMUTT. His research interest is in the area of manufacturing applications, production planning and control, optimization, logistic and supply chain and green industrial. His e-mail address is : suksan.pro@gmail.com.

Simulation of Intelligent Systems

ON THE EFFECT OF NEIGHBORHOOD SCHEMES AND CELL SHAPE ON THE BEHAVIOUR OF CELLULAR AUTOMATA APPLIED TO THE SIMULATION OF SUBMARINE GROUNDWATER DISCHARGE

Christoph Tholen and Lars Nolle
Department of Engineering Science
Jade University of Applied Science
Friedrich-Paffrath-Straße 101
26389 Wilhelmshaven, Germany
Email:
{christoph.tholen|lars.nolle}@jade-hs.de

Oliver Zielinski
Institute for Chemistry and Biology
of the Marine Environment
Carl von Ossietzky University of Oldenburg
Schleusenstraße 1
26382 Wilhelmshaven, Germany
Email: oliver.zielinski@uol.de

KEYWORDS
SGD, CDOM, FDOM, coastal recirculation, cellular automata, neighbourhood, cell shape.

ABSTRACT

In order to design new search strategies for collaborating autonomous underwater vehicles, a novel simulator was developed to model the diffusion of groundwater discharge in shallow coastal waters. The simulation allows for the evaluation of new search strategies without running the risk of losing expensive hardware during the field testing.

The developed simulation is based on cellular automata. In order to reduce computational complexity, a novel two-dimensional cellular automaton with additional depth-information for each cell is used to simulate a three-dimensional nearshore environment.

The influence of different neighbourhoods and cell shapes on the behaviour of the cellular automaton is examined and discussed. Results show a faster rise of discharged fluorescent dissolved organic matter for hexagon cells. Also all examined neighbourhoods converge to a stable state after a finite number of iterations.

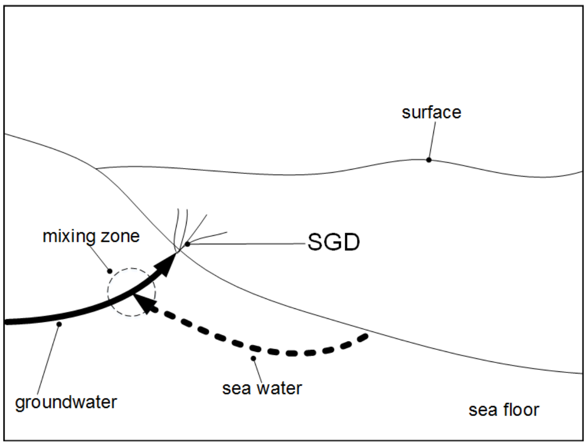
INTRODUCTION

The long term goal of this project is to develop a low cost and flexible environmental observatory, based on a swarm of autonomous underwater vehicles (AUV). AUVs can be used for the exploration of intermediate size areas and the precise measurement of parameters, for example oxygen, nutrients or Fluorescent Organic Matter (FDOM) (Zielinski et al. 2009). The interaction of the swarm should be managed by a search strategy, such as particle swarm optimisation (Nolle 2015). The search for submarine groundwater discharges (SGD) in coastal waters is one of the possible applications for such an observatory. Marine scientists are interested in locating and analysing these discharges because the nutrients discharged by SGD have a significant influence on the marine ecosystem (Dugan et al. 2010; Moore 2010; Nelson et al. 2015).

The area under investigation is a section of the north-western beach of the Spiekeroog island in the north-west of Germany. This area is in focus of many research groups because it represents a coastal transition zone at a barrier island with a freshwater lense (Röper et al. 2014; Beck et al. 2017). The model developed in this work simulates the three-dimensional environment using a two-dimensional cellular automaton.

Submarine Groundwater Discharge

Submarine groundwater discharge (SGD) consists of a flow of fresh groundwater and the recirculation of seawater from the sea floor to the coastal ocean (Moore 2010). The fresh water and the sea water discharges commingle in the so-called mixing zone (Figure 1) (Evans and Wilson 2016).



Figures 1: Submarine Groundwater Discharge of Fresh- and Recirculating-Water, modified after Evans and Wilson (2016)

The freshwater that flows to the ocean is a continuous and significant source of nutrients for the costal marine environment. Furthermore, the freshwater contains dissolved organic matter (DOM) (Nelson et al. 2015). The main source of this DOM is the dissolution of soil and terrestrial organic matter (Coble 2013). With a mass

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

of approximately 700 Gt, DOM represents one of the largest organic carbon reservoirs on Earth, equalling the bio mass on land surface (Hedges 1992).

Coloured Dissolved Organic Matter

Coloured dissolved organic matter (CDOM) is the part of the DOM-pool that interacts with solar radiation (wavelengths 280 – 700 nm). The gain size of the particles is smaller than 0.2 - 0.4 µm (Nelson and Siegel 2013). CDOM have a major influence on the light distribution in sea water (Stedmon et al. 2010; Coble 2013).

A small part of the CDOM-pool is also fluorescent. This part is called fluorescent dissolved organic matter (FDOM). Fluorescent methods are used to analyse the chemical composition and the amount of DOM present in a sea water sample. The fluoresce of a sample is measured in quinine sulphate equivalent units (QSE). The QSE relates the intensity of FDOM to the fluoresce intensity of a standard compound (Kowalczuk et al. 2010).

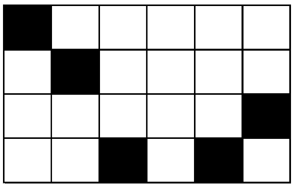
Optical methods are highly suitable for detecting spatio-temporal patterns of relevant biogeochemical parameters like dissolved organic matter or nutirents (Moore et al. 2009; Zielinski et al. 2011). In this work, the concentration of FDOM is modelled and simulated as described below.

SIMULATION

The simulation developed here is based on cellular automata (CA). This section introduces the basic concepts of CA and describes the main principles of the simulation developed. The rules used for the CA are introduced and an overview of different neighbourhood schemes is provided.

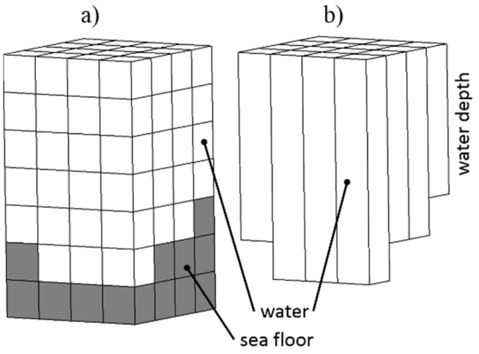
Cellular Automata

Cellular automata (CA) are mathematical models used for the simulation of complex systems. They consist of a finite collection of identical cells. Each cell has a current state, which is updated after one time step. The state of a cell in the next time step is based on its own current state and the current state of its neighbours. A CA discretizes a system in space and time (Wolfram 1984). Figure 2 shows an example of a two-dimensional cellular automaton with the dimensions 6 x 4. In the example, black cells are in state “true” and white cells are in state “false”.



Figures 2: Cellular Automaton of Size 6 x 4

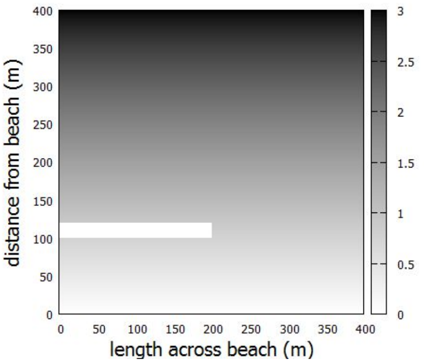
In principle, a three-dimensional model is needed to simulate the environment under investigation. However, due to the large amount of cells needed, this would be computationally expensive. Since the area to be simulated is very shallow and the water column is usually well mixed due to strong currents and waves in coastal waters (Röper et al. 2014), the simulation is based on a two-dimensional cellular automaton with additional depth-information for each cell (Figure 3). This reduces the number of cells significantly and therefore the computational effort needed. Figure 3a shows a three dimensional CA. It can be seen that the number of cells in the example is 112 including cells acting as boundaries (sea floor). In the two-dimensional representation (Figure 3b) on the other hand, only 16 cells are needed to model the same area.



Figures 3: Three-dimensional CA (left) and two-dimensional Representation (right)

The implemented simulation-environment covers an area of 400 m x 400 m. This area is divided into a number of symmetric cells. Each cell has an x- and a y-position as well as a depth and a FDOM level.

The depth increases from the beach (y-value = 0 m) to the open sea (y-value = 400 m) in steady steps. Furthermore, there is an obstacle (sandbank) with the depth of 0 m (Figure 4).



Figures 4: Depth-Profile of the Simulation Developed

The FDOM values in sea water are subject to wide variations (Kowalczuk et al. 2010). Therefore, in the simulation, seawater is assigned a FDOM level of one arbitrary unit, while the FDOM level of discharges was chosen to be 100 arbitrary units. Each cell is initialized with a FDOM-level of one unit.

To simulate the groundwater discharge two springs are added to the simulation. These springs have a volume flow rate of 0.125 m^3/iteration and a FDOM level of 100 units. The springs are located at position (100 m /200 m) and (200 m / 200 m). The left, right and top boundaries are located in the open sea. To simulate the exchange of water and nutrients with the open sea, cells located at this borders are acting as sinks. If a cell is a sink, it has a constant FDOM level of one unit. This ensures a steady flow of fresh seawater to the simulated environment and it prevents an enrichment of FDOM in the simulated environment.

Rules

The interaction of cells with their neighbours is based on a set of application specific rules. These rules define the dynamic behaviour of the model (Wolfram 1984; Nolle et al. 2016).

The developed CA is based on one simple rule only; the FDOM-value of a cell x at time step t+1 is calculated as the weighted average of the FDOM-values of the cell x and its neighbouring cells y_n and, if present, the FDOM-value of an existing spring at the position of the cell x. All values will be multiplied with the volume of the cells respectively with the volume flow rate of the spring. The sum will be divided by the sum of the volume of all cells in the neighbourhood. The rule is given in Equation (1), where I is representing the FDOM level in arbitrary units and V is representing the volume of the cells.

$$I_x^{t+1} = \frac{I_x^t * V_x^t + \sum\left(I_y^t * V_y^t\right) + I_s^t * V_s^t}{V_x^t + \sum\left(V_y^t\right) + V_s^t} \qquad (1)$$

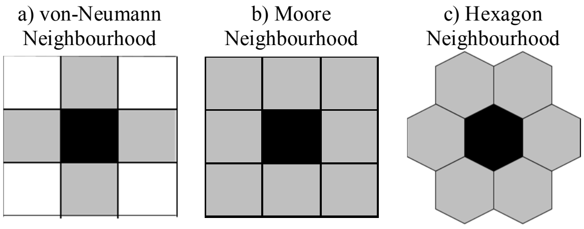
Where:
I_x^{t+1}: FDOM level in cell x in iteration $t+1$
I_x^t: FDOM level in cell x in iteration t
V_x^t: Volume of cell x in iteration t
I_y^t: FDOM level in neighbour cell y_n in iteration t
V_y^t: Volume of neighbour cell y_n in iteration t
I_S^t: FDOM level of spring located at cell x in iteration t
V_S^t: Volume flow of spring located at cell x in iteration t

In addition to the rules, the behaviour of a CA also depends on the chosen neighbourhood and cell shape.

Neighbourhoods and Cell Shapes Used

The most popular neighbourhoods are the von-Neumann neighbourhood (Figure 5a) and the Moore neighbourhood (Figure 5b) (Jiménez et al. 2005). Both are using square-shaped cells (Tzedakis et al. 2015).

In order to study the influence of the cell shape and neighbourhood on the behaviour of the CA, hexagon shaped cells (Figure 5c) (Gerhardt and Schuster 2012) are compared to square cells using the von-Neumann neighbourhood and the Moore neighbourhood.



Figures 5: Different Neighbourhood Definitions for two-dimensional Cellular Automata

Cells have defining properties, like size of the area covered or the length of the perimeter, which depend on each other. In the two-dimensional model developed here, the area size represents the volume of the cell whereas the length of the perimeter represents the surface area of the cell, through which it interacts with its neighbours. The ratio between area size and perimeter length depends on the shape of the cells. However, the ratio is different for square-shaped- and hexagon-shaped cells (Birch et al. 2007), i.e. they cannot be kept the same. In order to allow for a fair comparison of the different neighbourhood schemes, two hexagon-shaped automata are used in the experiments, one with the same area size as the square-shaped cells and one with the same perimeter length.

EXPERIMENTS

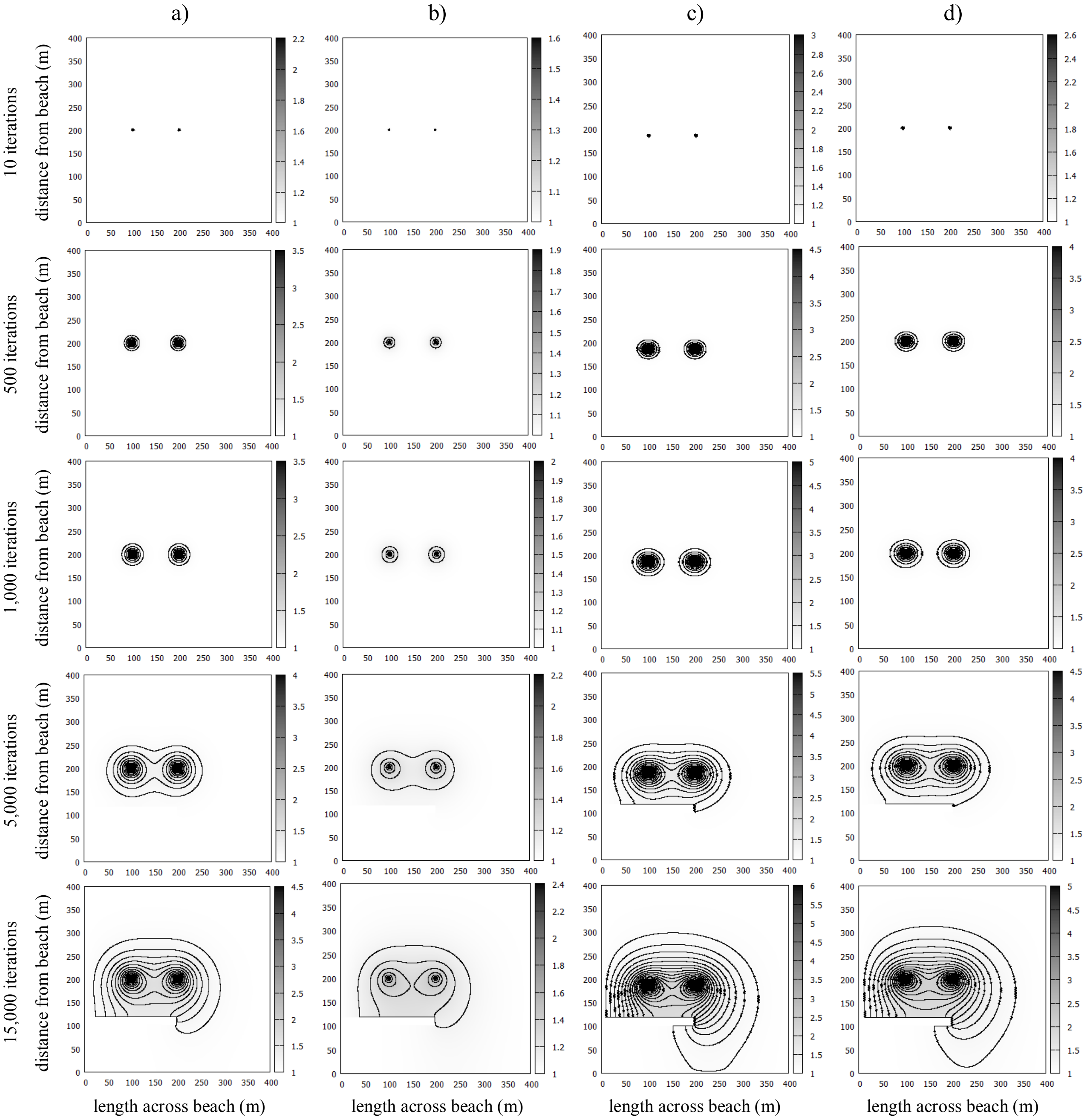
For the experiments, four different neighbourhoods were used, the von-Neumann neighbourhood, the Moore neighbourhood and two different versions of the hexagon neighbourhood. Each simulation was allowed to run for 15,000 iterations. The FDOM-level of each cell was logged for every iteration.

RESULTS

Figures 6a - d provide the FDOM-distributions after 10, 500, 1,000, 5,000 and 15,000 iterations for the different neighbourhoods used. The iso-lines show FDOM levels in steps of 0.1 units.

Each neighbourhood exhibits a dispersion of freshwater into the simulated costal area over the run time of the simulation. As expected, all neighbourhoods yield a symmetric distribution of FDOM around both springs. When the sandbank obstructs the distribution of FDOM it results in an asymmetric distribution of FDOM in the simulated area (Figure 6a - d).

At the beginning, the distribution of FDOM in both hexagon shaped neighbourhoods seems similar. Only when the sandbank obstructs the distribution the behaviour of the hexagon shaped neighbourhoods differ (Figure 6c and d).



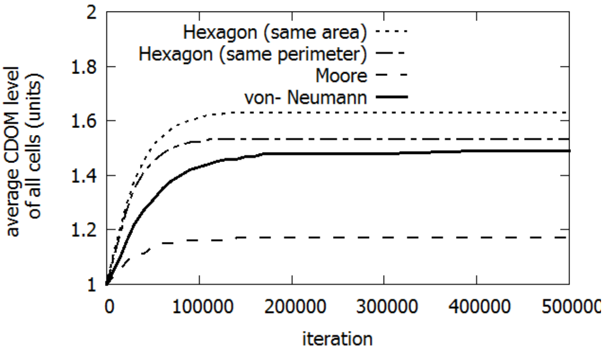
Figures 6: Distribution of FDOM over Time using the von-Neumann Neighbourhood (a), the Moore Neighbourhood (b), a Hexagon Neighbourhood with Same Area Size (c) and a Hexagon Neighbourhood with Same Perimeter Length (d)

Nearby the springs the FDOM level increases significant, while in all other areas the FDOM level increase only sparsely (Figure 6a - d). This behaviour agrees with the expected behaviour of small submarine groundwater discharges in coastal waters (Nelson et al. 2015).

The two springs represent a continuous inflow of FDOM units into the simulated area over the run. Because of that the average FDOM level for all cells has to increase over time. Although the FDOM input flow is the same for all four CAs, the development of the average FDOM levels differ (Figure 7) and hence, the neighbourhood schemes used have an influence of the distribution of FDOM in the area simulated.

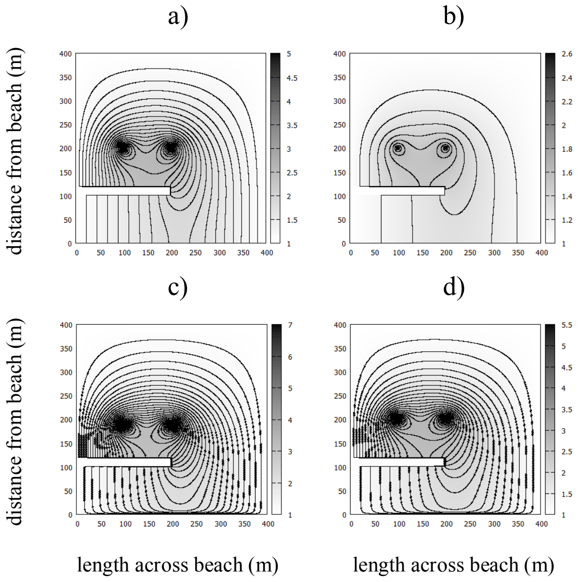
In such a state the amount of FDOM units supplied by the sources equals the amount of FDOM units that are lost on the boundaries of the CA. In this state the average of the FDOM level increase not anymore and the simulations converge towards stable states after a finite number of iterations.

To examine this behaviour of the different neighbourhoods, each simulation was run for 500,000 iterations. During the runs the average FDOM levels for all cells have been logged (Figure 7).



Figures 7: Convergence of Average FDOM level of all Cells for the Different Neighbourhoods

At the beginning the average increase fast for the hexagon shaped cells and the von-Neumann neighbourhood, while the average of the simulation using the Moore neighbourhood increase slower. However all used neighbourhoods converge to different stable states, i.e. the average value, the number of iterations and the appearance of the FDOM distribution (Figure 8) differ for the four used neighbourhoods.



Figures 8: Distribution of FDOM in the Stable State for the von Neumann (a)-, the Moore (b)-, the Hexagon-shaped same Area (c)- and the Hexagon-shaped same Perimeter (d)-neighbourhood

The convergence behaviour of the neighbourhoods using hexagon-shaped cells are similar (Figure 8c and d) while the behaviour of the neighbourhoods using square-shaped cells differ (Figure 8a and b). It can be observed that the FDOM level in the simulation using the von-Neumann neighbourhood expand much faster than that using the Moore neighbourhood (Figure 6 a and b).

Furthermore the average FDOM level in the stable state for the simulation using the von-Neumann neighbourhood is 1.49 units, while the average in the simulation using the Moore-neighbourhood is 1.17 units. However, the von-Neumann neighbourhood needs around 400,000 iterations to converge, while the Moore neighbourhood converges after around 140,000 iterations.

The parameter used to compare the hexagon shaped cells with the square shaped cells, i.e. either same area covered or same perimeter length, have an influence on the behaviour of the CA. While the simulation using hexagon shaped cells with the same area converges after around 130,000 iterations, the other simulation using hexagon shaped cells converges after around 110,000 iterations. A summary of these results is presented in Table 1.

Table 1: Convergence of Different Neighbourhoods

Neighbourhood	Iterations	Average
von-Neumann	400,000	1.49
Moore	140,000	1.17
Hexagon same Area	130,000	1.63
Hexagon same Perimeter	110,000	1.53

CONCLUSION AND FUTURE WORK

The aim of this research was to develop and determine a simulation based on cellular automata, which can be used for the evaluation of different search strategies for a swarm of autonomous underwater vehicles. The focus of this study was on the influence that different neighbourhoods might have on the behaviour of cellular automata. In order to reduce the computational effort, a novel two-dimensional CA was implemented that simulates a three-dimensional environment by treating depth, i.e. the third dimension, as a property of the cells.

From the experiments it can be seen that both, cell shape and neighbourhood affects the behaviour of the CAs. As depicted in Figure 7, the FDOM levels rise faster and the average level in stable state is higher for CAs with hexagon shaped cells. The behaviour of the simulations using hexagon shaped cells seems similar. It is estimated that the chosen parameters to compare the different cell shapes (i.e. area covered or perimeter length) have no significant influence on the behaviour of the CA (Figure 6c and d, Figure 8c and d).

In the next phase of this research it is proposed to use real FDOM measurements from the island of Spiekeroog to fine-tune the rule base of the Cellular Automaton and to incorporate waves and tides into the model.

The latter would allow for the simulation of additional transportation of water and FDOM between cells in the direction of waves. This will be realised by changing the volume of the cells periodically, with different amplitudes and frequencies for waves and tides. Also, some non-deterministic behaviour of the waves will be modelled and incorporated into the simulation.

This would then in turn allow for the cost-effective evaluation of different search strategies for swarms of autonomous underwater vehicles before such swarms are physically deployed to search for FDOM sources near the island of Spiekeroog.

REFERENCES

Beck,M., Reckhardt, A., Amelsberg, J., Bartholomä, A., Brumsack, H.J., Cypionka, H., Dittmar, T., Engelen, B., Greskowiak, J., Hillebrand, H., Holtappels, M., Neuholz, R., Köster, J., Kuypers, M.M.M., Massmann, G., Meier, D., Niggemann, J., Paffrath, R., Pahnke, K., Rovo, S., Striebel, M., Vandieken, V., Wehrmann, A., Zielinski, I., "The drivers of biogeochemistry in beach ecosystems: A cross-shore transect from the dunes to the low water line." *Marine Chemistry* 2017, doi: 10.1016/j.marchem.2017.01.001.

Birch, C.P., Oom, S.P. and Beecham, J.A. 2007 "Rectangular and hexagonal grids used for observation, experiment and simulation in ecology", *Ecological Modelling*, 206(3-4), pp. 347–359.

Coble, P.G. 2013 "Colored dissolved organic matter in seawater". In Subsea Optics and Imaging 2013, J. Watson and O. Zielinski (Eds.). WOODHEAD PUBLISHING Oxford, Cambridge, Philadelphia and New Delhi, 98-118.

Dugan, J.E., Defeo, O. Jaramillo, E., Jones, A.R., Lastra, M., Nel, R., Peterson, C.H., Scapini, F., Schlacher, T., Schoeman, D.S. 2010 "Give Beach Ecosystems Their Day in the Sun" in *Science*, Vol. 329, ISSUE. 5996, p 1146.

Evans, T.B. and Wilson, A.M. 2016 "Groundwater transport and the freshwater–saltwater interface below sandy beaches", *Journal of Hydrology*, 538, pp. 563–573.

Gerhardt, M. and Schuster, H. 2012 *Das digitale Universum*, Vieweg & Teubner.

Hedges, J.I. 1992 „Global biogeochemical cycles: progress and problems", *Marine Chemistry*, Vol. 39, Issues 1-3, pp 67-93.

Jiménez, A., Posadas, A.M. and Marfil, J.M. 2005 "A probabilistic seismic hazard model based on cellular automata and information theory", *Nonlinear Processes in Geophysics*, 12(3), pp. 381–396.

Kowalczuk, P., Zabłocka, M., Sagan, S. and Kuliński, K. 2010 "Fluorescence measured in situ as a proxy of CDOM absorption and DOC concentration in the Baltic Sea", *OCEANOLOGIA*, 52(3), pp. 431–471.

Moore, C., Barnard, A., Fietzek, P., Lewis, M.R., Sosik, H.M., White, S. and Zielinski, O. 2009 "Optical tools for ocean monitoring and research", *Ocean Science*, Vol. 5, pp 661-684.

Moore, W.S. 2010 "The effect of submarine groundwater discharge on the ocean", *Annual Review of Marine Science*, 2, pp. 59–88.

Nelson, C.E., Donahue, M.J., Dulaiova, H., Goldberg, S.J., La Valle, F.F., Lubarsky, K., Miyano, J., Richardson, C., Silbiger, N.J. and Thomas, F.I. 2015 "Fluorescent dissolved organic matter as a multivariate biogeochemical tracer of submarine groundwater discharge in coral reef ecosystems", *Marine Chemistry*, 177, pp. 232–243.

Nelson, N.B. and Siegel, D.A. 2013 "The global distribution and dynamics of chromophoric dissolved organic matter", *Annual Review of Marine Science*, 5, pp. 447–476.

Nolle, L. 2015 "On a search strategy for collaborating autonomous underwater vehicles",In *Proceedings of Mendel 2015, 21st International Conference on Soft Computing*, Brno, CZ, pp. 159–164.

Nolle, L., Thormählen, H. and Musa, H. 2016. "Simulation of Submarine Groundwater Discharge of Dissolved Organic Matter Using Cellular Automata". In *Proceedings 30th European Conference on Modelling and Simulation ECMS* 2016. Regensburg (Germany), 31 May - 3 June.

Röper, T., Greskowiak, J. and Massmann, G. 2014 "Detecting small groundwater discharge springs using handheld thermal infrared imagery", *Ground Water*, 52(6), pp. 936–942

Stedmon, C.A., Osburn, C.L. and Kragh, T. 2010 "Tracing water mass mixing in the Baltic–North Sea transition zone using the optical properties of coloured dissolved organic matter", *Estuarine, Coastal and Shelf Science*, 87(1), pp. 156–162.

Tzedakis, G., Tzamali, E., Marias, K. and Sakkalis, V. 2015 "The Importance of Neighborhood Scheme Selection in Agent-based Tumor Growth Modeling", *Cancer Informatics*, 14(Suppl 4), pp. 67–81.

Wolfram, S. 1984 "Universality and complexity in cellular automata", *Physica D: Nonlinear Phenomena*, Vol. 10, Issue 1-2, pp 1-35.

Zielinski, O., Busch, J.A., Cembella, A.D., Daly, K.L., Engelbrektsson, J., Hannides, A.K. and Schmidt, H. 2009 "Detecting marine hazardous substances and organisms: sensors for pollutants, toxins and pathogens", *Ocean Science,* Vol. 5, pp 329-349.

Zielinski, O., Voß, D., Saworski, B., Fiedler, B. and Körtzinger, A. 2011 "Computation of nitrate concentrations in turbid coastal waters using an in situ ultraviolet spectrophotometer", *Journal of Sea Research,* Vol. 65, pp 456-460.

AUTHOR BIOGRAPHIES

CHRISTOPH THOLEN graduated from the Jade University of Applied Science in Wilhelmshaven, Germany, with a Master degree in Mechanical Engineering in 2015. Since 2016 he is a research fellow at the Jade University of Applied Science in a joint project of the Jade University of Applied Science and the Institute for Chemistry and Biology of the Marine Environment (ICBM), at the Carl von Ossietzky University of Oldenburg for the development of a low cost and intelligent environmental observatory.

LARS NOLLE graduated from the University of Applied Science and Arts in Hanover, Germany, with a degree in Computer Science and Electronics. He obtained a PgD in Software and Systems Security and an MSc in Software Engineering from the University of Oxford as well as an MSc in Computing and a PhD in Applied Computational Intelligence from The Open University. He worked in the software industry before joining The Open University as a Research Fellow. He later became a Senior Lecturer in Computing at Nottingham Trent University and is now a Professor of Applied Computer Science at Jade University of Applied Sciences. His main research interests are computational optimisation methods for real-world scientific and engineering applications.

OLIVER ZIELINSKI is Director of the Institute for Chemistry and Biology of the Marine Environment (ICBM) at the Carl von Ossietzky University of Oldenburg and head of the research group "Marine

Sensor Systems.” After receiving his Ph.D. degree in Physics in 1999 from University of Oldenburg, he moved to industry where he became scientific director and CEO of “Optimare Group,” an international supplier of marine sensor systems. In 2005, he was appointed Professor at the University of Applied Science in Bremerhaven, Germany. In 2007, he became Director of the Institute for Marine Resources (IMARE). He returned to the Carl von Ossietzky University of Oldenburg in 2011. His area of research covers marine optics and marine physics, with a special focus on coastal systems, marine sensors, and operational observatories involving different user groups and stakeholders.

Application of Genetic Optimization Algorithms to Lumped Circuit Modelling of Coupled Planar Coils

Jens Werner[1], Lars Nolle[1], Jennifer Schütt[2]
[1] Jade University of Applied Science Wilhelmshaven/Oldenburg/Elsfleth,
Friedrich-Paffrath-Str. 101, D-26389 Wilhelmshaven,
Email: jens.werner@jade-hs.de, lars.nolle@jade-hs.de
[2] Nexperia Germany GmbH, Stresemannallee 101, D-22529 Hamburg
Email: jennifer.schuett@nexperia.com

KEYWORDS

Coupled planar coils; Common mode filter; Lumped network model; Genetic algorithm; SPICE model

ABSTRACT

In portable electronic devices, like smart phones, coupled planar coils are often used as common mode filters (CMF). The purpose of these CMF is to suppress electromagnetic interference (EMI) between wireless communications systems (e.g. WIFI) and digital high-speed interfaces (e.g. USB 3). A designer of such an electronic device usually carries out a signal integrity (SI) analysis, using models of the system components. There are two alternative ways of modelling the CMF: One is based on matrices (called S-parameters) that describe the behaviour in the frequency domain and are either derived from measurements or simulation tools. The other is using a representation based on lumped circuit networks. In this work, a lumped network is generated manually based on expert knowledge. The advantage of this approach is the reduced number of only passive network components compared to traditional methods that produce much larger networks comprising of many active and passive devices. On the other hand, suitable component values of the lumped network need to be found so that the network exhibits the same frequency response as the physical device. Since there are many interacting parameters to be tuned, this cannot be achieved manually. Hence, a genetic algorithm is applied to this optimisation problem. Two sets of experiments were carried out and a sensitivity analysis has been conducted. It has been shown that the proposed method is capable of finding near optimal solutions within reasonable computation time.

INTRODUCTION

Modern portable electronic devices have to be as compact as possible whilst being efficient. In such devices, miniaturized planar coils can be found, for example, in common mode filters which are built directly into integrated circuits.

For the design of these coils, sophisticated simulation tools based on the method of moments (Keysight, 2016; Harrington, 1968) are often used. Usually, these simulations return frequency dependent scattering parameters (S-parameters, (Pozar, 2012)). This is a black-box approach describing the electrical behaviour in the frequency domain at the ports (or: terminals) of the device without revealing details about the physics of the internal structure (1).

$$[\underline{S}] = \begin{bmatrix} \underline{S}_{dd,11} & \underline{S}_{dd,12} & \underline{S}_{dc,11} & \underline{S}_{dc,12} \\ \underline{S}_{dd,21} & \underline{S}_{dd,22} & \underline{S}_{dc,21} & \underline{S}_{dc,22} \\ \underline{S}_{cd,11} & \underline{S}_{cd,12} & \underline{S}_{cc,11} & \underline{S}_{cc,12} \\ \underline{S}_{cd,21} & \underline{S}_{cd,22} & \underline{S}_{cc,21} & \underline{S}_{cc,22} \end{bmatrix} \quad (1)$$

In matrix (1) all elements are complex valued and frequency dependent. The two parameters $\underline{S}_{dd,21}$ and $\underline{S}_{cc,21}$ describe the transmission characteristics between the two ports for two different modes of operation, differential mode (dd) and common mode (cc). Those parameters with mixed indices ($\underline{S}_{dc,ij}$ and $\underline{S}_{cd,ij}$) are relevant for the conversion from one mode to another; for that reason the elements of matrix (1) are also called mixed-mode S-parameters (Bockelman and Eisenstadt, 1995). For passive and reciprocal devices the forward and reverse transmission parameters $\underline{S}_{x,21}$ and $\underline{S}_{x,12}$ are identical. The focus in this work is on $\underline{S}_{dd,21}$ and $\underline{S}_{cc,21}$, which are the most important parameters for real world applications.

Fig. 1 shows the two modes of operation for a two-port device, like a CMF. The two ports are denoted by the indices 1 and 2 as used in (1). In differential mode (Fig. 1 a)) opposing currents (I_+, I_-) are applied to the pins of port 1. There is no current in the ground path ($I_{gnd} = 0$). The transfer characteristics of this mode from port 1 to port 2 is described by $\underline{S}_{dd,21}$.

In common mode (Fig. 1 b)), the pins of port 1 are driven commonly, resulting in a return current of the same amplitude in the ground path. The transfer characteristics of this mode from port 1 to port 2 is described by $\underline{S}_{cc,21}$.

However, sometimes a designer needs to analyse the behaviour in the time domain as well. Here, an equivalent circuit model is usually required, which exhibits the

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

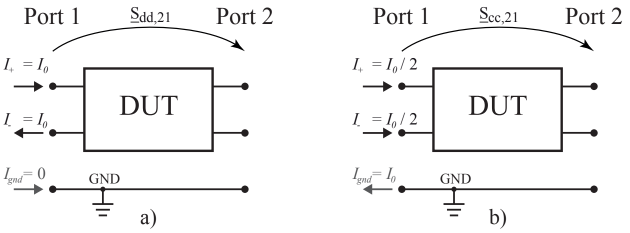


Fig. 1. Differential mode and common mode currents of a two-port device under test (DUT).

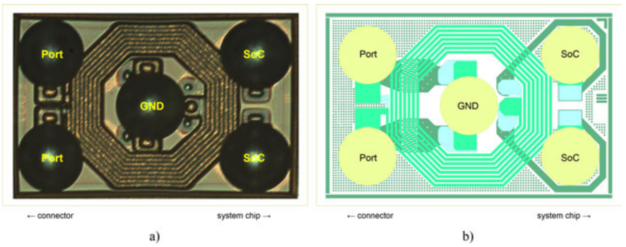


Fig. 2. Planar top view: a) die photograph; b) layout in EM tool (Werner et al., 2016).

same characteristics as the coil in both, the time and the frequency domain. In general, broadband SPICE generators (Simulation Program with Integrated Circuit Emphasis) produce networks of idealised components (so called lumped networks), which exhibit the required behaviour at their terminals (Stevens and Dhaene, 2008). In practice, these networks consist of a large number of active and passive components, which do not resemble the physics of the coils. The approach presented in this paper uses expert knowledge to generate a network which is closer to the actual physics whilst being much smaller. The challenge in this approach is that the component values of this network have to be determined. However, even for networks with few components the search space is too vast to find the optimum values manually. Therefore, a computational optimisation approach is used in this work and applied to a common mode filter (CMF) that consists of two planar coupled coils.

APPLICATION

In high-speed differential data lines (e.g. USB, SATA, PCIe) the spectrum used overlaps with wireless radio communication bands (e.g. GSM, LTE, WIFI). Electromagnetic interference (EMI) for example caused by USB 3 data lines can potentially disturb wireless applications in the 2.4 GHz band as reported in (Intel, 2012; Chen et al., 2013; USB, 2016). In the same manner, the spectrum of an USB 2 signal might interfere with the GSM 900 MHz downlink spectrum (Werner et al., 2015). This can cause a degradation of the receiver sensitivity in a mobile phone.

As wireless and wired signals utilise the same frequency band, a low pass filter can not be used to reduce unwanted interference from data lines into an antenna and finally the receiver. Instead, a common mode filter is applied. The purpose of a CMF is to suppress unwanted common mode (CM) currents on the data lines, since those currents are responsible for the EMI. Ideally, the differential mode (DM) currents of the wanted high-speed signals are not affected. Typically, CMFs are built using two coupled coils.

Common mode filter with on-chip planar coils

Fig. 2 a) shows the top view on a planar copper coil as part of an integrated on-chip CMF (Werner et al., 2016). In addition Fig. 2 b) presents the layout view from the EM design tool (Keysight, 2016).

For all kind of digital high-speed transmission systems, a signal integrity (SI) analysis is performed in the time domain. This involves the simulation of a large number of transmitted data bits and the evaluation of their electrical characteristics at each sampling point in the time domain. The segmented and overlapped representation of this time domain data is known as eye diagram (Gao et al., 2010; Ahmadyan et al., 2015). Here, the simulation in the time domain benefits from an exact and compact SPICE model of the involved planar structures.

Circuitry

The fundamental common mode filtering is achieved by two planar copper coils. The chosen device-under-test (DUT) provides furthermore ESD protection in order to avoid destruction of the sensitive CMOS system-on-chip (SoC). This protection is accomplished by two low-voltage triggered semiconductor controlled rectifiers (LVTSCR) which will shunt the current of an ESD pulse into the ground (GND) connection. These LVTSCR are depicted by the diode symbols in Fig. 3. In an real application, like a smart phone with USB 3 connector, the diodes would be mounted towards the external interface, in order to protect the SuperSpeed receive (SSRX) or transmit (SSTX) signal lines.

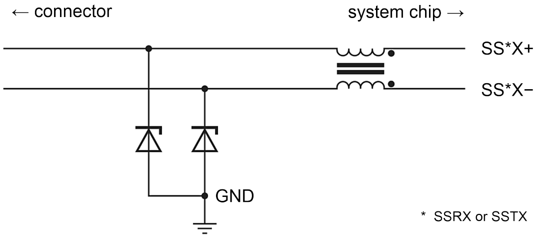


Fig. 3. Functional diagram (Werner et al., 2016).

Physical realisation

The filter device is fabricated in a planar bipolar semiconductor process with two metal layers made by aluminium. In addition, three layers of polyimide and three layers of copper are added on top of the planar wafer. Two of these copper layers form the planar coupled coils. There is no plastic package surrounding the silicon die, instead the device is attached with solder balls and manufactured as wafer level chip scale package. The contact side of the CMF is depicted in Fig. 2 a). The five solder balls marked with "Port", "SOC" and "GND" can be clearly identified as well as the octagonal shaped coil around the GND ball. The layout of this CMF, generated in an EM simulation tool, is shown in Fig. 2 b). It marks in the same manner the five ball pads. Due to the top view perspective only the upper coil layer can be seen.

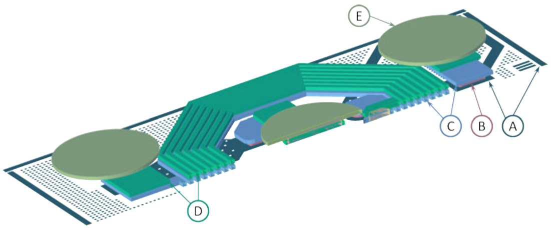
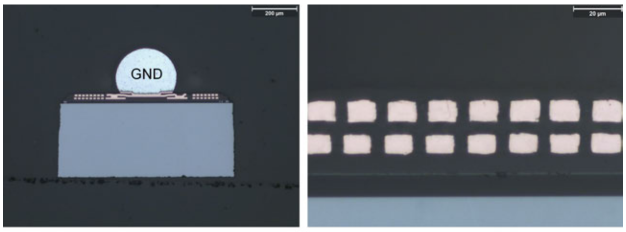


Fig. 4. Cross sectional view from EM simulation tool showing all metal layers (without solder balls) (Werner et al., 2016).



a)

b)

Fig. 5. Cross sectional view from scanning electron microscope: a) symmetric cut through GND ball; b) detailed view on upper and lower copper coil layer (Werner et al., 2016).

The cross sectional view in Fig. 4 provides a more detailed perspective into the structure:

The symbol "A" marks the top level metal of the semiconductor process which is needed to implement the basic semiconductor schematic. Furthermore it provides the interconnection to the 3-metal/3-polyimide-process. "B" denotes the interconnect between first copper layer and top aluminium layer. "C" and "D" mark the two copper layers which are the core of this common mode filter. These layers determine dominantly the inductance and resistance of the coils. The third copper layer, marked by "E", enables the so-called under bump metal, a drop area for the solder ball.

Fig. 5 depicts the physical cross section of the manufactured CMF. Fig. 5 a) shows a cross section that is orthogonal to the cut plane as shown in Fig. 4. This picture gives a good impression about the real relation between silicon thickness, solder ball and copper structures. As it can be seen in Fig. 5 b) the planar coil windings are rectangular shaped with a larger width than height and a spacing in the same range as the height. The typical spacing between the coil windings is 5 µm.

STRUCTURAL MODELLING

As mentioned above, when a designer needs to analyse the behaviour of an entire system in the time domain, for example for a signal integrity analysis, a lumped network can offer a solution. In contrast to transforming S-parameters directly into the time domain, this approach has the advantage that causality, passivity and reciprocity (Triverio et al., 2007) are inherently fulfilled.

One of the simplest schematic diagrams, one could think of, that represents the electric function as shown in Fig. 3 is given in Fig. 6: The two planar coils are represented by inductors L1a and L1b, while the magnetic coupling is modelled by the coupling coefficient kval. Electric and dielectric losses are summed up into the two resistors R1a, R1b. The two LVTSCRs are in a high impedance state as long as no ESD pulse is discharged. Thus their junction capacitance (Cval) is the dominant characteristic to be considered here. Finally, this simple lumped circuit model is described by only four parameters: Lval, Rval, kval and Cval. After a swift manual tuning of those four values, a comparison of the tuned model and the measured frequency response of the CMF was done.

Both, differential and common mode response are calculated and plotted by their magnitude and phase in Fig. 7. While the phase is matched for frequencies up to around 1 GHz, the amplitude response starts to show significant deviation already at ≈ 600 MHz. Since this CMF is applied to an USB 3 signal with 5 GBit/s, it is obvious that this model is not suited to perform a proper signal integrity (SI) analysis in the time domain. A more complex, yet still comprehensible model is given in Fig. 8.

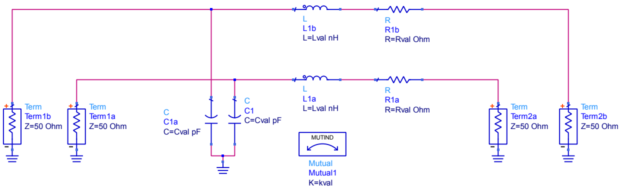


Fig. 6. Basic schematic diagram of common mode filter with ESD protection.

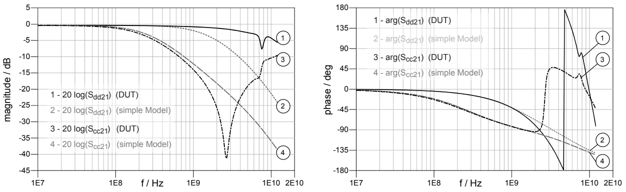


Fig. 7. Amplitude and phase response of CM and DM: Measured data of DUT vs. simulated response of simple lumped model.

The core of this model is a set of seven cascaded pairs of coupled coils. They are meant to mimic the 7.5 windings of the DUT. From the outer to the inner windings it is assumed that resistance will decrease and inductance will increase slightly. This is modelled by a factor c_1, applied from one segment to another. Each single inductor is characterised by its inductance ($\propto L_1$) and its resistance ($\propto R_1$). The magnetic coupling k_2 is assumed to be identical for all paired inductors. The total capacitance between upper and lower copper layer of the coils is described by six capacitors and two parameters: the fundamental capacitance C_1 and another scaling factor c_2. Inter-winding capaci-

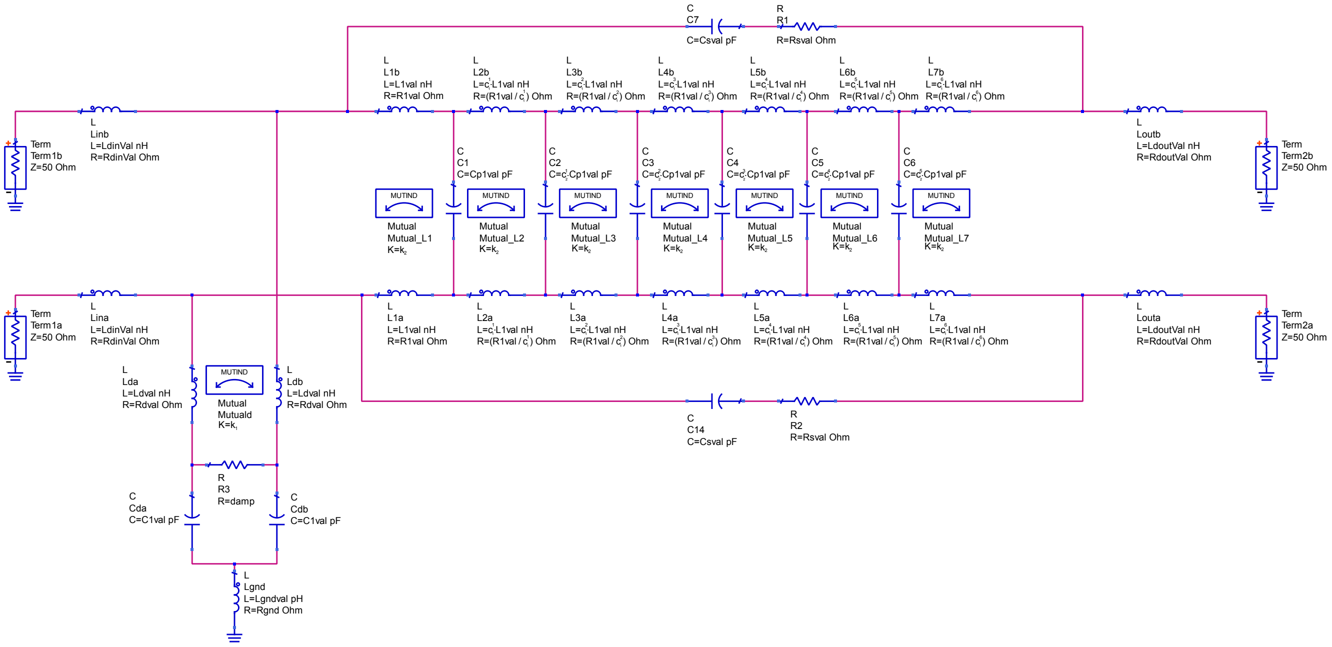


Fig. 8. Lumped element circuit diagram as used in the simulation and optimisation process.

tance is neglected. Cross-talk from the CMF input to the output is modelled by a series network C_s, R_s on each signal line. In general the impedance of each signal trace at the terminals is modelled by a series resistor and inductor. The values of these components (L_{in}, R_{in} and L_{out}, R_{out}) may differ for input and output since the LVTSCRs are only present at one side. For the junction capacitance of the LVTSCRs the capacitance C_d is used. The impedance of the inter-connect metal is modelled by L_d, R_d while the shunt resistor R_{ds} represents losses and is applied to tune the quality factor of the LC-circuit formed by L_c and C_d. Finally, the connection to the GND ball is defined by L_{gnd} and R_{gnd}. The challenge now is to determine the 19 component values so that the behaviour of the lumped model matches that of the measurements from the actual device. This is achieved in this work by computational optimisation.

MODEL OPTIMISATION

In computational optimisation a given error function f is driven towards an extreme value by a direct search algorithm in an iterative process (see Fig. 9). The error function maps an error value onto a given design vector $\boldsymbol{d}$. Here, this vector contains the 19 component values to be optimised (2).

The error e is then used by the optimisation algorithm to create new candidate solutions, i.e. design vectors.

The challenge in this research is to find an optimum design vector $\boldsymbol{d}'$ that produces a DM and CM frequency response as close as possible to the measured data of the DUT by minimising the error function (3).

$$\boldsymbol{d} = \{L_{in}, R_{in}, L_d, R_d, k_1, R_{ds}, \cdots, L_{out}, R_{out}\}^T \quad (2)$$

$$\boldsymbol{d}' = \arg\min f(\boldsymbol{d}) \quad (3)$$

The error function and the algorithm used in this work are described below.

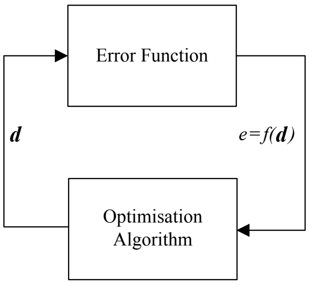


Fig. 9. Optimisation loop

Error function

For the problem at hand the error is defined as difference between a calculated and a measured frequency response (5)-(13). Such a difference can be calculated by comparing both graphs at arbitrary discrete frequencies. Since the given frequency responses exhibit a strong variation at rather high frequencies, the sampling rates have been chosen differently for low and high frequencies according to (4), see Fig. 10.

Since the interest is in both, differential mode and common mode response, two error terms e_{DM} and e_{CM} are combined using the weighted sum (5).

Both frequency responses are complex valued, i.e. amplitude and phase have to be optimised. With (6) and (10) again a weighted sum is used to combine these contributions to the error value.

$$f\left(i\right)=\begin{cases}10\,\mathrm{MHz}\cdot\left(10^{\frac{1}{10}}\right)^{i} & \forall\,0\leq i\leq 27, i\in\mathbb{N}\\ 5\,\mathrm{GHz}\cdot\left(10^{\frac{1}{92}}\right)^{i-27} & \forall\,28\leq i\leq 62, i\in\mathbb{N}\end{cases} \tag{4}$$

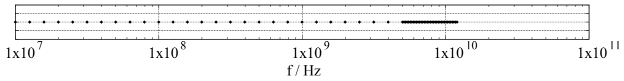


Fig. 10. Spacing of the 63 discrete frequency points for the error function.

$$e=\frac{1}{2}\cdot e_{DM}+\frac{1}{2}\cdot e_{CM} \tag{5}$$

$$e_{DM}=\sum_{i=0}^{62}10\cdot e_{DM}^{mag}\left(i\right)+e_{DM}^{phase}\left(i\right) \tag{6}$$

$$e_{DM}^{phase}\left(i\right)=\left|\arg\left(S_{dd,21}^{DUT}(f(i))\right)-\arg\left(S_{dd,21}^{opt}(f(i))\right)\right| \tag{7}$$

$$e_{DM}^{mag}\left(i\right)=\begin{cases}5\cdot\Delta_{DM}^{mag}\left(i\right)\text{ if } & i<54\wedge\Delta_{DM}^{mag}(i)\geq 0.1\\ 1\cdot\Delta_{DM}^{mag}\left(i\right)\text{ else}\end{cases} \tag{8}$$

$$\Delta_{DM}^{mag}\left(i\right)=\left|20\cdot\log\left(\left|\frac{S_{dd,21}^{DUT}(f(i))}{S_{dd,21}^{opt}(f(i))}\right|\right)\right| \tag{9}$$

$$e_{CM}=\sum_{i=0}^{62}10\cdot e_{CM}^{mag}\left(i\right)+e_{CM}^{phase}\left(i\right) \tag{10}$$

$$e_{CM}^{phase}\left(i\right)=\left|\arg\left(S_{cc,21}^{DUT}(f(i))\right)-\arg\left(S_{cc,21}^{opt}(f(i))\right)\right| \tag{11}$$

$$e_{CM}^{mag}\left(i\right)=\begin{cases}2\cdot\Delta_{CM}^{mag}\left(i\right)\text{ if } & i<8\\ 1\cdot\Delta_{CM}^{mag}\left(i\right)\text{ else}\end{cases} \tag{12}$$

$$\Delta_{CM}^{mag}\left(i\right)=\left|20\cdot\log\left(\left|\frac{S_{cc,21}^{DUT}(f(i))}{S_{cc,21}^{opt}(f(i))}\right|\right)\right| \tag{13}$$

Computational Optimisation

For engineering design optimisation, it is recommended to use a fixed number of decimal places (Nolle et al., 2016). Thus, the optimisation problem was transformed into a discrete optimisation problem. For this, a genetic algorithm (GA) (Goldberg, 1989) was used to minimise the error function (Fig. 9).

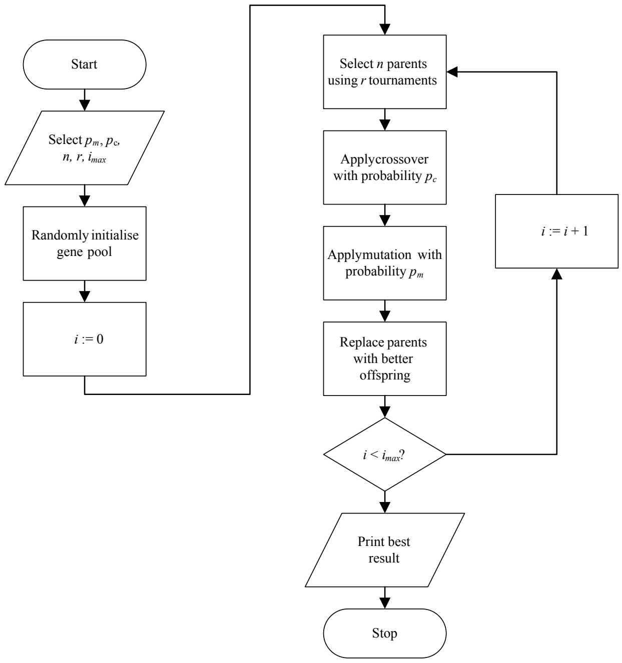
Genetic algorithms are discrete optimisation algorithms which simulate the evolutionary mechanism found in nature by using heredity and mutation. They were first introduced in 1975 by Holland (Holland, 1975) who also provided a theoretical framework for genetic algorithms, the Schemata Theorem.

For the optimisation problem under consideration, an integer coded genetic algorithm (Abbas, 2006) was used. Fig. 11 shows the chromosome structure employed.

L_{in}	R_{in}	L_d	R_d	...	L_s	C_s	L_{out}	R_{out}
1	2	3	4		16	17	18	19

Fig. 11. Chromosome structure used.

Each gene holds integer numbers from the range 1...2000 and is decoded into its phenotype by multiplication with a scaling factor and addition of an offset value in order to represent the fractional component values of the network. This results in a lower limit and an upper limit for the search space of each component value as listed in Table I.



Start
Select p_m, p_c, n, r, i_{max}
Randomly initialise gene pool
$i := 0$
Select n parents using r tournaments
Applycrossover with probability p_c
Applymutation with probability p_m
Replace parents with better offspring
$i < i_{max}$?
$i := i + 1$
Print best result
Stop

Fig. 12. Flowchart of the genetic algorithm used.

Fig. 12 shows a flowchart of the basic algorithm. After choosing the mutation probability p_m, the crossover probability p_c, the number n of individuals in the gene pool, the number r of tournaments used for the selection of one parent, and the maximum number of iterations i_{max}, the gene pool is randomly initialised. In each of the n generations, parents from the current generation are selected for the mating pool using tournament selection with r tournaments each (Miller and Goldberg, 1996). Then, uniform crossover (Syswerda, 1989) is used with probability p_c to produce offspring from the mating pool. After power mutation (Deep and Thakur, 2007) is applied to the offspring with the probability p_m, parents that perform worse than their offspring are replaced by their offspring. The algorithm terminates after i_{max} generations.

The next section provides the results of the experiments conducted.

TABLE I: Optimisation parameters and results of the second simulation set with $N = 25$ runs

Parameter	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Name	L_{in}	R_{in}	L_d	R_d	k_1	R_{ds}	C_d	L_{gnd}	R_{gnd}	L_1	R_1	k_2	c_1	C_p	c_2	R_s	C_s	L_{out}	R_{out}
unit	nH	Ohm	nH	Ohm	1	kOhm	pF	pH	Ohm	nH	Ohm	1	1	pF	1	Ohm	pF	nH	Ohm
Lower limit	0.5	0.0	1.0	2.0	-1.0	2.0	0.0	200	3.0	1.0	0.0	0.4	1.0	0.0	0.6	100	0.0	0.5	3.0
Upper limit	2.5	2.0	7.0	14.0	-0.4	4.0	2.0	600	5.0	7.0	2.0	1.0	1.8	2.0	1.0	500	2.0	2.5	9.0

EXPERIMENTS

For the experiments, the number of generations was set to $i_{max} = 20{,}000$ and the population size to $n = 1000$. The crossover probability was chosen to be $p_c = 0.6$ and the mutation probability was $p_m = 0.001$. Five individuals competed in each tournament. All these parameters were determined empirically and have been applied successfully in previous work (Werner and Nolle, 2016). With respect to the search space of the optimisation variables, two sets of experiments were conducted, each with 25 simulation runs.

The error values are normalized by the lower bound of the error values, i.e. 880. Fig. 13 shows a convergence plot for the average errors and the best solutions of each generation over 25 runs. It can be observed that most populations have converged after approximately 4000 generations.

After the first set of 25 runs it could be observed that 14 runs ended with an L_{out} value of 0 nH, which is exactly the lower limit of this optimisation parameter. The solutions related to this result are labelled group I. The other eleven solutions, denoted group II, resulted in 0.7 nH< L_{out} <1.3 nH. Results for both groups and the measured data are shown in Fig. 14. Even though the group I solutions yield a lower error value than those of group II they show a rather large deviation in the magnitude against the response of the DUT. Since the response of group II is more desirable and the value of L_{out} = 0 nH is physically unreasonable, the lower limit for this parameter was adjusted to 0.5 nH for the second set of 25 runs.

After the adjustment of the lower limit for L_{out} a clustering could still be observed; solutions belonging to one cluster converged towards the (new) lower limit whereas the others converged towards 1.15 nH. Therefore, the solutions were analysed separately for each cluster (Table II). Both clusters of this second set were labelled group I and II in the same way as it was done in the first set: Group I solutions are the ones that converged towards the lower limit of L_{out}, the remaining solutions belong to group II.

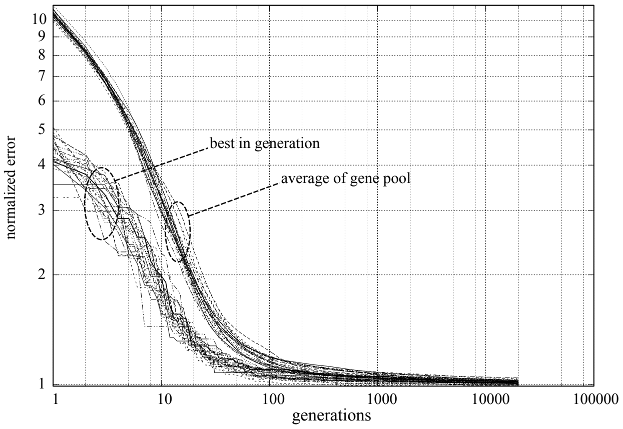


Fig. 13. Convergence plot of 25 optimisation runs. The error value is normalized by 880.

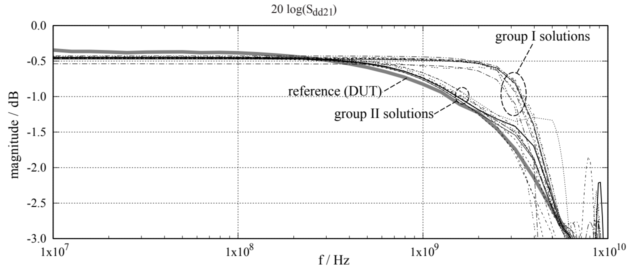


Fig. 14. Differential mode frequency response of DUT and results of first optimisation run. Two groups of local optima can be identified.

DISCUSSION

The best solution of group II (Table II) was used as overall solution and compared with the measurement data. Fig. 15 and Fig. 16 depict the frequency responses for both, the differential and the common mode, of the simulation respectively the measured data.

Fig. 15 shows that, for the differential mode, amplitude $|S_{dd,21}|$ and phase $\arg(S_{dd,21})$ are matched very well up to approximately 7 GHz. At the resonance frequency of 7.6 GHz, amplitude and phase show a certain deviation and above 9 GHz the model and measurements do not agree. For the common mode (Fig. 16) the amplitude response $|S_{cc,21}|$ is matched quite well in general with the exception of the resonance frequency around 2.5 GHz and frequencies above 5 GHz.

The phase $\arg(S_{cc,21})$ is matched accurately up to 600 MHz. At higher frequencies it still follows the measurements in principle, but with a higher error.

For both modes, the chosen model matches the measurements well in amplitude and phase over wide frequency ranges. This can be observed in more detail from Fig. 17, which shows the logarithmic magnitude of the complex differences. For frequencies below 7 GHz both differences are below -20 dB.

In order to analyse the complexity of this optimisation problem, a sensitivity analysis of the component values has been performed. In the optimum solution from group II (best (II) in Table II) each element of the optimum design vector $\boldsymbol{d'}$ is varied one at a time by a step of $(h \cdot d'_i)$ with $h = 1 \cdot 10^{-3}$. This allows to approximate the derivative of the error function with respect

TABLE II: Optimisation parameters and results of the second simulation set with $N = 25$ runs

Parameter	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Name	L_{in}	R_{in}	L_d	R_d	k_1	R_{ds}	C_d	L_{gnd}	R_{gnd}	L_1	R_1	k_2	c_1	C_p	c_2	R_s	C_s	L_{out}	R_{out}
unit	nH	Ohm	nH	Ohm	1	kOhm	pF	pH	Ohm	nH	Ohm	1	1	pF	1	Ohm	pF	nH	Ohm
Best (I)	0.96	1.01	3.10	2.01	-.86	2.24	0.07	494	3.06	4.27	0.03	0.87	1.06	0.19	0.86	233	0.05	0.50	3.10
Mean (I)	0.81	0.55	4.38	2.97	-.66	3.17	0.06	318	3.03	3.58	0.09	0.84	1.13	0.15	0.89	239	0.06	0.89	3.53
Std. dev. (I)	0.26	0.41	0.45	1.43	0.06	0.34	0.01	98	0.06	0.69	0.08	0.02	0.06	0.02	0.03	13	0.00	0.42	0.42
Best (II)	0.60	0.45	4.10	2.02	-.60	2.60	0.06	210	3.01	2.65	0.01	0.82	1.22	0.13	0.90	226	0.06	1.12	3.78
Mean (II)	0.60	0.48	4.45	2.24	-.63	3.04	0.06	312	3.02	2.76	0.07	0.82	1.20	0.13	0.91	227	0.06	1.15	3.59
Std. dev.(II)	0.08	0.28	0.32	0.48	0.04	0.27	0.00	125	0.02	0.17	0.05	0.00	0.02	0.01	0.02	5	0.00	0.16	0.31

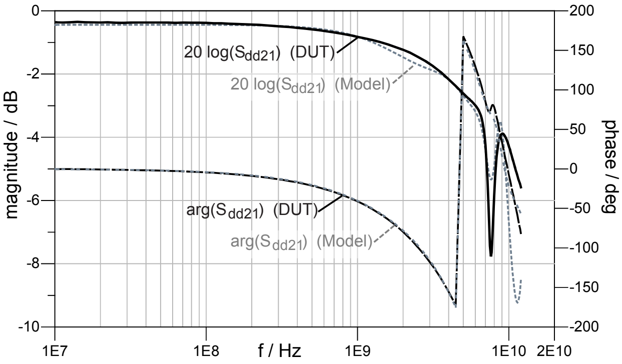


Fig. 15. Differential mode frequency response $S_{dd,21}$ of DUT and model.

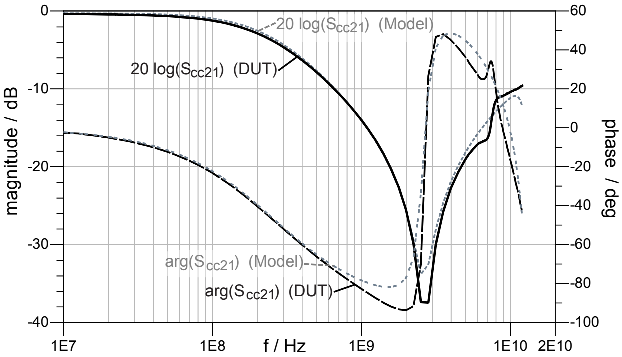


Fig. 16. Common mode frequency response $S_{cc,21}$ of DUT and model.

to the ith component of $\boldsymbol{d}$ (Bischof and Carle, 1998).

$$\left.\frac{\partial f(\boldsymbol{d})}{\partial d_i}\right|_{\boldsymbol{d}=\boldsymbol{d}'} \approx \frac{f(\boldsymbol{d}' + h \cdot d_i' \cdot \boldsymbol{e}_i) - f(\boldsymbol{d}')}{h \cdot d_i'} \tag{14}$$

Here, f is the error function and $\boldsymbol{e}_i$ is the ith Cartesian basis vector. The analysis is done for both error functions of the differential and the common mode.

Table III presents the sensitivities for the common mode (CM sens.) and the differential mode (DM sens.) for each component. It can be observed that some sensitivities are equal or close to zero, i.e with respect to the corresponding parameter, a (local) optimum has been reached, whereas others show a significant deviation from zero. This indicates that further improvements might be possible.

Another interesting aspect is that in the common mode there is no electric current through components 6, 14 and 15, meaning that their respective component values are irrelevant. Likewise in differential mode, no current is traversing through components 8 and 9, hence they have no effect on the resulting frequency responses. Both observations agree with the actual physics of the network.

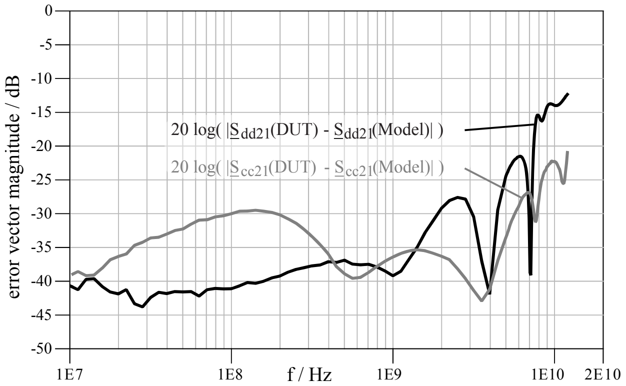


Fig. 17. Error vector magnitude for differential and common mode frequency response comparing DUT and model.

TABLE III: Sensitivity analysis of the component values of the optimum solution from group II.

Nr	Comp.	CM sens.	DM sens.
1	L_{in}	-355.5	290.9
2	Rin	-6.5	2.4
3	Ld	-0.1	350.9
4	Rd	3.4	-0.8
5	k1	-1.3	-987.3
6	Rds	0.0	-0.01
7	Cd	-2418.6	28330.8
8	Lgnd	0.0	0.0
9	Rgnd	6.7	0.0
10	L1	39.2	-264.2
11	R1	-30.1	1.9
12	k2	67.0	-353.7
13	c1	313.1	-324.4
14	Cp	0.0	-5829.7
15	c2	0.0	-281.3
16	Rs	-0.4	0.2
17	Cs	2041.2	-707.8
18	Lout	-89.1	-748.3
19	Rout	-7.6	0.9

CONCLUSION AND FUTURE WORK

In this work, the behaviour of a coupled coil device was modelled by a lumped element network. The advantage is that this model describes the device not only in the frequency domain, but also in the time domain. The latter is important for the signal integrity analysis of high-speed interfaces, like USB 3. The topology of the network was generated manually using expert knowledge. It should be noted that the modelling of a complex physical device by a simplified (lumped) network can never achieve an exact representation but only a close approximation. The challenge of this approach is that the component values of the network have to be determined.

The aim is to find a local optimum which produces a frequency response as close as possible to the DUT characteristic.

Since this optimisation problem is complex, a computational optimisation method, namely a genetic algorithm, was employed. The solution found produced good matching between the measurements and model responses for frequencies up to approximately 7 GHz. Whilst this agreement would be sufficient for many applications it is not adequate enough to be applied to interfaces with data rates of 5 Gbit/s and above. This is due to the fact that the fifths harmonic of the fundamental wave of the signal (5·2.5 GHz) is of importance.

The fact that the results were clustered into two groups indicates that the global optimum was not found. This assumption is support by the results from the sensitivity analysis presented here.

The next stage of this research will foucs on three different aspects: tuning the error function, adjusting the GA control parameters and testing alternative optimisation algorithms.

REFERENCES

Bluetooth & USB 3.0 - A guide to resolving your Bluetooth woes, 2016. URL `http://www.bluetoothandusb3.com/`. [2017-03-30].

H. M. Abbas. Accurate resolution of signals using integer-coded genetic algorithms. In *2006 IEEE International Conference on Evolutionary Computation*, pages 2888–2895, July 2006.

S. N. Ahmadyan, C. Gu, S. Natarajan, E. Chiprout, and S. Vasudevan. Fast eye diagram analysis for high-speed CMOS circuits. In *2015 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1377–1382, March 2015.

Christian Bischof and Alan Carle. Automatic differentiation principles, tools, and applications in sensitivity analysis. In *Second International Symposium on Sensitivity Analysis of Model Output*, pages 33–36, April 1998. ISBN 92-828-3498-0.

D. E. Bockelman and W. R. Eisenstadt. Combined differential and common-mode scattering parameters: theory and simulation. *IEEE Transactions on Microwave Theory and Techniques*, 43(7):1530–1539, July 1995. ISSN 0018-9480.

C. H. Chen, P. Davuluri, and D. H. Han. A novel measurement fixture for characterizing USB 3.0 radio frequency interference. In *Electromagnetic Compatibility (EMC), 2013 IEEE International Symposium on*, pages 768–772, Aug. 2013.

Kusum Deep and Manoj Thakur. A new mutation operator for real coded genetic algorithms. *Applied Mathematics and Computation*, 193(1):211–230, 2007.

W. Gao, L. Wan, S. Liu, L. Cao, D. Guidotti, J. Li, Z. Li, B. Li, Y. Zhou, F. Liu, Q. Wang, J. Song, H. Xiang, J. Zhou, X. Zhang, and F. Chen. Signal integrity design and validation for multi-GHz differential channels in SiP packaging system with eye diagram parameters. In *Electronic Packaging Technology High Density Packaging (ICEPT-HDP), 2010 11th International Conference on*, pages 607–611, Aug. 2010.

David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1989. ISBN 0201157675.

R. F. Harrington. *Field Computation by Moment Methods.* The Macmillan Company, 1968.

J.H. Holland. Adaptation in natural and artificial systems. 1975.

Intel. USB 3.0* Radio Frequency Interference Impact on 2.4 GHz Wireless Devices, 2012. URL `http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/usb3-frequency-interference-paper.pdf`. [2017-03-30].

Keysight. ADS - Advanced Design System, 2016. URL `http://www.keysight.com/find/eesof-ads`. [2017-03-30].

B. L. Miller and David E. Goldberg. Genetic algorithms, tournament selection, and the effect of noise. *Complex Systems*, (9):193–212, 1996.

Lars Nolle, R. Krause, and R. J. Cant. On practical automated engineering design. In K. Al-Begain and A. Bargiela, editors, *Seminal Contributions to Modelling and Simulation.* Springer, 2016.

David M. Pozar. *Microwave engineering.* J. Wiley & Sons, 4th edition, 2012. ISBN 978-0-470-63155-3.

N. Stevens and T. Dhaene. Generation of rational model based spice circuits for transient simulations. In *2008 12th IEEE Workshop on Signal Propagation on Interconnects*, pages 1–4, May 2008.

Gilbert Syswerda. Uniform crossover in genetic algorithms. In J. David Schaffer, editor, *ICGA*, pages 2–9. Morgan Kaufmann, 1989. ISBN 1-55860-066-3.

P. Triverio, S. Grivet-Talocia, M. S. Nakhla, F. G. Canavero, and R. Achar. Stability, causality, and passivity in electrical interconnect models. *IEEE Transactions on Advanced Packaging*, 30(4):795–808, Nov. 2007. ISSN 1521-3323.

J. Werner, J. Schütt, and G. Notermans. Sub-miniature common mode filter with integrated ESD protection. In *2015 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, pages 386–390, Aug. 2015.

J. Werner, J. Schütt, and G. Notermans. Common mode filter for USB 3 interfaces. In *2016 IEEE International Symposium on Electromagnetic Compatibility (EMC)*, pages 100–104, July 2016.

Jens Werner and Lars Nolle. Spice Model Generation from EM Simulation Data Using Integer Coded Genetic Algorithms. In Max Bramer and Miltos Petridis, editors, *2016 SGAI*, pages 355–367. Springer International Publishing, Dec. 2016.

AUTOMATIC BEAM HARDENING CORRECTION FOR CT RECONSTRUCTION

Marina Chukalina
FSRC "Crystallographyand Photonics"
RAS
IMT RAS
Moscow
Russia

Anastasia Ingacheva
National Research University Higher
School of Economics

Moscow 3
Russia

Alexey Buzmakov
FSRC "Crystallographyand Photonics"
RAS
Russian Academy of Science
Moscow
Russia

Igor Polyakov
Institute for Information Transmission
Problems
RAS
Moscow
Russia

Andrey Gladkov
Visillect Service

Moscow
Russia

Ivan Yakimchuk
Schlumberger Moscow
Research Center
Moscow
Russia

Dmitry Nikolaev
Institute for Information
Transmission Problems RAS
Moscow
Russia

KEYWORDS

Computer tomography, beam hardening correction, Radon invariant.

ABSTRACT

In computed tomography (CT) the quality of reconstructed images depends on several reasons: the quality of the set-up calibration; precision of the mathematical model, used in the reconstruction procedure; the reconstruction technique used; quality of the numerical implementation of the reconstruction algorithm. Most of fast reconstruction algorithms use the X-ray monochromaticity assumption. If we apply the algorithms to reconstruct the images from the projections measured in polychromatic mode then the reconstructed images will be corrupted by so-called cupping artifacts due to beam hardening. There are several approaches to take into account polichromaticity. One of them is to correct the sinograms before reconstruction. This paper presents the usage of Radon invariant to estimate the gamma parameter in the procedure of the sinogram correction for beam hardening (BH).

INTRODUCTION

Today the existing methods that take into account polychromaticity effect can be jointed in several groups. The methods with usage of hardware filtering (Jennings 1988) can be classified as the first group. The second one includes so-called pre-processing techniques (Herman 1979). The next proposes dual energy correction (Kyriakou et al. 2010, Yu et al. 2012). The forth group of methods relates to an iterative correction in reconstruction stage (De Man et al. 2001, Elbarki and Fessler 2002, Menvielle et al. 2005). The last group works with artifacts in-line (Suk Park et al. 2016) or post-processing of the reconstructed images. This work aims to optimize the procedure of CT sinogram correction for beam hardening (pre-processing case). Due to the complex task of CT, it is hard to determine the optimal measurement settings (X-ray energy (spectrum), filters, etc.) for a given experiment. Different operators might apply different measurement strategies. Methods that reduce the dependence of experiment results on the applied machine settings are preferable. This work aims to optimize the procedure of CT sinogram correction for beam hardening. The presence of polychromaticity in the X-ray probe generates discrepancy between sinograms calculated from the projection data (detector data) and the Radon transform. The necessity of the compensation for the linearization of the detector data was pointed out in 1975 (McCollough 1975, Brooks et al. 1976). The quality of the CT reconstruction dramatically depends on the level of compensation. The CT reconstruction procedure is time consuming. Choosing the level of compensation based on the reconstruction results evaluation is expensive. We propose to use the concept of Radon invariant for evaluation of the quality of compensation. We give the grounds for approach taken. We suggest a new criterion for automatically choosing the level of compensation. We discuss the results obtained for simulated and real CT data. In conclusion

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

we discuss the results of applying the quantitative criteria for cupping effect pronouncement (Nikolaev et al. 2016). The cupping artifacts corrupt the reconstructed image if the CT polychromatic data used in the reconstruction procedure were undercompensated.

LINK BETWEEN SINGLE-PARAMETER CORRECTION AND RADON INVARIANT

Fast reconstruction algorithms assume that the attenuation of the incident X-ray beam is exponentially related to the thickness of the object due to Beer's law. It becomes incorrect for polychromatic X-ray sources. In this case lower energy photons are more attenuated than higher ones (hard X-rays) when the beam passes through the object. Hence the attenuation produced by homogeneous sample, defined as the negative logarithm of the ratio of transmitted to incident beam, is not strictly proportional to its thickness. The measured nonlinear relationship can be fitted (Herman 1979). The process results to reconstructed image distortion called beam hardening due to the underlying phenomena. To state the problem of beam hardening correction let's write the mathematical model of the signal formation (point in the projection space) for polychromatic case assuming that object under study is unicomponent, but its distribution inside the volume is nonuniform:

$$I=\int dEI_0(E)\nu(E)exp(-\int_L\mu(E)C(l)dl). \quad (1)$$

Here I is a point in the projection space or value measured by a detector pixel. $I_0(E)$ is original intensity of the spectrum component with energy E. Function $\nu(E)$ describes the detector response at energy E. Direction to the detector pixel is determined by $\int_L$. $\mu(E)$ is linear attenuation coefficient of the component at energy E, $C(l)$ assigns its concentration along L. Assuming that $I_0(E)$, $\nu(E)$ and $\mu(E)$ are known from the reference measurements, we can rewrite

$$I=f(\int C(l)dl)=f(R(C)), \quad (2)$$

where R is Radon transform. Then the general statement of the problem is to find f^{-1}

$$R(C)=f^{-1}(I). \quad (3)$$

For practical cases we propose to use the following considerations. If the reconstructed image is described by function H, then the reconstruction problem is

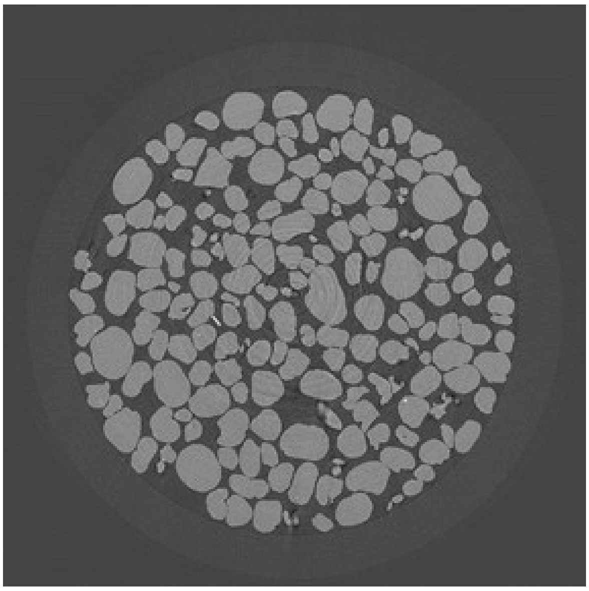
$$H=R^{-1}(I^*), \quad (4)$$

Where the following approximation is often used to correct the beam hardening:

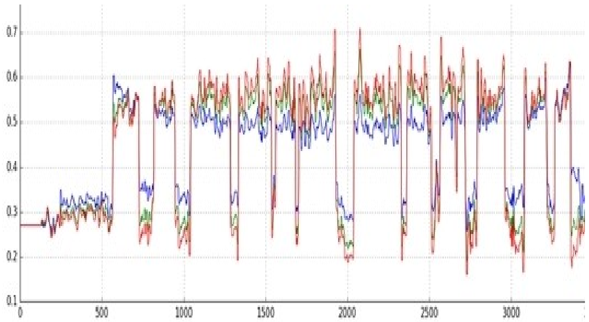
$$I^*=(ln(\int dEI_0(E)\nu(E)/I))^\gamma \quad (5)$$

As the sample does not cover the view field of the detector completely (for example, see Figure 1), the value $\int dEI_0(E)\nu(E)$ is estimated in the pixels free from the object.

We used CT real projection data from test sample placed in vial to demonstrate the influence of different values γ on reconstruction result (Figures 1, 2). FDK reconstruction algorithm (Feldkamp et al, 1984) was applied. Visual control for finding an optimal value of the parameter γ can be rather complicated. In Figure 2 presence of the cupping effect (blue line) informs us about undercompensating for the linearization. We propose the following algorithm to find the optimal γ value. Let's calculate the sum I^* values over all pixels for each CT projection. We have several realizations of random value. One value SUM^ϕ for one projection angle ϕ. Its mean $M(SUM^\phi)$ is close to the Radon invariant. Origin of concept of Radon invariant is monochromatic measurement case. As we try to compensate polychromatic CT data by (5) to use further the reconstruction algorithms developed for monochromatic CT data, we should minimize total error distance. To estimate the optimal value γ (γ∈(1, 2)) we minimize the root-mean-square deviation of the random value $SUM\gamma^\phi$ from $M(SUM\gamma^\phi)$ divided by the mean $M(SUM\gamma^\phi)$ (NRMSD).



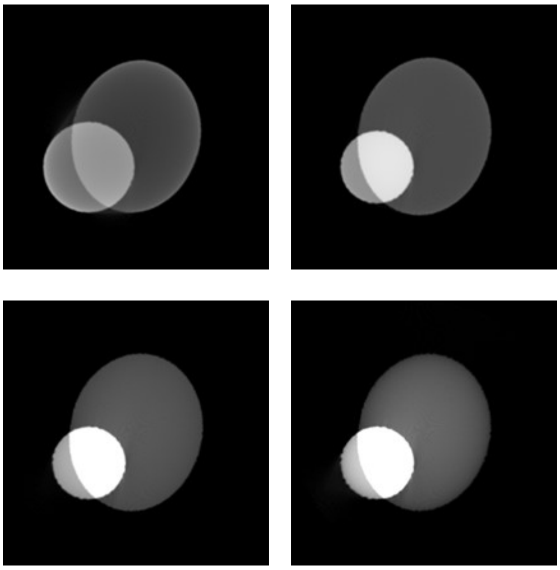
Figures 1: Reconstruction result for γ=1.6.



Figures 2: Reconstruction result (cross-section) with different γ values.

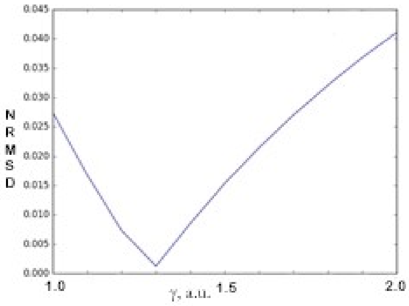
RESULTS FOR SIMULATED DATA

To test the behavior of the algorithm we used the simulation data. Concept of the geometry of the phantom used to calculate the polychromatic sinograms becomes clear if we refer to Figure 3. Image size is 208×208 pixels. The reconstruction results for different γ values are presented. We calculated the monochromatic CT projections in fun geometry. The value used to simulate the polychromaticity is 1.3.



Figures 3: Reconstruction results for the simulated data with different γ:1; 1.3; 1.6; 1.9.

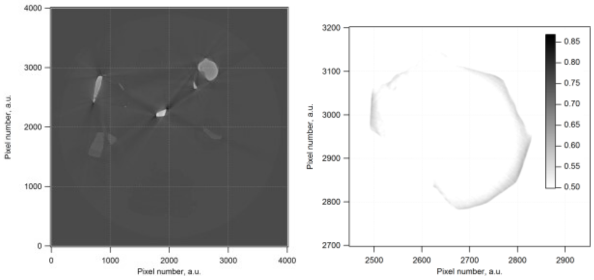
Figure 4 presents the NRMSD criterion behavior for the simulated CT projections. As one can see the minimum corresponds to the proper γ=1.3 value.



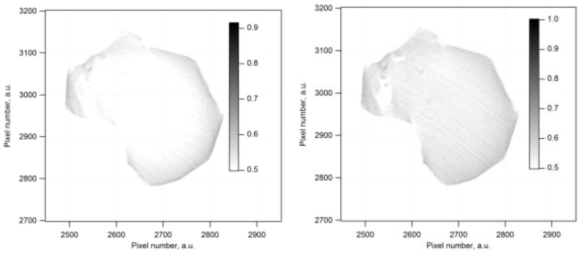
Figures 4: NRMSD dynamics for simulated CT projections.

RESULTS FOR REAL CT DATA

We present the results for real CT data below. The test sample was measured with CT scanner. The projections were collected at 2030 rotation angles. The projection size is 2096×4000. The reconstruction results with different γ values are presented in Figure 5.



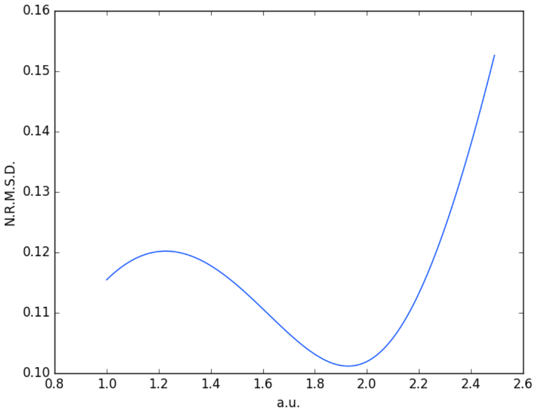
a) b)



c) d)

Figures 5: Reconstruction results for CT data. a) Full cross-section, γ=1. b) Part of the inverted image γ=1. c) Part of the inverted image γ=1.3. d) Part of the inverted image γ=1.6.

Figure 6 presents the criterion behavior for real CT projections.

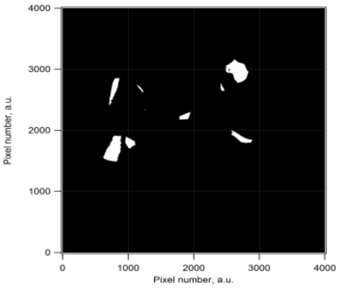


Figures 6: NRMSD dynamics for real CT projections.

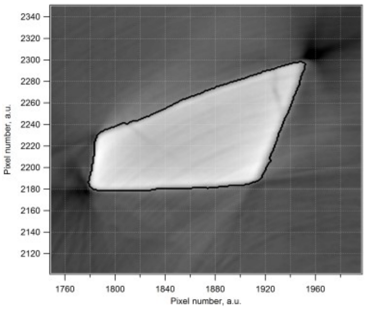
We can see well pronounced minimum on the curve. This confirms the correct procedure used for construction of the criterion

CUPPING ARTIFACT AND BEAM HARDENING COMPENSATION

Beam hardening correction should decrease the cupping artifact in the reconstructed images. Earlier we suggested the criteria to estimate the cupping artifact severity (Nikolaev et al. 2016). As cupping artifacts occur only inside dense objects, we used a mask designed for the object. To estimate the value of the artifact, we calculate the morphometric parameter CA over selected region corresponding to the mask.



a)



b)

Figures 7: a) Mask. b) Central mask object. Reconstruction with γ=1.

Suppose we have a mask *M* with *K* homogeneous objects as illustrated in Figure 7. The reconstruction result corresponding to the central *i*th object of the mask with mask boundary is presented in Figure 7b. The distance transform (DT^1, ….DT^k) is calculated for each object of the mask based on the Euclidian metric. If the beam hardening fully compensated, the pixels inside the homogeneous object should have the same values. Let I_j^k be a value of the *j*-th pixel of M^k mask area. We calculate a base value as the mean value within a stable region

$$BaseV^k = mean\{I_j^k | DT_j^k > Val^k\}, \qquad (6)$$

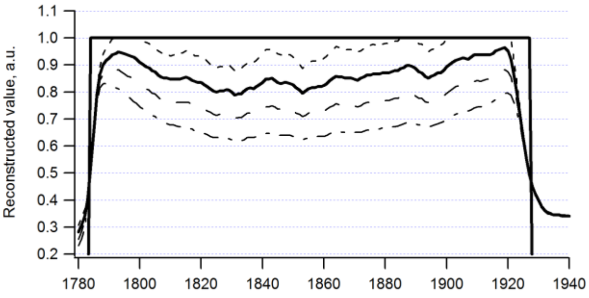
where Val^k=0.9 $max(DT^k)$. We calculate the distance transform histogram of M^k as

$$CE^k = \sum_{v=1}^{0.2max(DTk)} |mean\{I_j^k | v\text{-}1 < DT_j^k < v\} - BaseV^k \quad (7)$$

The average value over all objects is the areal morphometric parameter CA that represents the measure of the cupping artifact severity, i.e.

$$CA = 1/\mathrm{k}/\sum_{k=1}^{K} CE^k \qquad (8)$$

The reconstruction results for different γ values are presented in Figure 8 to illustrate the execution of algorithm. As the criteria takes into account all masked parts of the object under investigation, detailed analysis of the criteria behavior will be presented in the talk.



Figures 8: Cross-sections for different γ values: 1, 1.3, 1.6 and 1.9 (down-up) in accordance with Figure 7b (horizontal slice 2220).

CONCLUSION

The problem we are solving now, from an engineering standpoint, is the task of blind. This kind of technique, we think, should be used to solve most ill-posed inverse problems, since the quality of reconstruction is significantly dependent on calibration accuracy (Chukalina et al. 2016). For example, the problems, similar to tomographic problems, arise when we try to reconstruct the environment according to (from) the sonar sinals (Svets et al. 2016), where the parameters of the sonar significantly affect the signal model. The level of the compensation for the beam hardening is only one from the calibratable CT parameters.This work aims to optimize the procedure of CT sinogram correction for beam hardening. We propose to use the concept of Radon invariant for evaluation of the quality of correction. We give the grounds for approach taken. We suggest a new criterion for automatic selection the level of correction. We discuss the results for simulated and real CT data packages. As an uncorrected sinogram produces the images with well pronounced cupping effect we check the quantitative criteria for the cupping effect pronouncement to estimate the correction quality.

The research leading to results on the development of mathematical models of distortions arising in tomographic systems with polychromatic probe has funded from the Russian Science Foundation Grant (project #14-50-00150).

REFERENCES

Brooks, R.A. and G. Di Chiro. 1976. “Beam hardening in X-ray reconstructive Tomography.” *Phys. Med. Biol* 21, No.3, 390-398.

Chukalina, M., D. Nikolaev, A. Buzmakov, A. Ingacheva, D. Zolotov, A. Gladkov, V. Prun, B. Roshin, I. Shelokov, V. Gulimova, S. Saveliev and V. Asadchikov. 2016. “The Error Formation in the Computed Tomography: from a Sinogram to the Results Interpretation.” *Russian Foundation for Basic Research J.*, No 4(92), 73-83.

De Man, B.; Nuyts J.; P. Dupont; G. Marchal and P. Suetens. 2001. “An iterative maximum-likelyhood polychromatic algorithm for CT.” *IEEE Trans. Med. Imag*. 20, No.10, 999-1008.

Elbarki I.A. and J.A. Fessler. 2002. “Statistical image reconstruction for polyenergetic X-ray computed tomography.” IEEE Trans. Med. Imag. 21, No.2, 89-99.

Feldkamp, L. A., L. Davis, and J. Kress. 1984. “Practical Cone-beam Algorithm. ” J. of the Optical Society of America 1, 612–619.

Herman, G.T. 1979. “Correction for beam hardening in Computer Tomography.” *Phys. Med. Biol* 24, No.1, 81-106.

Jennings, R.J. 1988. “A method for comparing beam hardening ¯lter materials for diagnostic radiology.” *Medical Physics* 15, No 4, 588-599.

Kyriakou Y.; E. Meyer; D. Prell and M. Kachelriess. 2010. “Emperical beam hardening correction (EBHC) for CT.” *Med. Phys*. 37, 5179-5187.

McCullough E.C. 1975. “Photon attenuation in computed tomography.” *Med. Phys*. 2, 6, 307-320.

Menvill N.; Y. Goussard; D. Orban and G. Soulez. 2005. “Reduction of beam-hardening artifacts in X-ray CT.” in *Proc. 27th Annu. Int. Conf. IEEE EMBS*, 1865-1868.

Nikolaev, D.P.; A. Buzmakov; M. Chukalina; I. Yakimchuk; A. Gladkov and A. Ingacheva. 2016. “CT Image Quality Assessment based on Morphometric Analysis of Artifacts”. In *Proceedings of the International Conference on Robotics and Machine Vision (ICRMV 2016)* (Moscow, Russia, Sept.14-16). Proc. SPIE 10253, 2016 International Conference on Robotics and Machine Vision, 102530B (February 8, 2017); doi:10.1117/12.2266268; http://dx.doi.org/10.1117/12.2266268.

Shvets E., D. Shepelev and D. Nikolaev. 2016. “Occupancy grid mapping with the use of a forward sonar model by gradient descent.” J. of Comm. Technology and Electronics, V. 61, №12, pp. 1474-1480.

Suk Park, H.; D. Hwang and J. K. Seo. 2016. “Metal artifacts Reduction for Polychromatic X-ray CT Based on a Beam-Hardening Corrector.” *IEEE Trans. Med. Imag*. 35, No.2, 480-487.

Yu, L.; S. Leng and C.H. McCollough. 2012. “Dual-energy CT-based monochromatic imaging.” *Am. J. Roentgenol*. 199, S9-S15.

AUTHOR BIOGRAPHIES

MARINA CHUKALINA was born in Cheboksary, Russia and went to the Moscow Physical Engineering Institute, where she studied physics and mathematics. She received her PhD in physics at the Institute of Microelectronics Technology and High Purity Materials RAS where she has been working since 1988. From 2013 she is Vice-head of Cognitive Technologies Department of Faculty of Innovation and Higher Technology of Moscow Institute of Physics and Technology (State University). Her interests include the development of signal and image processing tools for X-ray Microscopy and Tomography. Her e-mail address is: chukalinamarina@gmail.com. Her Web-page can be found at http://tomo.smartengines.biz/peoples.en.html.

ANASTASIA INGACHEVA was born in Kazan, Russia and went to Kazan (Volga region) Federal University, where she studied Economic Cybernetics and obtained her degree in 2012. She obtained master's degree of Higher School of Economics (National Research University) at computer science faculty in 2015. Now Anastasia is a PhD student in the same place. Since March 2013 Anastasia has been worked in the X-ray reflectivity laboratory of the Shubnikov Institute of Crystallography RAS. Her research activities are in the areas of poly- and monochromatic X-Ray computed tomography and analysis of obtained CT data. Her e-mail address is: ingacheva@gmail.com.

ALEXEY BUZMAKOV was born in Kirs, Russia. He studied physics, obtained his Master degree in 2006 and Ph.D. degree in 2009 from Lomonosov Moscow State University. Now he work in the Institute of Crystallography Federal Research Scientific Centre "Crystallography an Photonics" RAS in a research group in the field of x-ray tomography and x-ray optics. His e-mail address is: buzmakov@gmail.com and his web-page can be found http://tomo.smartengines.biz/peoples.en.html.

IGOR POLYAKOV was born in Moscow, Russia. He studied mathematics and mechanics, obtained his Master degree in 2008 and Ph.D. degree in 2012 from Moscow State University. Since 2016 he is a researcher of the Vision Systems Lab. at the Institute for Information Transmission Problems, RAS. His research activities are in the area of computer vision with primary application to image classification. His e-mail address is igorp86@mail.ru.

ANGREY GLADKOV was born in Stavropol, Russia. He studied automatic control and informatics at the Moscow State Institute of Radio Engineering, Electronics and Automation and obtained his Master’s degree in 2012. Now Andrey is a PhD student at the Institute for Information Transmission Problems, where he studies and works. His research activity is focused on image processing in X-ray tomography and radiography. His e-mail address is: gladkov.ap@iitp.ru.

IVAN YAKIMCHUK was born in Protvino, Russia and went to the Lomonosov Moscow State University, where he obtained his Master degree in general physics in 2009 and received PhD degree in X-ray physics in 2012. Since then he worked as a research scientist in the group of Imaging and Image Processing at the Schlumberger Moscow Research Center. Now he is leading this group making the focus on X-ray micromography and Electron Microscopy of rock samples with corresponding digital processing of results including image segmentation, morphometry, pattern recognition, spatial registration, etc. His e-mail address is: IYakimchuk@slb.com.

DMITRY NIKOLAEV was born in Moscow, Russia. He studied physics, obtained his Master degree in 2000 and

Ph.D. degree in 2004 from Moscow State University. Since 2007 he is a head of the Vision Systems Lab. at the Institute for Information Transmission Problems, RAS. His research activities are in the areas of computer vision with primary application to color image understanding. His e-mail address is dimonstr@iitp.ru.

AN INTELLIGENT WINCH PROTOTYPING TOOL

Robin T. Bye[*], Ottar L. Osen[*,†] Webjørn Rekdalsbakken[*],
Birger Skogeng Pedersen[**] and Ibrahim A. Hameed[*]
[*] Software and Intelligent Control Engineering Laboratory
[*] Department of ICT and Natural Sciences
[**] Mechatronics Laboratory
[**] Department of Ocean Operations and Civil Engineering
[*,**] NTNU, Norwegian University of Science and Technology
[*,**] Postboks 1517, NO-6025 Ålesund, Norway
[†] ICD Software AS
[†] Hundsværgata 8, NO-6008 Ålesund, Norway

KEYWORDS

Virtual prototyping; Product optimisation; Computer-automated design; Maritime winch design; Genetic algorithm; Artificial intelligence.

ABSTRACT

In this paper we present a recently developed intelligent winch prototyping tool for optimising the design of maritime winches, continuing our recent line of work using artificial intelligence for intelligent computer-automated design of offshore cranes. The tool consists of three main components: (i) a winch calculator for determining key performance indicators for a given winch design; (ii) a genetic algorithm that interrogates the winch calculator to optimise a chosen set of design parameters; and (iii) a web graphical user interface connected with (i) and (ii) such that winch designers can use it to manually design new winches or optimise the design by the click of a button. We demonstrate the feasibility of our work by a case study in which we improve the torque profiles of a default winch design by means of optimisation. Extending our generic and modular software framework for intelligent product optimisation, the winch calculator can easily be interfaced to external product optimisation clients by means of the HTTP and WebSocket protocols and a standardised JSON data format. In an accompanying paper submitted concurrently to this conference, we present one such client developed in Matlab that incorporates a variety of intelligent algorithms for the optimisation of maritime winch design.

INTRODUCTION

NTNU in Ålesund is located on the west coast of Norway in the heart of the Global Centre of Expertise (GCE) Blue Maritime Cluster.[1] This industrial cluster is a world leader in design, construction, equipment and operation of advanced special vessels for the global ocean industry, with an annual turnover of about 62 billion NOK (GCE Blue Maritime Cluster, 2016). In close cooperation with the maritime industry, NTNU in Ålesund offers courses on 3D modelling, visualisation and VP, training of maritime personnel in advanced simulators, and takes part in research projects. Together with two companies in the maritime cluster, ICD Software AS[2] (provider of industrial control systems software) and Seaonics AS[3] (designer and manufacturer of offshore equipment), the Software and Intelligent Control Engineering (SoftICE) Laboratory[4] has received funding from the Research Council of Norway and its Programme for Regional R&D and Innovation (VRI) for two independent but related research projects for using artificial intelligence (AI) for intelligent computer-automated design (CautoD) of offshore cranes and winches, respectively. Our main focus is on the development of a generic and modular software framework for intelligent CautoD of maritime products, exemplified by offshore cranes and winches. We have previously presented the software framework with respect to the design of cranes (Bye, Osen, Pedersen, Hameed and Schaathun, 2016) and how various intelligent algorithms can be applied to optimise the design (Hameed, Bye, Osen, Pedersen and Schaathun, 2016; Hameed, Bye and Osen, 2016*a*,*b*).

In this paper, we extend this framework with the inclusion of new product calculator for maritime winches that together with a GA optimisation module and a web GUI constitute what we refer to as a winch prototyping tool (WPT). Submitted concurrently in an accompanying paper, we present a Matlab winch optimisation client (MWOC) that we use to test a number of algorithms within this same framework (Hameed et al., 2017). (Hameed, Bye, Pedersen and Osen, 2017). Whilst we have achieved the goals of the specific two research projects mentioned above, we wish to emphasise that our work easily can be extended to other products and CautoD methodologies.

In the following, we begin with a background overview of virtual prototyping (VP) in general and CautoD in

Corresponding author: Robin T. Bye, robin.t.bye@ntnu.no.

[1] http://www.bluemaritimecluster.no

[2] http://www.icdsoftware.no

[3] http://www.seaonics.com

[4] http://blog.hials.no/softice

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

particular, VP of maritime winches, and the motivation for our work. Next, we outline the method we have used, including details about the software architecture of our product optimisation system and its main components, and the new intelligent WPT. Finally, we present a case study where we use the WPT to optimise a given winch design and discuss our work and potential future directions.

BACKGROUND

Virtual Prototyping (VP)

VP may be defined as the computer-aided construction of digital product models, usually virtual prototypes or digital mockups, and realistic graphical simulations for the purpose of design and functionality analyses in the early stages of the product development process (Pratt, 1995). Common VP methodologies include computer-aided design (CAD), realistic virtual environments (VEs), VR, and CautoD, with modelling, simulation, and visualisation as key underlying themes. In our work, the main focus is on applying AI methods such as genetic algorithms (GAs), simulated annealing (SA), particle swarm optimisation (PSO), and grey wolf optimisation (GWO) for CautoD in order to automate and optimise the design phase of product development.

Computer-Automated Design (CautoD)

CautoD traces back at least to the 1960s, when Kamentsky and Liu (1963) created a computer programme for determining suitable logic circuits satisfying certain hardware constraints while at the same time evaluating the ability of the logics to perform character recognition. Since then, there have been many contributions of CautoD, particularly in the field of structural engineering (see Hare et al., 2013, for a survey).

The general paradigm of CautoD is that of *optimisation*, where one formulates the design problem as the optimisation of an objective function. The objective function is either a cost function that must be minimised, or a fitness function that must be maximised. Parameterising the design, the goal is to find suitable values for the design parameters such that the objective function is optimised.

Whilst some optimisation problems can be formulated such that analytical or exact solutions can be found, more complex optimisation problems, including non-deterministic polynomial time (NP) problems, may require heuristic or intelligent methods from the field of AI, such as machine learning and evolutionary computation, to find satisfactory solutions (see Zhang et al., 2011, for a survey).

Virtual Prototyping of Maritime Winch Systems

Figure 1 shows a winch system in the Seaonics Big Drum Trawlwinch series, which is one of several kinds of maritime winch systems offered by Seaonics AS. In addition to trawling, maritime winches are used for anchor handling, mooring, towing, and more. The winch may at first sight appear insignificant and be conceived as a taken-for-granted piece of machinery, however, winches are indispensable for many tasks, including the precise monitoring of various operating conditions (e.g., cable payout length, speed, and tension), active motion compensation, integrated cable cleaning systems, remote control, and as a computer interface (Pearlman et al., 2017). Thus, maritime winch systems are typically complex, come in many flavours, and consist of many different parts and components. With the advent of new technologies, major improvements in the drive systems, cable handling, safety and reliability are possible, particularly in motor and hydraulic controls (Pearlman et al., 2017). Examples of recent relevant research include model-based control designs for offshore hydraulic winch systems (Skjong and Pedersen, 2016), the influence of fishing grounds on trawler winch design (e.g., see Carral et al., 2015, for a review), winch design interventions for safety and entanglement hazard prevention (Lincoln et al., 2016), analysis of trawl winches barrels deformations (Solovyov and Cherniavsky, 2013), and high performance winch and synthetic rope systems for workboats, tug boats, and commercial marine applications (Griffin, 2004), to mention some.

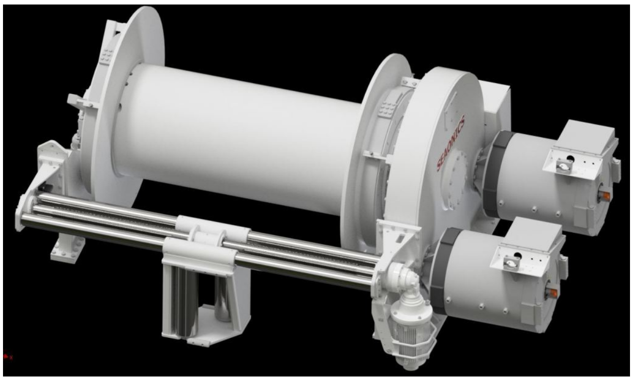


Figure 1: Seaonics Big Drum Trawlwinch PM, designed for trawling on deep water and rough bottoms, in arctic conditions. The winch is delivered with permanent magnet (PM) or conventional AC motors for both demersal and pelagic trawling. Image courtesy of Seaonics AS.

In the work we present here, we focus on winches with two kinds of motors, namely electric and hydraulic. Together with the drum and the wire, these four components and their properties yield a number of design parameters that must be appropriately chosen by the designer to achieve a winch design with desired measures of performance, or key performance indicators (KPIs). As noted in a review by Pearlman et al. (2017), many of these parameters are dependent on each other and the winch designer must apply an iterative process to obtain a satisfactory design.

In this paper, we are mainly concerned with torque performance but emphasise the many other concerns must be taken into account by the winch designer, including adhering to laws, regulations, and the use of design codes such as the standards provided by classification socities like DNV GL, Lloyd's Register Group Limited, and the American Bureau of Shipping.

Motivation

Designing an optimal winch requires deep knowledge about its intended application. For example, an optimal winch for trawling will not be optimal for heave-compensated cranes, since heave compensation will operate in a sinusoidal mode around a working area whereas trawling will require high capacity for bringing the catch on-board in a continuous operation. Also, within any one application there are usually many conflicting requirements. For instance, for trawling it is important to set the net quickly, which requires high wire velocity and a winch drum with a large inner diameter. However, when the net is full of fish, one needs high torque, which requires lower wire velocity and a smaller drum diameter.

Moreover, the design process traditionally has involved rather complicated spreadsheets that are difficult to use and maintain and have very limited visualisation features. In addition, in order to improve the versatility of the winches and enhance their performance, it has become popular to design hybrid winches that employ both electrical and hydraulic motors, combining the advantages of both kinds of motors. The merit of hybrid winches comes at a cost though, as the design process become even more complex, and even with a fully functional spreadsheet, the most difficult part remains, namely finding the optimal parameters. Due to the large number of parameters, the task of improving the design through trial and error is very time consuming and difficult. Hence, it is a difficult task to engineer a winch with desired specifications due to the large number of possibly conflicting design parameters and the lack of suitable optimisation tools for problems that may be NP-hard in nature.

In the next sections we present our WPT that has support for hybrid winches and with built-in support for “automagic” parameter optimisation. Unlike many other automatic parameter optimisation tools, this tool also support selection from predefined components, such as a catalogue of commercially available motors. Hence, designers are free to limit parameters to an interval or to a predefined set of components. Another rather unique feature is that designers can limit the scope of the component library. For example, the designers may choose to let predefined component sets such as pairs of motors and gears be lumped together as bigger units, or they may choose to have the software search for motors and gears independently of each other.

METHOD

Product Optimisation System

The diagram in Figure 2 shows a high-level overview of the software architecture of our product optimisation system. The system employs a server-client software architecture. The main component of the server is a product calculator, e.g., for offshore cranes or winches, that contains a number of different product design parameters, of which many are interdependent through electrical, hydraulic, and mechanical interactions in a highly complex, and often nonlinear, manner. Different parameter values constitute different designs of the same product. When new parameter values are set, the product calculator calculates a number of KPIs. The goal of the product designer is to determine the parameter values that yields a product design with desirable KPIs. Using a client, the product designer can manually set the parameter values in the product calculator via two different communication interfaces: the Hypertext Transfer Protocol (HTTP) or the WebSocket (WS) protocol. In return, the client can obtain the values of the KPIs, as well as other measures of interest, calculated by the product calculator. These bidirectional messages are transferred as JavaScript Object Notation (JSON), which is a lightweight human-readable data-interchange format.

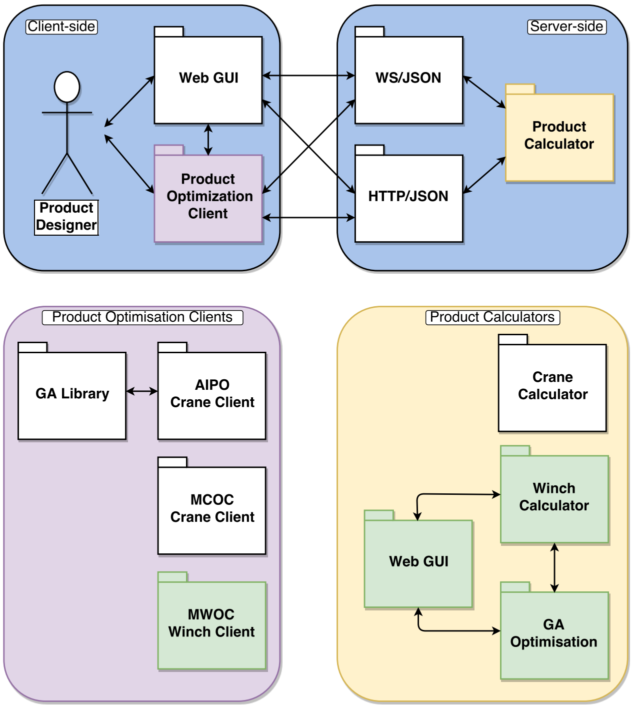


Figure 2: Software architecture for intelligent CautoD of offshore cranes, winches, or other products. Green boxes indicate work not presented previously.

Determining a suitable product design by manual trial-and-error is a tedious task for the product designer. Instead, one can opt to use a product optimisation client (POC) that automates this process. In addition, it may be beneficial to use a graphical user interface (GUI) both for interacting with optimisation software and with the server-side product calculator.

The design of this software framework is generic and modular. On the server side, we can develop new product calculators as long as they conform to the HTTP/JSON or WS/JSON communication interfaces and message formats that we have defined. Likewise, on the client side, users can develop GUIs and POCs for different products as needed, again as long as they conform to said communication interfaces and message formats.

Recently, we have experimented with various client solutions and developed both a GUI and several POCs for optimisation of offshore cranes, including the Artificial Intelligence for Product Optimisation (AIPO) client written in Haskell that uses a GA for the optimisation (Bye et al., 2016), as well as the Matlab Crane Optimisation Client (MCOC) that uses several evolutionary algorithms for the optimisation, including the GA, SA, PSO, and GWO algorithms (Hameed, Bye, Osen, Pedersen and Schaathun, 2016; Hameed, Bye and Osen, 2016*a*,*b*).

In the following sections, we present our new intelligent winch prototyping tool, or WPT. The interested reader may also wish to refer to our accompanying paper, in which we present a Matlab winch optimisation client implemented with several intelligent algorithms, the MWOC, and test it within this same framework (Hameed et al., 2017).

Intelligent Winch Prototyping Tool (WPT)

Implementing all the necessary design parameters in a winch calculator based on detailed models of the physics involved, we are able to calculate the theoretical physical properties for a given winch design as defined by the chosen set of parameter values. The aim of the winch designer is choose the parameter values that result in a winch design with desirable properties, usually expressed as KPIs, while simultaneously meeting requirements by laws, regulations, codes and standards. Our industrial partner, Seaonics AS, has identified a subset of the most important design parameters that the winch designer is free to experiment with. Via a web GUI (see below), the designer can set and manually tune these design parameters, or use a GA to optimise the design based on some desired optimisation criteria (see Figure 3). Seaonics AS has tested the WPT and the accuracy of the tool has been verified against other existing tools such as spreadsheets currently in use in the industry.

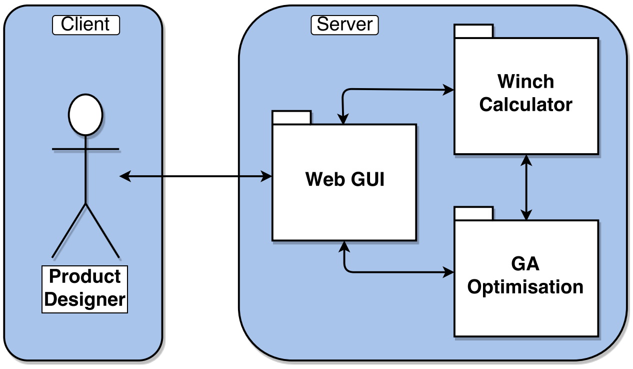


Figure 3: Winch prototyping tool (WPT).

Web Graphical User Interface (GUI)

To simplify practical use of the winch calculator, we have implemented a web GUI (see Figure 4). The GUI has two main panes: one for user input (left-hand pane) that allows the user to select one of three tabs: *Specify*, *View*, and *Optimize*; and one for displaying a graph of key torque characteristics (right-hand pane).

Under the Specify tab, the user can enter values for the winch design parameters, and observe how newly entered values will update the graphical key torque characteristics in the right-hand pane. Parameters are grouped together and categorised as belonging to one of four major components of the winch, namely the drum, the electrical motor, the hydraulic motor, or the wire. Under the View tab, the user can set the resolution (number of data points) of the graph; the total number of winch layers; and the current layer to be observed in the graph. The user can also generate a file in portable document format (PDF) that contains a plot for all the winch layers. Under the Optimize tab, the user can use a GA to optimise a winch design based on a user-defined objective function. To do so, the user must (i) set a number of settings for the GA (see Table 1); (ii) define a suitable objective function (more details in following sections); and (iii) set the allowable ranges (constraints) for each design parameter (optimisation variable) to be optimised.

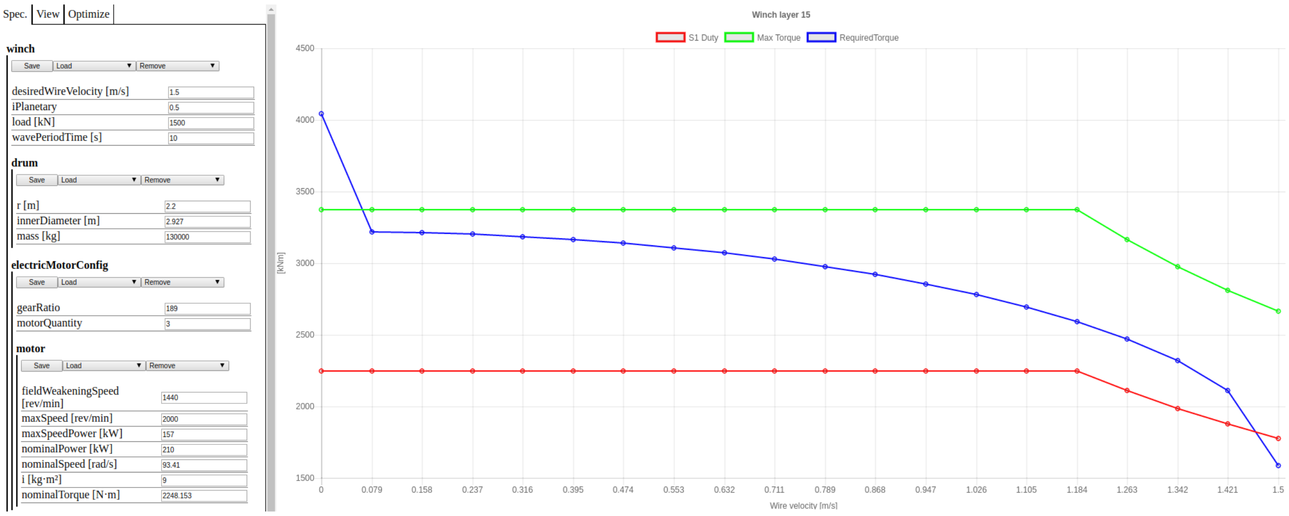
The right-hand pane shows graphically the S1 continuous duty cycle[5] torque (red), the maximum torque (green), and the required torque (blue), as functions of the wire velocity, for a given set of winch specifications and for the particular winch layer defined under the Specify and View tabs, respectively. When a parameter value changes, or after an optimisation has been run, the plots are automatically updated to reflect the effect on the three torque profiles determined by the winch calculator.

The Genetic Algorithm (GA)

The GA (Holland, 1975) is an intelligent algorithm inspired by natural evolution and principles such as inheritance, mutation, selection, and crossover. GAs are well suited for hard optimisation problems (e.g., where solutions are difficult or impossible to obtain in polynomial time) and can also conveniently handle constraints. Since its popularisation in the 1980s, the GA has continued to be a very popular optimisation tool across many different disciplines (e.g., see Haupt and Haupt, 2004). Indeed, in addition to our work on design of offshore cranes and winches, the authors and colleagues have themselves used GAs for a number of diverse real-world optimisation problems, including a general optimisation and machine learning framework for pedagogical and industrial use (Hatledal et al., 2014), boids swarm models (Alaliyat et al., 2014), and dynamic resource allocation with maritime application (DRAMA) (e.g., Bye, 2012; Bye and Schaathun, 2014, 2015).

We assume that the reader is somewhat familiar with GA and refer to a previous paper (Bye et al., 2015) and relevant literature (e.g., see Haupt and Haupt, 2004) for pseudocode and more details. Table 1 shows a summary of some basic GA parameters with typical values that must be set in the web GUI before the GA is run. For the particular objective function we use here, we must also set a *resolution* parameter N_r (see next section).

[5] One of eight duty cycle classifications (S1–S8) provided by the International Electrotechnical Commission in the IEC 60034-1 standard.



winch

desiredWireVelocity [m/s] 1.5
iPlanetary 0.5
load [kN] 1500
wavePeriodTime [s] 10

drum

r [m] 2.2
innerDiameter [m] 2.927
mass [kg] 130000

electricMotorConfig

gearRatio 189
motorQuantity 3

motor

fieldWeakeningSpeed [rev/min] 1440
maxSpeed [rev/min] 2000
maxSpeedPower [kW] 157
nominalPower [kW] 210
nominalSpeed [rad/s] 93.41
i [kg·m²] 9
nominalTorque [N·m] 2248.153

Figure 4: Web GUI for winch prototyping tool (WPT).

parameter	typical
candidates, N_c	100
parents, N_p	50
elites, N_e	10
mutations, N_m	10
generations, N_g	500
resolution, N_r	20

Table 1: GA parameters with typical values.

CASE STUDY

As a simple case study, the main KPI that Seaonics AS is interested in consists of the three torque profiles that result from a given winch design, namely the S1 continuous duty cycle torque T_{S1}, the maximum torque $T_{\max}$, and the required torque T_{req}, which are all functions of the wire velocity v, which has a resolution of N_r sample points between zero and the maximum wire velocity $v_{\max}$ (see Figure 4). The S1 torque is the maximum continuous duty cycle with constant load that the electric motors can safely operate under. The maximum torque is an upper threshold at which the electric motors can safely operate under but only for shorter periods of time. The required torque is the minimum torque required for safe operation for a given constant load.

The torque profiles are also dependent on the winch layer of interest. The number of winch layers, as well as the winch layer to be inspected, can be set in the View tab in the web GUI. Since the torque requirements of the winch increase with winch layers, we conservatively optimise the design parameters for the outermost winch layer. In this paper, the winch is designed with 15 layers and we optimise with respect to layer 15. Notably, however, our GA is able to take all layers into consideration if needed.

Design Parameters

Seaonics AS has provided us with 28 design parameters that can used for optimisation of winch design (see Table 2).[6] As indicated, these parameters can further be divided into five subsets as being *general*, or related to the *drum*, the *electric motors* or *hydraulic motors*, or the *wire*. Choosing the default values for each parameter results in a winch with the same torque profiles as depicted previously in Figure 4. This default design acts as a baseline winch that was designed by a human operator at Seaonics and hereafter will be subject to optimisation and comparison.

For optimisation, we first need to decide which parameters to include as optimisation variables. Keeping the four general parameters such as the wave period and load constant at their default values makes sense, since these parameters relate to the kind of operating scenario for which we need to determine optimised solutions, and the GA should not be allowed to modify these. The 24 remaining parameters relating to the drum, motors, and wire may all have an influence on the main KPI we are interested in, namely the torque profiles. However, for illustration purposes, we limit ourselves to only five parameters that intuitively should strongly influence the torque profiles, namely the inner diameter of the drum, and the gear ratios and quantities of the electric and hydraulic motors (shown in bold in Table 2).

Whilst our GA easily can optimise over the entire set of parameters, one should ideally have a more realistic library of components with fixed parameters, and the GA should optimise the composition of several components put together rather than individual parameters. Our GA has been implemented to be able to perform such component-wise optimisation, however, Seaonics AS has not yet been able to provide us with a useful library of components and we therefore perform optimisation over the set of parameters mentioned above instead. Conceptually, this approach is no different from component-wise optimisation.

Finally, for each design parameter to be optimised, we need to add constraints, that is, minimum and maximum values. Table 2 summarises the parameter settings, includ-

[6] Due to space considerations, we only provide an explanation for selected relevant parameters.

ing default, minimum, maximum, and optimised parameter values.

Objective Function

As mentioned previously, the main KPI that we are concerned with here relates to the torque profiles of T_{req}, T_{max}, and T_{S1} as shown in Figure 4. Because the torque required to rotate the drum and the inertia of the drum and wire increase with the lever arm (the perpendicular distance from the axis of rotation to the line of action of the force), we focus on the worst case when most of the wire is on the drum, in this case winch layer 15.

In order to define a suitable objective function we need to establish what the torque profiles of T_{req}, T_{max}, and T_{S1} should look like. As per information provided by Seaonics AS, a set of guidelines could for instance be given by the following:

- T_{req} should be lower than T_{max} for all wire velocities v_k, except for standstill where $v_0 = 0$, where it could be allowed to be higher.
- T_{req} should be lower than T_{S1} at wire velocities used for continuous operation, that is, typically from half the maximum wire velocity v_{max} and higher.
- conversely, T_{req} should preferably lie between T_{max} and T_{S1} for wire velocities *not* used for continuous operation, that is, typically from half the maximum wire velocity v_{max} and lower.

The rationale for this is that we do not want to use bigger and more expensive motors than necessary, which could lead to T_{req} being below T_{S1} for all velocities, including low velocities not suitable for continuous operation. This rationale is motivated not only by cost, but also by weight and performance, since bigger motors will have higher mass and inertia. Hence, the possibility of operating above the nominal S1 duty rating for the motors should be utilised. Operating above the S1 duty cycle is only possible for shorter periods of time due to heat accumulation in the motors. Consequently, after a short period of operating above S1, the motors must be allowed to operate below S1 in order to cool down.

It is important to understand that the guidelines above are merely an example of suitable guidelines for a specific application. Say, if the winch would be used in constant tension mode, then either T_{req} must be lower than T_{S1} at zero wire velocity to avoid overheating and failure, or other heat-preventing precautions must be taken, such as additional cooling by installing more fans or by installing water cooling.

Based on the above guidelines we have devised the following objective function, which is in fact a cost function:

$$f_{\text{cost}} = \sum_{k=1}^{N_{\text{r}}} a \cdot \text{R}_k + (1-a) \cdot \text{S}_k \tag{1}$$

where

$$\text{R}_k = \begin{cases} \delta\text{R}_k^2 \text{ for } \delta\text{R}_k > 0 \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

$$\text{S}_k = \delta\text{S}_k^2 \tag{3}$$

$$\delta\text{R}_k = T_{\text{req}}(v_k) - T_{\text{max}}(v_k) \tag{4}$$

$$\delta\text{S}_k = T_{\text{req}}(v_k) - T_{\text{S1}}(v_k) \tag{5}$$

$$a = 0.5 \tag{6}$$

and v_k is the kth sample of the wire velocity. That is, this cost function is a weighted sum of the squared difference between T_{req} and T_{max} for only those velocities v_k where T_{req} is higher than T_{max}, and the squared difference between T_{req} and T_{S1}. The effect of the first term is that an intolerable torque T_{req} higher than T_{max} is punished severely. For the second term, the smallest cost of zero at any wire velocity is achieved for T_{req} equal to T_{S1}, whilst T_{req} being either higher or lower than T_{S1} is punished severely, and more so the bigger the difference, due to squaring. Because T_{req} will typically have a falling torque profile, due to higher torque requirements at lower velocities, the intention of the second term is to obtain a profile for T_{req} that is higher than T_{S1} for low velocities and lower than T_{S1} for high velocities. The two terms can be weighted relative to each other by the weighting factor a, here set to $a = 0.5$.

Results

Figure 5 shows the default torque profiles for T_{max} (red), T_{req} (blue), and T_{S1} (green) for winch layers 1, 5, 10, and 15 before optimisation, and for $T_{\text{req,GA}}$ (black) after GA optimisation (T_{max} and T_{S1} remain unchanged).

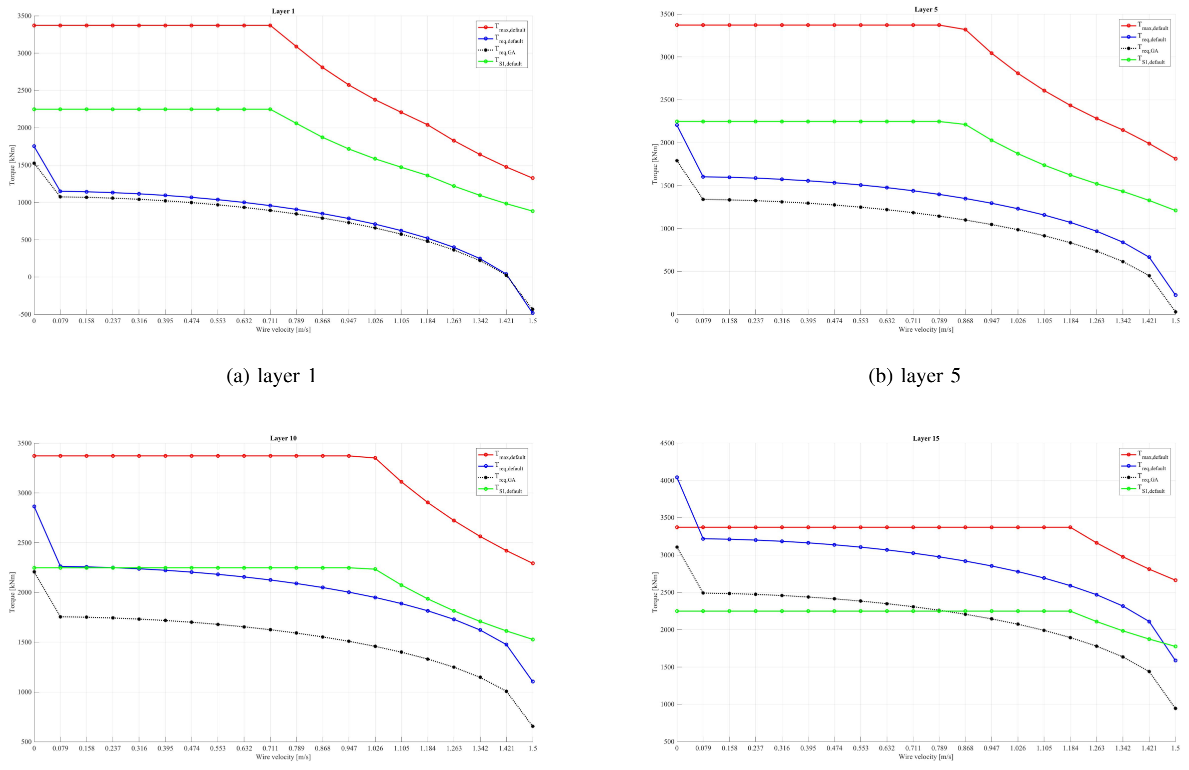
For winch layers 1 and 5, there is not much difference between T_{req} and $T_{\text{req,GA}}$, whereas for winch layers 10 and 15 there is a big improvement in reduced required torque from GA optimisation resulting in $T_{\text{req,GA}}$. Comparing with the guidelines we used to define the objective function, we observe that all three requirements for the worst case of layer 15 are satisfied.

The optimised values for the five design parameters that were chosen as optimisation variables are provided in boldface in Table 2. After optimisation, the size of the inner diameter of the drum has increased from its default value of 2.927 m to 2.97 m, which is close to the parameter maximum constraint of 2.99 m. The number of electric motors has increased from 3 to 4, whereas the number of hydraulic motors is unchanged at 4. Finally the gear ratio of the electric motors has increased from their default value of 189 to 199.8, which is very close to the parameter maximum constraint of 200. The gear ratio of the hydraulic motors has increased slightly from their default value of 159.16 to 167.90, which is close to the middle of the constrained parameter ranged from 150 to 190.

It is not surprising that having an extra electric motor with a better gear ratio reduces the required torque, however, it is less intuitive that the gear ratio of the hydraulic motor seems less important (it was not driven towards its

subset	number	name	units	default	min	max	optimised
general	1	wavePeriodTime	s	10	-	-	10
	2	load	kN	1500	-	-	1500
	3	iPlanetary	-	0.50	-	-	0.50
	4	desiredWireVelocity	m/s	1.50	-	-	1.50
drum	5	r	m	2.20	-	-	2.20
	6	**innerDiameter**	**m**	**2.927**	**2.70**	**2.99**	**2.97**
	7	mass	kg	130000	-	-	130000
electric motor	**8**	**gearRatio**	**-**	**189.0**	**170.0**	**200.0**	**199.8**
	9	**motorQuantity**	**-**	**3**	**1**	**5**	**4**
	10	fieldWeakeningSpeed	rev/min	1440	-	-	1440
	11	maxSpeed	rev/min	2000	-	-	2000
	12	maxSpeedPower	kW	157.0	-	-	157.0
	13	nominalPower	kW	210	-	-	210
	14	nominalSpeed	rad/s	93.410	-	-	93.410
	15	i	kg·m^2	9.00	-	-	9.00
	16	nominalTorque	N·m	2248.15	-	-	2248.15
hydraulic motor	**17**	**gearRatio**	**-**	**159.16**	**150.00**	**190.00**	**167.90**
	18	**motorQuantity**	**-**	**4**	**1**	**5**	**4**
	19	friction	-	0.950	-	-	0.950
	20	staticEfficiency	-	0.789	-	-	0.789
	21	dynamicEfficiency	-	0.916	-	-	0.916
	22	maxSpeed	rev/min	1600	-	-	1600
	23	displ	cm^3	1000	-	-	1000
	24	pressureDrop	bar	280	-	-	280
	25	i	kg·m^2	0.550	-	-	0.550
	26	nominalTorque	N·m	4457.71	-	-	4457.71
wire	27	diameter	mm	77.0	-	-	77.0
	28	diameterReductionFactor	-	0.866	-	-	0.866

Table 2: Default and optimised (bold) winch design.



(a) layer 1

(b) layer 5

(c) layer 10

(d) layer 15

Figure 5: Default torque profiles for $T_{\max}$ (red), T_{req} (blue), and T_{S1} (green) for winch layers 1, 5, 10, and 15 before optimisation, and for $T_{\text{req,GA}}$ (black) after GA optimisation ($T_{\max}$ and T_{S1} remain unchanged).

maximum value constraint like its electric counterpart). This may be because the hydraulic motors have a higher nominal torque and are not suffering from field weakening at higher speeds such as electric motors. Due to different speeds and gear ratios on hydraulic and electric motors, one kind of motor might be operating below maximum whilst the other kind is at maximum, and this might explain why the gear ratio for the hydraulic motors is not driven to its maximum by the GA as for the electric motors. That the inner diameter of the drum is nearly maxed out from a default value of 2.927 to a value of 2.97, with the maximum parameter constraint being 2.99, is counter-intuitive when considering torque alone. However, we suspect that it has been increased to ensure that the default speed requirements of the motors are satisfied, as the wire velocity increases with the diameter of the drum for the same rotational speed.

The results show that the GA indeed is capable of improving a default unsatisfactory design to yield an optimised design that satisfies the design guidelines we set out previously.

DISCUSSION

In this paper, we have expanded our software framework from earlier work on design optimisation of offshore cranes to also include maritime winches by means of a WPT. All our work, previous and current, have been performed in cooperation with two industrial partners, ICD Software AS and Seaonics AS, to ensure correctness and relevance of our projects. We have successfully tested the WPT on a design optimisation problem of reducing the required torque for a winch in a desired manner described previously by letting a GA determine suitable values for a subset of five design parameters. Choosing a suitable objective function and which design parameters to optimise is dependent on the kind of operation the winch is intended for, physical limitations such as available components (e.g., motors) and weight, size, and cost requirements, to name a few.

Implementation Details

The WPT differs from our previous work in that the optimisation module and the web GUI is also implemented on the server-side, thus offering a complete solution to end-users with no need for a local installation, but more importantly, removing communication overhead between a client-side POC and the server-side product calculator. Nevertheless, simultaneously, we have ensured that the WPT is compatible within the client-server architecture of our framework, which means it is still possible to implement winch optimisation clients in any language of choice (e.g., Matlab Hameed et al., 2017) that can connect to the winch calculator through the HTTP/JSON and WS/JSON interfaces. Nevertheless, we note the benefit of letting the software modules for the POC and the product calculator co-exist on the same server to avoid communication overhead.

Furthermore, we have implemented optional authentication for the WS communication interface and for the server-side WPT, requiring users to be registered and enter a password for access. This feature can useful for licensing of software, e.g., on a time-limited basis, and other models of commercialisation that our industrial partners want to proceed with.

Finally, we wish to re-iterate that our software framework is highly modular and generic, as we have demonstrated here and in our earlier work.

Web GUI and Future Work

In our earlier work on crane design optimisation (Bye et al., 2016), the POC using a GA was not accessible via a web GUI, which raised the bar significantly for usage by a product designer without programming experience and/or AI knowledge. In the WPT we present here, we have incorporated application of a GA by means of a simple user interface where a winch designer can perform winch optimisation by the press of a button. The web GUI also offers some useful defaults for GA settings and parameter values and boundaries that the designer can modify as needed.

For the future, we would like to implement some improvements to the web GUI. First, as short electronic manual outlining the basics of GAs as well as the effect of the GA settings should be provided, possibly integrated in the web GUI (e.g., by mouseovers and/or a separate webpage). Second, the manual should also include notes on how to design useful objective functions, and the web GUI should store a library of such functions, with explanations, for different optimisation purposes. Third, Seaonics AS should provide a library of real-world components with pre-defined sets of parameters that the GA should combine in an optimal manner. Fourth, the web GUI should allow for import and export of optimisation parameters to allow for batch processing and analysis in the design process. Finally, the auto-generated report tool, which currently exports plots of the torque profiles of all the winch layers could be expanded to contain more information and quantitative data.

ACKNOWLEDGEMENTS

The SoftICE lab at NTNU in Ålesund wishes to thank ICD Software AS for their contribution in the software development process, and Seaonics AS for providing documentation and insight into the design and manufacturing process of offshore cranes. We are also grateful for the support provided by Regionalt Forskningsfond (RFF) Midt-Norge and the Research Council of Norway through the VRI research projects *Artificial Intelligence for Crane Design (Kunstig intelligens for krandesign (KIK))*, grant no. 241238, and *Artificial Intelligence for Winch Design (Kunstig intelligens for vinsjdesign (KIV))*, grant no. 249171.

REFERENCES

Alaliyat, S., Yndestad, H. and Sanfilippo, F. (2014), Optimisation of Boids Swarm Model Based on Genetic Algorithm and Particle Swarm Optimisation Algorithm (Comparative Study), Proceedings of the 28th European Conference on Modelling and Simulation (ECMS '14), pp. 643–650.

Bye, R. T. (2012), A receding horizon genetic algorithm for dynamic resource allocation: A case study on optimal positioning of tugs., *Series: Studies in Computational Intelligence* 399, 131–147. Springer-Verlag: Berlin Heidelberg.

Bye, R. T., Osen, O. L. and Pedersen, B. S. (2015), A computer-automated design tool for intelligent virtual prototyping of offshore cranes, Proceedings of the 29th European Conference on Modelling and Simulation (ECMS '15), pp. 147–156.

Bye, R. T., Osen, O. L., Pedersen, B. S., Hameed, I. A. and Schaathun, H. G. (2016), A software framework for intelligent computer-automated product design, Proceedings of the 30th European Conference on Modelling and Simulation (ECMS '16), pp. 534–543.

Bye, R. T. and Schaathun, H. G. (2014), An improved receding horizon genetic algorithm for the tug fleet optimisation problem, Proceedings of the 28th European Conference on Modelling and Simulation (ECMS '14), pp. 682–690.

Bye, R. T. and Schaathun, H. G. (2015), Evaluation heuristics for tug fleet optimisation algorithms: A computational simulation study of a receding horizon genetic algorithm, Proceedings of the 4th International Conference on Operations Research and Enterprise Systems (ICORES '15), pp. 270–282.

Carral, J., Carral, L., Lamas, M. and Rodríguez, M. J. (2015), Fishing grounds' influence on trawler winch design, *Ocean Engineering* 102, 136–145.

GCE Blue Maritime Cluster (2016), Breaking Waves: Operations Report 2016, Accessed on 6 February 2017 from http://www.bluemaritimecluster.no.

Griffin, B. (2004), High performance winch and synthetic rope systems for workboats, tug boats, and commercial marine applications, OCEANS'04. MTTS/IEEE TECHNO-OCEAN'04, Vol. 4, IEEE, pp. 1900–1903.

Hameed, I. A., Bye, R. T. and Osen, O. L. (2016*a*), A Comparison between Optimization Algorithms Applied to Offshore Crane Design using an Online Crane Prototyping Tool, Proceedings of the 2nd International Conference on Advanced Intelligent Systems and Informatics (AISI '16), Vol. 533 of *Advances in Intelligent Systems and Computing (AISC)*, pp. 266–276.

Hameed, I. A., Bye, R. T. and Osen, O. L. (2016*b*), Grey wolf optimizer (GWO) for automated offshore crane design, Proceedings of the IEEE Symposium Series on Computational Intelligence (IEEE SSCI '16), pp. 1–6.

Hameed, I. A., Bye, R. T., Osen, O. L., Pedersen, B. S. and Schaathun, H. G. (2016), Intelligent computer-automated crane design using an online crane prototyping tool, Proceedings of the 30th European Conference on Modelling and Simulation (ECMS '16), pp. 564–573.

Hameed, I. A., Bye, R. T., Pedersen, B. S. and Osen, O. L. (2017), Evolutionary winch design using an online winch prototyping tool, Proceedings of the 31st European Conference on Modelling and Simulation (ECMS '17). Under review.

Hare, W., Nutini, J. and Tesfamariam, S. (2013), A survey of non-gradient optimization methods in structural engineering, *Advances in Engineering Software* 59, 19–28.

Hatledal, L. I., Sanfilippo, F. and Zhang, H. (2014), JIOP: a java intelligent optimisation and machine learning framework, Proceedings of the 28th European Conference on Modelling and Simulation (ECMS 14), pp. 101–107.

Haupt, R. L. and Haupt, S. E. (2004), *Practical Genetic Algorithms*, 2nd edn, Wiley.

Holland, J. H. (1975), *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*, University of Michigan Press, Oxford, England.

Kamentsky, L. and Liu, C. (1963), Computer-automated design of multifont print recognition logic, *IBM Journal of Research and Development* 7(1), 2–13.

Lincoln, J. M., Woodward, C. C., King, G. W., Case, S. L., Lucas, D. L. and Teske, T. D. (2016), Preventing fatal winch entanglements in the US southern shrimp fleet: A research to practice approach, *Journal of Safety Research* .

Pearlman, S. M., Gordon, D. R. and Pearlman, M. D. (2017), Winch Technology — Past Present and Future: A Summary of Winch Design Principles and Developments, Technical report, InterOcean Systems LLC. Accessed on 14 February 2017 from http://www.interoceansystems.com/winch_article.htm.

Pratt, M. J. (1995), Virtual prototypes and product models in mechanical engineering, *Virtual Prototyping–Virtual Environments and the Product Design Process* 10, 113–128.

Skjong, S. and Pedersen, E. (2016), Model-based control designs for offshore hydraulic winch systems, *Ocean Engineering* 121, 224–238.

Solovyov, V. and Cherniavsky, A. (2013), Computational and experimental analysis of trawl winches barrels deformations, *Engineering Failure Analysis* 28, 160–165.

Zhang, J., Zhan, Z., Lin, Y., Chen, N., jiao Gong, Y., hui Zhong, J., Chung, H., Li, Y. and Shi, Y. (2011), Evolutionary computation meets machine learning: A survey, *IEEE Computational Intelligence Magazine* 6(4), 68–75.

AUTHOR BIOGRAPHIES

ROBIN T. BYE[7] is an IEEE Senior Member who graduated from the University of New South Wales, Sydney with a BE (Hons 1), MEngSc, and a PhD, all in electrical engineering. Working at NTNU in Ålesund since 2008, Dr. Bye is an associate professor in automation engineering. His research interests belong to the fields of artificial intelligence, cybernetics, neuroengineering, and education.

OTTAR L. OSEN is MSc in Engineering Cybernetics from the Norwegian Institute of Technology in 1991. He is the head of R&D at ICD Software AS and an assistant professor at NTNU Ålesund.

WEBJØRN REKDALSBAKKEN is an associate professor in engineering cybernetics at NTNU Ålesund. Norway. He has a master in physics from former NTH, Norway.

BIRGER SKOGENG PEDERSEN graduated from NTNU Ålesund with a BE in automation engineering and is a former employee at ICD Software AS. He is currently a MSc student of simulation and visualisation and employed as a research assistant in the Mechatronics Laboratory at NTNU Ålesund.

IBRAHIM A. HAMEED is an IEEE Senior Member and has a BSc and a MSc in Industrial Electronics and Control Engineering, Menofia University, Egypt, a PhD in Industrial Systems and Information Engineering from Korea University, S. Korea, and a PhD in Mechanical Engineering, Aarhus University, Denmark. He has been working as an associate professor at NTNU Ålesund since 2015. His research interests includes artificial intelligence, optimisation, control systems, and robotics.

[7] www.robinbye.com

RUSSIAN LICENSE PLATE SEGMENTATION BASED ON DYNAMIC TIME WARPING

Mikhail A. Povolotskiy
Moscow Institute of Physics and Technology
Institutskiy per., 141700, Russia;
Institute for Information
Transmission Problems, RAS
Bolshoy Karetny per., 127994, Russia
E-mail: mikhail.povolotskiy@iitp.ru

Elena G. Kuznetsova
Institute for Information
Transmission Problems, RAS
Bolshoy Karetny per., 127994, Russia
E-mail: kuznetsova@iitp.ru

Timur M. Khanipov
Institute for Information
Transmission Problems, RAS
Bolshoy Karetny per., 127994, Russia
E-mail: timur.khanipov@iitp.ru

KEYWORDS

License plate recognition, license plate segmentation, dynamic time warping, dynamic programming.

ABSTRACT

We propose a fast plate segmentation algorithm for automatic license plate recognition systems which is stable to plate rectangle localization inaccuracies and image brightness distortions. The algorithm uses a priori information about the geometry of standard plate types and additionally makes adjustments to symbols positions through localization errors estimation and correction. We introduce a plate localization error model and compute its optimal parameters using dynamic programming. We also suggest a modification of the algorithm for simultaneous segmentation and optimal type selection (from a known set of types). Experimental results are presented which show the efficiency of this approach.

INTRODUCTION

Most commonly (Du et al. 2013) plate recognition on an image is performed with the following stages:
1. License plate extraction — at this stage for each plate present on the source image a bounding quadrangle is searched and its inside area is transformed projectively to a rectangular view;
2. License plate segmentation — symbol regions detection and all symbols' images extraction;
3. Character recognition — obtaining each symbol value.

Segmentation algorithms are usually based on a priori knowledge of difference between symbols and background color. The first common method is connected components search on a binarized plate image (Uddin et al. 2016). This approach has certain problems specific to our task. In particular, brightness and contrast distribution on the image region may be non-uniform because of shadows, partial overexposure, dirt on a plate. In this case, binarization errors will occur. This leads to connected components of symbols getting split and merged (see figure 1 as an example).

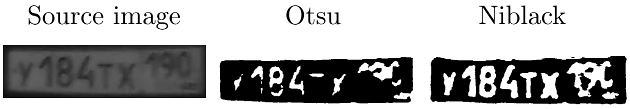


Figure 1. Examples of Common Used Binarization Algorithms Operation

Another well-known approach (Xia and Liao 2011) is the sequential search of vertical and horizontal symbol borders by maximizing deviations of vertical/horizontal projections (or other brightness functions) of rows/columns corresponding to symbols and the background (see figure 2 as an example of implementation). This method is also non-robust to brightness distortion.

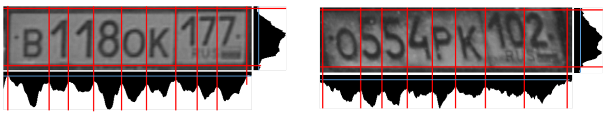


Figure 2. Histogram-based Algorithm Operation Example (It Uses the Histogram of Image Spatial Derivatives Moduli)

The third wide-spread group of methods is based on symbols localization using a priori information about position and size of plate regions known from various state standards. These algorithms have low ($O(1)$) computational complexity and brightness distortion robustness. If several plate types should be recognized, such methods cannot be used without prior type selection.

In (Gao et al. 2007) plate type is chosen using information about font and background color. This approach, however, cannot be used for distinguishing types with same background and symbols colors but different geometry which is the case, for example, for standard plate types in Russia (figure 3). This group of methods also does not provide robustness to unavoidable inaccuracies of the preceding recognition stages which cause inaccurate plate rectangle localization.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

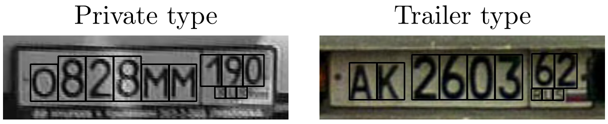


Figure 3. Examples of Vehicle Registration Plates of Russia

To tackle this problem a combined algorithm is used in (Paliy et al. 2004) which optimizes symbols localization result based on maximizing symbols and background brightness deviations described above. The algorithm utilizes information about plate geometry while checking various plate region shifts (see figure 6). However, increasing the accuracy of plate border shifts cannot compensate for localization imprecision caused by errors in determining plate rectangle border tilt angles.

A more complex distortion model is used in (Tian et al. 2015) which also considers various border tilt angles (assumed to be equal for the left and the right borders) as well as plate scales. The main disadvantage of an algorithm proposed in this paper is its high computational complexity because of the necessity to calculate image rotations for all possible tilt angles and to iterate through a large number of shifts and scales combinations. Besides, this model does not cover cases of non-equal borders rotations.

It is apparent that robustness to borders rotations can be reached by generalizing the above described approach on the case of projective plate rectangle distortions. That is, we should check all possible independent shifts for each rectangle vertex during optimal plate location search. This algorithm, however, would have unacceptable computational complexity due to a huge number of 4 vertices shifts combinations and a need to apply projective transform at each iteration to compensate distortions.

In this paper we propose a fast segmentation algorithm based on testing of various rectangle vertices shifts by approximating the localized number projective distortions with symbol rectangles shifts. We also suggest a generalization of this algorithm for automatic plate type selection from the set of alternatives determined by the recognition system requirements.

SEGMENTATION ALGORITHM FOR A FIXED TYPE

Let $p_i = [x_i, y_i, w_i, h_i]$ be a plate region of the i-th symbol with coordinates of top-left corner x_i, y_i, width w_i, and height h_i. We will model plate distortions by changing symbol regions with the following limitations:

1. All plate symbols regions $p_0, p_1, \ldots, p_{n-1}$ can be shifted in-parallel on an arbitrary number of pixels

$$p'_i = [x'_i, y'_i, w'_i, h'_i] = [x_i + \Delta x, y_i + \Delta y, w_i, h_i] \\ \forall i \in [0, n-1]$$

in vertical and horizontal axes. Symbol regions must not exceed image borders:

$$\begin{aligned} x_i + \Delta x \geq 0, x_i + w_i - 1 + \Delta x \leq \mathcal{W}, \\ y_i + \Delta y \geq 0, y_i + h_i - 1 + \Delta y \leq \mathcal{H}, \end{aligned} \tag{1}$$

where $\mathcal{W}, \mathcal{H}$ are image width and height.
2. Symbols are allowed to be shifted relatively to each other if the relative distance changes between the centers of adjacent symbols are limited by δ (method parameter);

$$\begin{gathered} p''_i = [x''_i, y''_i, w''_i, h''_i] = [x'_i + \Delta x_i, y'_i + \Delta x_i], \\ |\Delta x_i - \Delta x_{i-1}| < \delta \cdot |Cx_i - Cx_{i-1}|, \\ |\Delta y_i - \Delta y_{i-1}| < \delta \cdot |Cy_i - Cy_{i-1}| \\ \forall i \in [1, n-1], \end{gathered} \tag{2}$$

where (Cx_i, Cy_i) are the center coordinates of rectangle p''_i.
3. Symbol region size change is neglected.

$$w_i = w'_i = w''_i, \ h_i = h'_i = h''_i \ \forall i \in [0, n-1] \tag{3}$$

Expressions (1)-(3) correspond to plate projective distortion approximation by shifts with a limited variance. Expression (1) limits number shift in the modeled localization inaccuracy, (2) limits the distortions set to a projective transform, (3) limits the aspect ration for the modeled inaccuracy.

Optimal symbols location is chosen from the set of integer parameters of the described model satisfying conditions (1)-(3)

$$\Delta x, \Delta y, \Delta x_i, \Delta y_i, \forall i \in [0, n-1]$$

and maximizing total brightness inside symbol regions

$$\sum_{i=0}^{n-1} \sum_{x,y \in p''_i} I(x, y) \rightarrow \max_{\Delta x, \Delta y, \Delta x_i, \Delta y_i}, \tag{4}$$

where $I(x, y)$ is the brightness value for plate image pixel with coordinates x, y. It is assumed that symbols brightness for a given plate type is greater than background brightness. If this condition does not hold, the image I is inverted first (see figure 4).

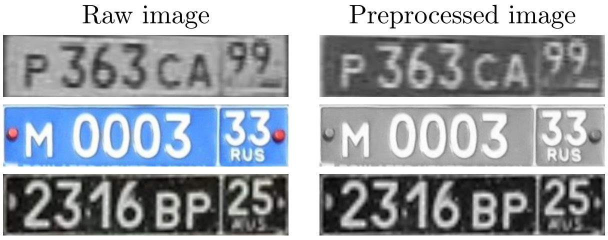


Figure 4. Image Preprocessing

Despite the symbol regions limitation imposed by the chosen model, optimal location exhaust search is computationally difficult. Indeed, the number of possible arrangements increases exponentially with the number of symbols, and for each arrangement one has to sum brightness for all pixels inside symbol regions.

The summation complexity can be trivially reduced using the fact that each symbol region is a rectangle. Its sum is computed by $O(1)$ by integrating the plate image

$$J(x,y) = \sum_{u=0}^{x} \sum_{v=0}^{y} I(u,v). \tag{5}$$

The expression (5) can be quickly computed by using the dynamic scheme

$$\begin{gathered} J(0,0) = I(0,0), \\ J(x,0) = J(x-1,0) + I(x,0), x > 0 \\ J(x,y) = J(x,y-1) + I(x,y), y > 0. \end{gathered} \tag{6}$$

The pixel brightness sum inside any rectangle $p = [x,y,w,h]$ will then be

$$\begin{gathered} S(p) = \sum_{x,y \in p} I(x,y) = \\ = J(x+w-1, y+h-1) - J(x-1, y+h-1) - \\ -J(x+w-1, y-1) + J(x-1, y-1). \end{gathered} \tag{7}$$

Optimal parameters search can also be accelerated using the iterative calculation scheme (with the number of iterations N being the algorithm parameter). Each iteration is the search of optimal shifts $\Delta u, \Delta u_i$ with fixed $\Delta v, \Delta v_i$, where $u = x, v = y$ for even iterations and $u = y, v = x$ for odd iterations. Optimal parameters $\Delta\widetilde{u}, \Delta\widetilde{u_i}$ computation at each step is performed using the dynamic time warping algorithm (DTW) (Vintsyuk 1968). Fixed parameters $\Delta v, \Delta v_i$ values for the current iteration are set to the optimal values of the $\Delta\widetilde{u}, \Delta\widetilde{u_i}$ parameters obtained in the previous iteration. For the first iteration $\Delta v = \Delta v_i = 0$.

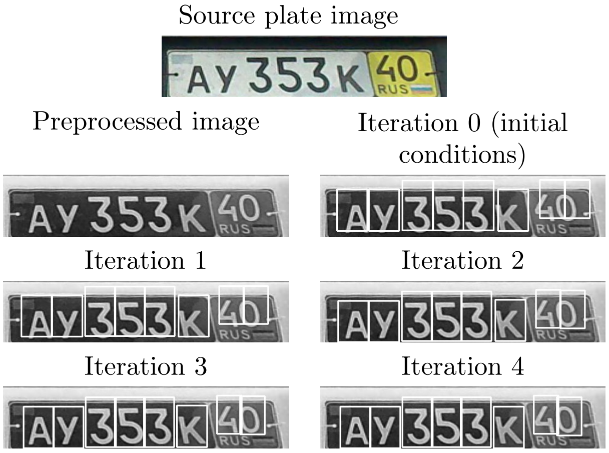
For every possible shift $U_{n-1} = \Delta u + \Delta u_{n-1}$ the algorithm's result is a sequence $T(U_{n-1})$ of shifts $[U_0, U_1, \dots U_{n-1}]$, satisfying conditions (1)-(3) and maximizing the functional $W(p''_{n-1})$ with fixed $\Delta v, \Delta v_i$, defined in (4). The algorithm for dynamic calculation of such states is described by the following scheme:
1. $T(U_0) = [U_0],\ W(p''_0) = S(p''_0)$
2. $T(U_i) = [T(argmax_{U_{i-1}} W(p''_{i-1})),\ U_i]$,
$W(p''_i) = \max_{U_{i-1}}(W(i-1, p_{i-1})'' + S(p''_i))$,
$|U_i - U_{i-1}| \le \delta \cdot |Cu_i - Cu_{i-1}|$
for all possible $p''_i, i = 0 \dots n-1$, satisfying conditions (1)-(2), where $S(p''_i)$ is defined by (7).

Note that in this scheme transition to each shift u_i of the i-th symbol demands computing a maximum on a segment of a fixed length for a given i, what can be done in $O(1)$ using the van Herk/Gil-Werman algorithm (Van Herk 1992).

We choose optimal set of parameters $\Delta\widetilde{u}, \Delta\widetilde{u_i}$ from the set of $T(U_{n-1})$ such that the corresponding functional $W(p'_{n-1})'$ is maximal.

In fig. 5 the convergence process of the described iterative scheme is shown.

The complexity of a single iteration of the suggested scheme is $O(n \cdot \mathcal{U})$, where $\mathcal{U}$ is the size (in pixels) of the image side along which finetuning is performed (i.e. $\mathcal{W}$ for even operations, $\mathcal{H}$ for odd). Indeed $n \cdot \mathcal{U}$ pairs of values $T(U_i), W(p''_i)$ are computed and each of them is determined in $O(1)$ with the acceleration from using integral images and the van Herk/Gil-Werman method.


Source plate image

Preprocessed image

Iteration 0 (initial conditions)

Iteration 1

Iteration 2

Iteration 3

Iteration 4

Figure 5. Iterative Model Parameters Adjustment

Therefore the total complexity of the suggested plate segmentation algorithm for a given type is $O(|N \cdot n \cdot \max(\mathcal{W}, \mathcal{H}))$.

ALGORITHM MODIFICATION FOR AUTOMATIC TYPE SELECTION

License Plates Description

In Russia several types of vehicle registration plates exist serving various purposes (figure 6). Each type has certain font and background colors, symbols location and size and a specific alphabet for each symbol. The alphabet consists of digits $D = \{0123456789\}$ and the intersection of the cyrillic and latin alphabets $L = \{ABEKMHOPCTYX\}$ for all types except the diplomatic ones. For the diplomatic types, the alphabet $A = \{CDT\}$ is used.

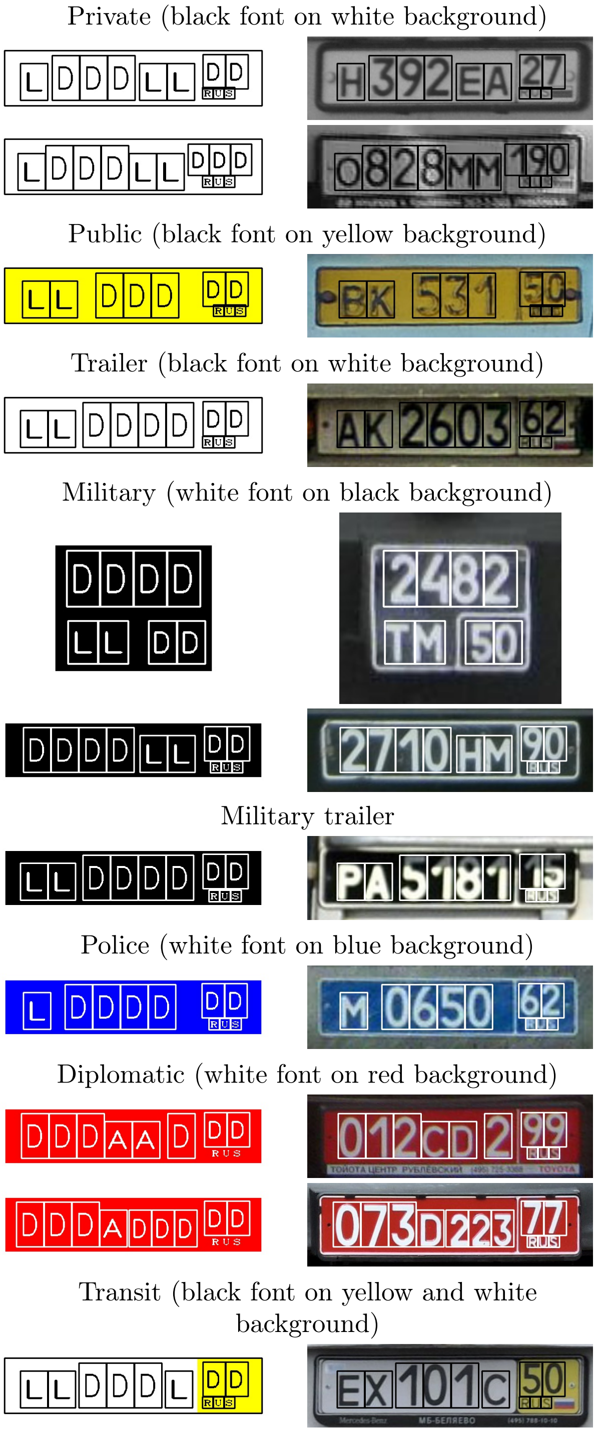
The segmentation algorithm described above needs a priori information about plate type. We will now describe the algorithm modification facilitating simultaneous type selection and optimal parameters of the plate distortion model search.

Proposed Algorithm

The proposed segmentation algorithm with type selection consists of two stages:

- Optimal symbol regions search for each number type from a known alternatives set $\mathscr{T}$ (figure 6) in accordance with the method described above;
- Plate type selection, for which the segmentation result from the previous step is optimal by some criterion which will be defined further.

Since the aggregate area inside symbol regions may vary from type to type, the total brightness maximized during the previous algorithm stage cannot be used as a type selection criterion. Using mean brightness for all pixels inside symbol regions instead of a sum is more sensible but also has a number of disadvantages.


Private (black font on white background)

Public (black font on yellow background)

Trailer (black font on white background)

Military (white font on black background)

Military trailer

Police (white font on blue background)

Diplomatic (white font on red background)

Transit (black font on yellow and white background)

Figure 6. Types of Vehicle Registation Plates of Russia

First, it is possible to select a type which has a smaller number of symbols than the correct one or a type with some symbol regions having smaller area and lying inside the ground truth regions.

Also during plate distortion model parameters computation for types where symbols and background brightness ratio does not correspond to the input plate image, symbol regions search is performed on the image with light background and dark symbols, hence the mean brightness might be high due to background pixels inside symbol regions. To tackle this kind of errors we need a criterion considering the brightness outside of symbol regions.

Let us consider that pixel brightness values in- and outside of symbol regions — are samples of independent normal random variables X_1 and X_2 with different expectations and equal variance, of sizes n_1 and n_2 resp. We will choose such a type for which the expectations equality hypothesis is rejected by the Student's t-test (Gosset 1908) with the smallest significance, i.e. the t-statistics (8) value is maximal.

$$
\begin{aligned}
t &= \frac{\overline{X}_1 - \overline{X}_2}{s_X \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \\
s_X &= \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}},
\end{aligned}
\tag{8}
$$

where s_1^2 and s_2^2 are unbiased variance estimations for the two samples. This test is stable to changes of the total area inside symbol regions for various plate types and has the property of suppressing selection errors for types where the symbols and background brightness ratio does not correspond to the input plate image.

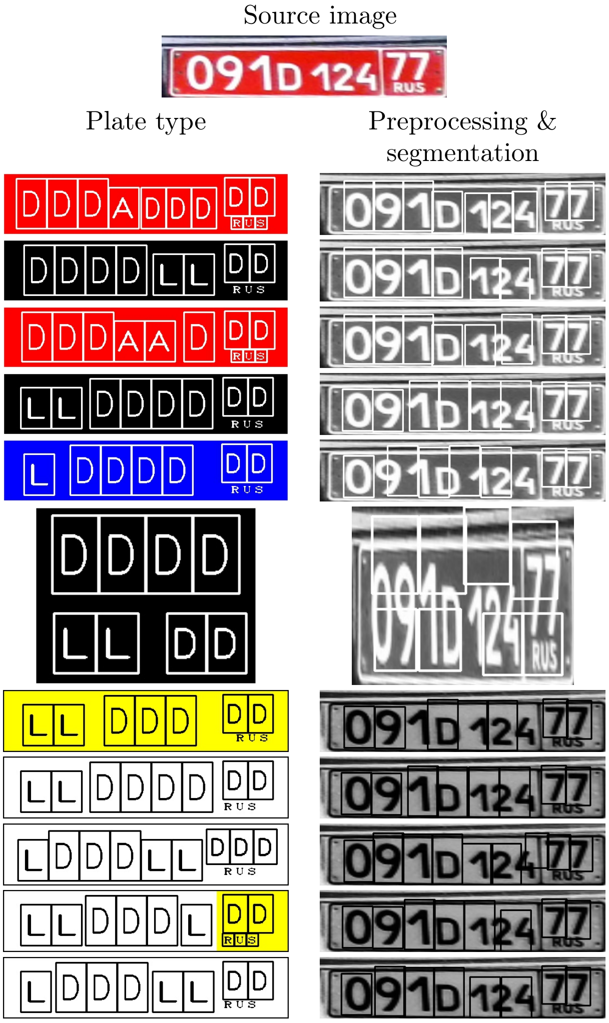

Source image

Plate type

Preprocessing & segmentation

Figure 7. Segmentation Algorithm Results for Russian Type Table (Ordered by Decreasing t-statistics)

Taking into account all the above mentioned, the total complexity of the segmentation algorithm with simultaneous type selection is $|\mathscr{T}| \cdot O(\cdot N \cdot n \cdot max(\mathcal{W}, \mathcal{H})) = O(|\mathscr{T}| \cdot N \cdot n \cdot \max(\mathcal{W}, \mathcal{H}))$, where $|\mathscr{T}|$ is the number of plate types in the selection set.

Note that the suggested algorithm can easily be accelerated by parallel calculations, because the optimal parameters of the plate distortion model are computed independently for each plate type.

Also note that plate type detection during the segmentation stage allows to reduce complexity and increase the quality of the next plate recognition stage - symbols recognition.

The background and symbols colors, as well as region size of each symbol extracted from information about plate type allow to unify input data for the algorithms used in symbol recognition. Also the information about a fixed alphabet for each symbol of the type (see fig. 6) is useful to symbol recognition algorithms allowing to prevent some of their errors.

The simultaneous optimal segmentation parameters computation for all plate types allows to use several alternatives in plate type selection for the subsequent operation stages where type selection can be performed more accurately. For example, recognition algorithms for previously segmented plate symbols (such as machine learning or template matching) usually have their own quality estimator for the symbols recognized, which can by utilized to check correctness for each of the alternatives found. In license plate recognition video systems the best alternative can be selected according to the motion model.

Another algorithm's advantage is that it does not depend on the Russian plate types specifics and may be used for other countries (or sets of countries) with fixed plate types.

EXPERIMENTAL RESULTS

To research the proposed algorithm's efficiency we conducted several computational experiments on its accuracy estimation using a set of vehicles and their plates images (8153 samples) with previously corrected radial distortion (Kunina et al. 2016, 2017). Some images of the set may contain more than one plate while the others may contain none (see figure 8).

We now introduce a metric for segmentation and plate type selection algorithm accuracy assessment.

Since the segmentation and plate type detection algorithm is based on optimal shifts search for each symbol region, it seems reasonable to estimate its accuracy by the distance between the centers of the detected and ground truth symbol regions. We will use the following functional to estimate the algorithm's quality:

$$Q_{shift} = \begin{cases} 0 & t \neq \widetilde{t} \\ \min_{i=0..n-1} S_i & t = \widetilde{t}, \end{cases} \tag{9}$$

where $\widetilde{t}$ is the ground truth plate type, t is the algorithm detected type, n is the number of symbols on a plate, S_i is the quality functional of i-th symbol center detection (assuming correctly found plate type), defined as

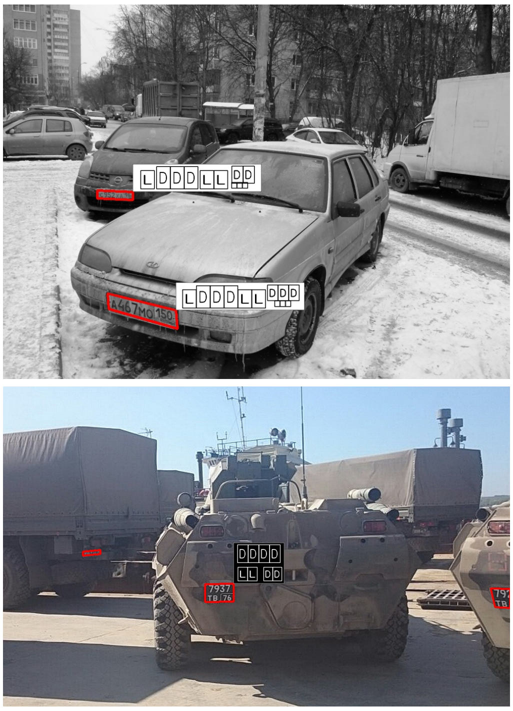


Figure 8. Samples of the Marked Testdata

$$S_i = \max(0, 1 - \frac{||\widetilde{C_i}, C_i||_2}{P(\widetilde{p_i})/4}), \tag{10}$$

where $\widetilde{C_i}$ and C_i are center coordinates of the detected by algorithm and true symbol rectangles resp.,
$P(\widetilde{p_i})$ is the ground truth rectangle $\widetilde{p_i}$ perimeter,
$||a, b||_2$ is the Euclidean distance between points a and b coordinates.

Also we will use the plate type detection quality functional Q_{type} defined as

$$Q_{type} = \begin{cases} 0 & t \neq \widetilde{t} \\ 1 & t = \widetilde{t}. \end{cases} \tag{11}$$

We will also estimate how accurately the suggested plate distortion model approximates the true projective transform (generated by plate rectangle vertices shift relatively to its true position). Let us introduce approximation accuracy Q_{approx} as the maximal among all plate symbols relative distance between the coordinates of true and detected symbol rectangle vertices.

$$Q_{approx} = \begin{cases} 0 & t \neq \widetilde{t} \\ \min_{i=0..n-1} A_i & t = \widetilde{t}, \end{cases} \tag{12}$$

where A_i is accuracy estimation for symbol region detection for the i-th plate symbol, defined as:

$$A_i = \max(0, 1 - \frac{\max_c(||\widetilde{c}, c||_2)}{P(\widetilde{p_i})/4}), \tag{13}$$

where c are the four rectangle vertices coordinates in the source image p_i coordinate system ($\widetilde{c}$ being ground truth coordinates, c - coordinates found by the algorithm).

To research the algorithm's robustness to plate localization imprecision we conducted several computational experiments (table 1).

We compared quality metrics Q_{type} (11), Q_{shift} (9), Q_{approx} (12) of the suggested algorithm with the algorithm similar to (Paliy et al. 2004), which optimizes the same quality functional (4) with the shift numer plate distortion model, i.e. (1)- (3), where in (2) $\Delta x_i = \Delta y_i = 0, \forall i$.

The experiment (1) demonstrates the algorithm's performance quality for a precisely localized plate rectangle (ground truth). Experiments 2-5 were conducted with the ground truth rectangle vertices distortion with normal random values with zero expectation and σ^2 variance shown in table 1. The dependence of mean approximation accuracy values Q_{approx} on σ for both models is shown in figure 9.

Also, to estimate the algorithm's performance in real conditions, it was implemented as a symbol segmentation and type selection module of the automatic license plate recognition system for images (MARINA) developed by our team (http://visillect.com/en/alpr) (experiment 6). The segmentation module in this system uses the results of a prior plate rectangle localization. The localization accuracy estimation (computed similarly to (13)) in this system is 94.2%.

Table 1: Experimental Results

Experiment	Model	Q_{type}	Q_{shift}	Q_{approx}
Marked data	Shift	97.15%	91.76%	91.51%
	Proposed	96.96%	88.16%	87.95%
Marked data with noise $\sigma = 0.005$	Shift	96.68%	90.17%	88.73%
	Proposed	96.99%	86.96%	85.70%
Marked data with noise $\sigma = 0.01$	Shift	91.60%	83.80%	81.16%
	Proposed	95.91%	85.95%	83.33%
Marked data with noise $\sigma = 0.015$	Shift	82.69%	74.01%	70.65%
	Proposed	91.89%	81.96%	78.06%
Marked data with noise $\sigma = 0.02$	Shift	73.65%	64.20%	60.28%
	Proposed	85.18%	75.25%	70.39%
Localization results	Shift	81.85%	71.90%	68.05%
	Proposed	88.67%	79.03%	74.50%

The figure 10 demonstrates the algorithm results for various σ values. The figure 11 shows algorithm operation examples in MARINA system.

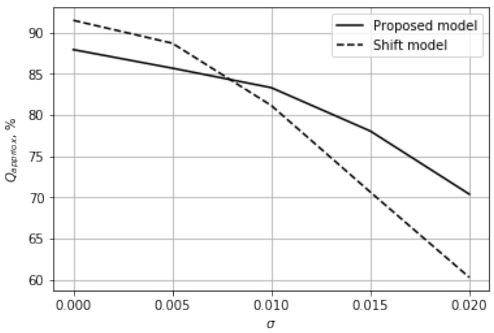
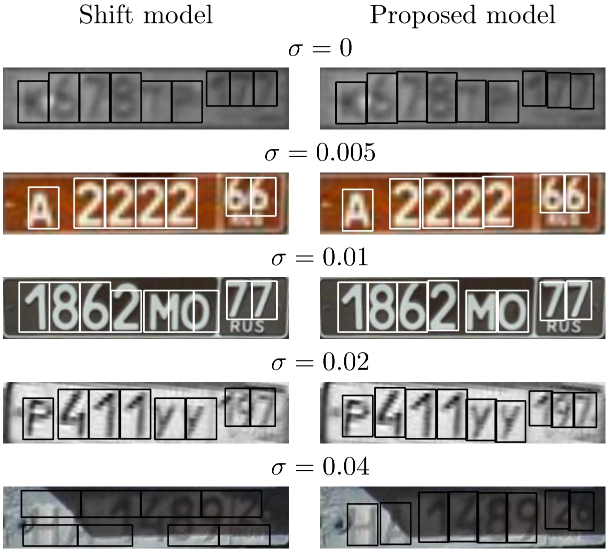


Figure 9. Dependence of Approximation Accuracy on Distortion Size


Shift model Proposed model

$\sigma = 0$

$\sigma = 0.005$

$\sigma = 0.01$

$\sigma = 0.02$

$\sigma = 0.04$

Figure 10. Examples of Segmentation Based on Shift and Proposed Distortion Models

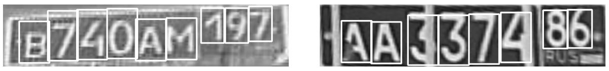


Figure 11. Examples of Segmentation Using MARINA Localization Results

According to experimental results, the proposed model demonstrates greater robustness to plate localization errors in comparison with the shift distortion model while having equal computational complexity. For $\sigma \leq 0.005$ the shift distortion model shows higher accuracy than the proposed one. However, with σ reaching 0.02, Q_{approx} decreases 4.7 times and Q_{type} drops 9.2 times for the shift distortion model whereas for the proposed model these numbers are 2.5 and 4.9 respectively. Moreover, our algorithm shows better accuracy by 13% for $\sigma = 0.01$, by 34% for $\sigma = 0.02$, and by 25% for MARINA localization results.

It may be inferred from the results, that the proposed model does not guarantee a precise approximation of the actual projective transform of the plate since $Q_{approx} < Q_{shift}$. This is due to to the restriction of symbol region resizing imposed by the model. However, the shift computation accuracy is usually more important than that of rotation and scale computation. In fact, segmentation results are used for character recognition, commonly solved with machine learning algorithms which can be trained to be robust to scaling and rotation. For example, with segmentation quality of 84% for plates with the correct type selected, a rate of successfully recognized plates is 97.7% (a fully-connected artificial neural network is used). As a further research, the model parameters might be approximated by projective distortion parameters in order to find more precise coordinates of symbol regions' corners.

In addition, a comparison with a well-known commercially available OCR engine ABBYY FineReader 14 (https://www.abbyy.com/finereader) was made. Since it does not consider plate types, a simplified functional Q_{length} was used for quality measurement:

$$Q_{length} = \begin{cases} 0 & n \neq \widetilde{n} \\ 1 & n = \widetilde{n}, \end{cases} \qquad (14)$$

where n is a number of characters acquired by an algorithm, excluding whitespaces, $\widetilde{n}$ is a ground truth number of characters. The mean Q_{length} value of 98% for proposed method on a subset of 247 marked samples was achieved, but it made only 33% for ABBYY FineReader 14 with all undesirable characters forbidden. The results show that ABBYY FineReader solution is inapplicable to the plate recognition without significant adaptation and it is reasonable to develop specific approaches like the proposed one.

The subset of 949 samples with $Q_{type} = 96\%$, $Q_{shift} = 86\%$, $Q_{approx} = 85.8\%$ on marked data for the proposed algorithm is available on ftp://vis.iitp.ru/license_plates.

CONCLUSIONS

In this paper, we propose a fast algorithm for plate segmentation based on information about plate type. The complexity of the algorithm linearly depends on image size. The algorithm is robust to plate quadrangle's localization errors and brightness distortions. High performance is achieved through approximation of quadrangle corners localization errors with a set of symbol regions shifts. We also propose an algorithm for plate type selection that allows to generalize the segmentation method for the case of multiple plate types. We provide experimental results demonstrating that our approach is more robust to plate localization errors in comparison with the known method which estimates the shift of the whole plate region.

ACKNOWLEDGEMENTS

The research was supported by the Russian Science Foundation grant (project No. 14-50-00150).

REFERENCES

Du, Shan; Mahmoud Ibrahim; Mohamed Shehata; and Wael Badawy. 2013. "Automatic license plate recognition (alpr): A state-of-the-art review." *IEEE Transactions on Circuits and Systems for Video Technology*, 23(2):311–325.

Gao, Qian; Xinnian Wang; and Gongfu Xie. 2007. "License plate recognition based on prior knowledge." In *Automation and Logistics, 2007 IEEE International Conference on*, 2964–2968.

Gosset, William Sealy. 1908. "The probable error of a mean." *Biometrika*, 6(1):1–25.

Kunina, IA; SA Gladilin; and DP Nikolaev. 2016. "Blind compensation of radial distortion in a single image using fast hough transform." *Computer Optics*, 40(3):395–403.

Kunina, IA; AP Terekhin; SA Gladilin; and DP Nikolaev. 2017. "Blind radial distortion compensation from video using fast hough transform." In *2016 International Conference on Robotics and Machine Vision*, 1025308–1025308.

Paliy, I; V Turchenko; V Koval; A Sachenko; and G Markowsky. 2004. "Approach to recognition of license plate numbers using neural networks." In *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, volume 4, 2965–2970.

Tian, Jiangmin; Ran Wang; Guoyou Wang; Jianguo Liu; and Yuanchun Xia. 2015. "A two-stage character segmentation method for chinese license plate." *Computers & Electrical Engineering*, 46:539–553.

Uddin, Md Azher; Joolekha Bibi Joolee; and Shayhan Ameen Chowdhury. 2016. "Bangladeshi vehicle digital license plate recognition for metropolitan cities using support vector machine." In *Proc. International Conference on Advanced Information and Communication Technology*.

Van Herk, Marcel. 1992. "A fast algorithm for local minimum and maximum filters on rectangular and octagonal kernels." *Pattern Recognition Letters*, 13(7):517–521.

Vintsyuk, Taras K. 1968. "Speech discrimination by dynamic programming." *Cybernetics and Systems Analysis*, 4(1):52–57.

Xia, Huadong and Dongchu Liao. 2011. "The study of license plate character segmentation algorithm based on vetical projection." In *Consumer Electronics, Communications and Networks (CECNet), 2011 International Conference on*, 4583–4586.

AUTHOR BIOGRAPHIES

MIKHAIL POVOLOTSKIY

was born in Moscow, Russia. He studied engineer science and mathematics, obtained his Bachelor degree in 2015 from Moscow Institute of Physics and Technology. Now he is a M.Sc. student. Since 2014 he works in Vision Systems Lab at the Institute for Information Transmission Problems RAS. His research activities are in the areas of computer vision. His e-mail address is mikhail.povolotskiy@iitp.ru.

ELENA KUZNETSOVA

was born in Zlatoust, Russia, studied engineering and mathematics at the National University of Science and Technology MISIS, graduated in 2013. Since 2013 she works as a junior researcher at the RAS Institute for Information Transmission Problems. Her research activities are in the areas of image processing, machine learning, development of computer vision systems. E-mail kuznetsova@iitp.ru.

TIMUR KHANIPOV

was born in St. Petersburg, Russia. He studied mathematics at the Moscow State University, graduated in 2008. Since 2010 he works as a researcher at the RAS Institute for Information Transmission Problems. Timur's research activities are in the areas of technical vision, industrial automation and recognition systems. His e-mail address is timur.khanipov@iitp.ru.

EVOLUTIONARY WINCH DESIGN USING AN ONLINE WINCH PROTOTYPING TOOL

Ibrahim A. Hameed,* Robin T. Bye,* Birger Skogeng Pedersen,** and Ottar L. Osen*,†
* Software and Intelligent Control Engineering Laboratory
* Department of ICT and Natural Sciences
** Mechatronics Laboratory
** Department of Ocean Operations and Civil Engineering
*,** NTNU, Norwegian University of Science and Technology
*,** Postboks 1517, NO-6025 Ålesund, Norway
† ICD Software AS
† Hundsværgata 8, NO-6008 Ålesund, Norway

KEYWORDS

Virtual prototyping; Product optimisation; Computer-automated design; Maritime winch design; Genetic algorithm; Particle swarm optimisation; Simulated annealing; Multi-objective optimisation using genetic algorithm; Artificial intelligence.

ABSTRACT

This paper extends the work of a concurrent paper on an intelligent winch prototyping tool (WPT) that is part of a generic and modular software framework for intelligent computer-automated product design. Within this framework, we have implemented a Matlab winch optimisation client (MWOC) that connects to the WPT and employs four evolutionary optimisation algorithms to optimise winch design. The four algorithms we employ are (i) a genetic algorithm (GA), (ii) particle swarm optimisation (PSO), (iii) simulated annealing (SA), and (iv) a multi-objective optimisation genetic algorithm (MOOGA). Here, we explore the capabilities of MWOC in a case study where we show that given a set of design guidelines and a suitable objective function based on these guidelines, we are able to optimise a particular winch design with respect to some desired design criteria. Our research has taken place in close cooperation with two maritime industrial partners, Seaonics AS and ICD Software AS, through two innovation and research projects on applying artificial intelligence for intelligent computer-automated design of maritime equipment such as offshore cranes and maritime winches.

INTRODUCTION

At the Software and Intelligent Control Engineering (SoftICE) Laboratory[1] at NTNU in Ålesund, virtual prototyping (VP) has been a key focus of research for several years. In particular, together with two industrial partners from the world-leading maritime industrial cluster in Norway, namely ICD Software AS[2] and Seaonics AS,[3] the SoftICE lab has recently completed two independent but related research projects on using artificial intelligence (AI) for intelligent computer-automated design (CautoD) of offshore cranes and maritime winches. In these projects, we have developed a generic and modular software framework for intelligent computer-automated design (CAutoD) of maritime cranes and winches (Bye et al., 2016) and examined how various intelligent evolutionary algorithms can be applied to optimise the design (Hameed, Bye, Osen, Pedersen and Schaathun, 2016; Hameed, Bye and Osen, 2016*a*,*b*). Here, we have implemented a Matlab winch optimisation client (MWOC) that connects with an online winch prototyping tool (WPT) that we present in an accompanying paper submitted concurrently to this conference (Bye et al., 2017).

In the remainder of this paper, we first provide some background on product design optimisation in general and VP of maritime winch systems in particular. We proceed with presenting our method, including a short overview of our product optimisation system, the MWOC, and the evolutionary algorithms that we have used. Then, we provide a case study of a winch design optimisation problem and present its results. Finally, we discuss our findings and directions for future work.

BACKGROUND

Product Design Optimisation

CAutoD revolves around the concept of *optimisation*, where the design problem is formulated as an optimisation problem where the best solution from all feasible solutions is determined by the optimisation of an objective function. In case of multiple criteria decision-making where there is a set of competing and conflicting objective functions, multi-objective optimisation (MOO) can also employed. To be able to perform such an optimisation, the design has to be broken down into a set of design parameters, and the goal of the designer is to determine suitable values for the design parameters such that the objective function is optimised. When the optimisation problem is difficult or impossible to solve using analytical or exact methods, a common solution is to apply population-based

Corresponding author: Ibrahim A. Hameed, ibib@ntnu.no.

[1] http://blog.hials.no/softice

[2] http://www.icdsoftware.no

[3] http://www.seaonics.com

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

evolutionary algorithms to get a satisfactory solution in a feasible time (e.g., see Zhang et al., 2011, for a survey).

Virtual Prototyping of Maritime Winch Systems

Maritime winches come in many shapes and uses, including trawling, anchor handling, mooring, towing, cranes of various sizes, and launch and recovery of remotely operated vehicles (ROVs). An example is depicted in Figure 1, which shows the Seaonics SCM-LARS, a new super compact launch and recovery system (LARS) that comes with a 3000 m umbilical on an electric winch with permanent magnet (PM) motors.

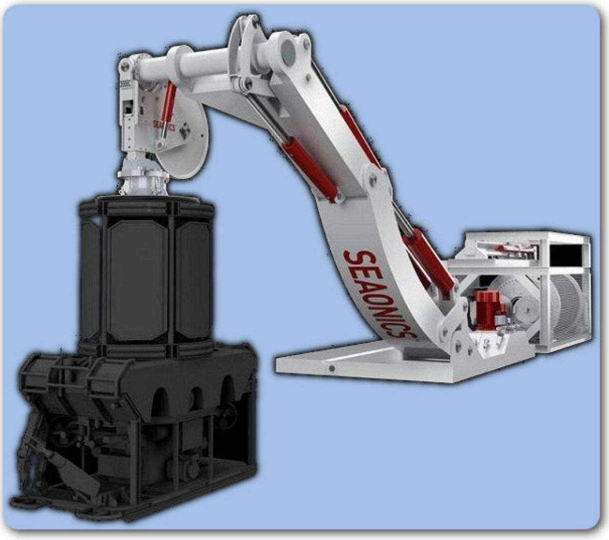


Figure 1: SCM-LARS with electric winch PM motors. Adapted from image courtesy of Seaonics AS.

As in our accompanying paper (Bye et al., 2017), we focus here on four important winch design components, namely the drum, electric motors, hydraulic motors, and the wire. Choosing appropriate values for design parameters related to these four components is necessary to yield winch designs that comply with desired design requirements and performance measures, or key performance indicators (KPIs). Typically, because many of these parameters are dependent on each other, human winch designers will have to iteratively try a large set of combinations of parameter settings to arrive at a satisfactory design (Pearlman et al., 2017). For more details about the design process of winches and the motivation for our work, please see Bye et al. (2017).

In this paper, we employ evolutionary algorithms in an attempt to automate this design optimisation process. We extend the case study we present in Bye et al. (2017) and are mainly concerned with torque performance as a KPI. Notably, however, many other concerns should be taken into account by the winch designer, such as manufacturing and operating costs, equipment weight, and adhering to laws, regulations, design codes, and standards.

METHOD

Product Optimisation System

In Bye et al. (2017), we present an overview of the client-server software architecture of our product optimisation system, reproduced in Figure 2 for convenience. The main component of the server is a product calculator, for example for offshore cranes or winches, that typically contains a large number of different and interdependent product design parameters with complex and nonlinear interactions. By setting different combinations of parameter values, the product calculator calculates a number of KPIs for each combination, which are effectively equivalent to different designs of the same product. Given a product design with desirable KPIs, the challenge is to determine the parameter values such that those KPIs goals being met.

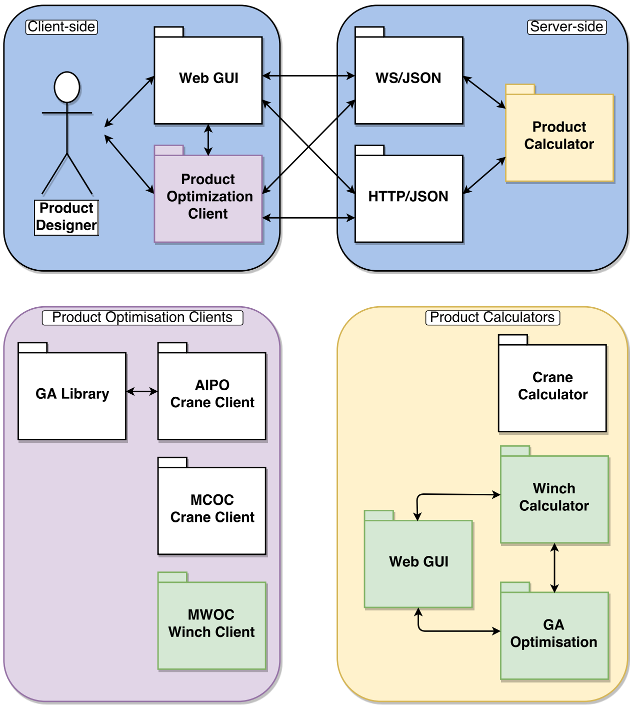


Figure 2: Software architecture for intelligent CautoD of offshore cranes, winches, or other products. Green boxes indicate work not presented previously. Adapted from Bye et al. (2017).

On the client-side, product designers can opt to use a web graphical user interface (GUI) to manually obtain suitable parameter values in the product calculator by trial-and-error, or implement a custom-made product optimisation client (POC) to automate the design process. Both the web GUI and POCs connect to the product calculator via the Hypertext Transfer Protocol (HTTP) or the WebSocket (WS) protocol and use JavaScript Object Notation (JSON). Consequently, the client can obtain KPIs, as well as other measures of interest, that result from setting various combinations of parameter values in the product calculator.

Here, we extend the case study presented in Bye et al. (2017) and implement a *winch* optimisation client in Matlab, the MWOC, that employs four different optimisation methods, namely GA, PSO, SA, and MOOGA.

Matlab Winch Optimisation Client (MWOC)

The MWOC module makes use of two libraries freely available from the MathWorks File Exchange[4] that were used for the WS/JSON interface, namely MatlabWebSocket, which is a simple library consisting of a websocket server and client for Matlab, and JSONlab, which is a toolbox to encode/decode JSON files in Matlab (Hameed, Bye and Osen, 2016*a*). For optimisation, we used a set of solvers for evolutionary optimisation algorithms available in the Global Optimization Toolbox (Mathworks, Inc., 2015), namely the GA Solver, the PSO Solver, the SA Solver, and the Multiobjective GA Solver.

The Genetic Algorithm (GA)

The GA (Holland, 1975) is the earliest, most well known, and probably most widely used evolutionary algorithm. Since its popularisation by Goldberg (1989), GAs have consistently served as an effective optimisation tool and therefore have become a very popular choice for optimisation problems in a variety of disciplines (e.g., see Haupt and Haupt, 2004; Simon, 2013).

A GA aims to exploit random search to solve optimisation problems and uses some genetic operators to direct the search into the region of better performance within the search space. Specifically, GAs have at least the following elements in common: a *population* of chromosomes (candidate solutions), *selection* according to fitness, *crossover* to produce new offspring, and random *mutation* of new offspring. Each iteration of this process is called a *generation*. A GA is iterated for a number of generations until a satisfactory solution is found. The entire set of generations is called a *run*.

Table 1 shows a summary of the GA parameter settings we used in our accompanying paper (Bye et al., 2017) and that we also use in this study.

setting	typical
candidates, N_c	100
parents, N_p	50
elites, N_e	10
mutations, N_m	10
generations, N_g	500

Table 1: GA settings with typical values.

Particle Swarm Optimisation (PSO)

PSO is a computational method that optimises a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. PSO is a population based stochastic optimisation technique inspired by social behaviour of bird flocking or fish schooling (Kennedy and Eberhart, 1995). Similar to GAs, PSO optimises a problem by having a population of candidate solutions, called *particles*, that constitutes a swarm moving around in the search space looking for the best solution according to a simple mathematical formula for the particles' position and velocity. Each particle's movement is influenced by its inertia, where each particle tends to maintain its velocity and direction; its local best known position (due to *personal influence/experience*), where a particle tends to return to its previous position if it is better than the current one; and finally by the global best known positions of its neighbour particles in the search-space (due to *social influence by neighbours*). Neighbours are defined topologically and are not based on the solution space. By following these simple rules, the swarm of particles will move toward the best solutions.

Table 2 shows a summary of the PSO settings we use in this paper.

setting	typical
swarm/population size, N_s	200
inertia range, W	[0.1 1.1]
min neighborhood fraction, σ	0.25
max iterations/generations, N_g	2000
max cognition learning rate, $\phi_{1,\max}$	1.49
max social learning rate , $\phi_{2,\max}$	1.49

Table 2: PSO settings with typical values.

Simulated Annealing (SA)

SA is an optimisation algorithm that mimics the cooling and crystallising behaviour of chemical substances (Kirkpatrick et al., 1983). The algorithm is initialised with a candidate solution x_0 to some minimisation problem and its *temperature* parameter, T_0, set to a high value so that the candidate solution is likely to change to some other configuration. It then randomly generates an alternative candidate solution x and measures its cost. If the cost of x is less than of x_0, the algorithm updates the candidate solution accordingly. Otherwise, it updates the candidate solution with a probability less than or equal to one. For every iteration, the temperature decreases, resulting in a tendency of the candidate solution to settle in a low-cost state. At the beginning of the algorithm, exploration is high and exploitation is low and vice versa at the end. A cooling schedule is used to control the rate of convergence.

Table 3 shows a summary of the SA settings used in this paper. Here, x_0 is row vector with the default values of the five optimisation parameters: (i) the drum inner radius, (ii) the gear ratio and (iii) quantity of electric motors, and (iv) the gear ratio and (v) quantity of hydraulic motors (more details in the Case Study section).

setting	typical
start point, x_0	[2.927 189.0 3 159.16 4]
reannealing interval, L	50
initial temperature, T_0	100
max iterations/generations, N_g	15000
cooling rate, α	0.95

Table 3: SA settings with typical values.

[4] http://www.mathworks.com/matlabcentral/fileexchange

Multi-Objective Optimisation GA (MOOGA)

Most realistic engineering problems have conflicting and competing multiple-objectives (MOs), for example, simultaneously minimising cost and maximising performance of a car engine. Combining individual objective functions into a single composite objective function, where each individual objective is weighted, is challenging and might not be realistic or even correct. An alternative approach is to determine an entire Pareto optimal solution set or a representative subset. A Pareto optimal set is a set of solutions that are non-dominated with respect to each other. While moving from one Pareto solution to another, there is always a certain amount of sacrifice in one objective to achieve a certain amount of gain in the other. Determining a set of Pareto solutions overcomes the problem of weight selection often used in the aforementioned composite objective functions used in traditional single objective GA.

CASE STUDY

We extend the case study in Bye et al. (2017), where we investigate a winch KPI consisting of the three torque profiles that result from a given winch design, namely the S1 continuous duty cycle[5] torque T_{S1}, the maximum torque T_{max}, and the required torque T_{req}. These torque profiles are all functions of the wire velocity v, which has a resolution of N_{r} sample points between zero and the maximum wire velocity v_{max}. As in Bye et al. (2017), we employ a resolution of $N_{\mathrm{r}} = 20$.

The S1 torque can be defined as the maximum continuous duty cycle with constant load that the electric motors can safely operate under; the maximum torque as an upper threshold at which the electric motors can safely operate under but only for shorter periods of time; and the required torque as the minimum torque required for safe operation for a given constant load.

As explained in Bye et al. (2017), the torque requirements of the winch increase with winch layers, and we therefore conservatively optimise with respect to the outermost layer 15.

Design Parameters

As for our work in Bye et al. (2017), we consider the 28 design parameters given in Table 4. The default values for each parameter yield the torque profiles depicted in Figure 3 and correspond to a baseline winch design obtained by a human operator at Seaonics AS and has not been subject to any optimisation.

As argued in Bye et al. (2017), we limit the winch design optimisation to only five parameters that influence the torque profiles, namely the inner diameter of the drum, and the gear ratios and quantities of the electric and hydraulic motors (shown in bold in Table 4).

For each design parameter to be optimised, we add constraints, that is, minimum and maximum allowed values. Table 4 summarises the parameter settings, including default, minimum, maximum, and optimised parameter values for each parameter p.

Objective Function

Our KPI of interest relates to the torque profiles of T_{req}, T_{max}, and T_{S1}. We adopt the same guidelines[6] as in Bye et al. (2017):

- T_{req} should be lower than T_{max} for all wire velocities v_k, except for standstill where $v_0 = 0$, where it could be allowed to be higher.
- T_{req} should be lower than T_{S1} at wire velocities used for continuous operation, that is, typically from half the maximum wire velocity v_{max} and higher.
- conversely, T_{req} should preferably lie between T_{max} and T_{S1} for wire velocities *not* used for continuous operation, that is, typically from half the maximum wire velocity v_{max} and lower.

These guidelines are intended to prevent ending up with a winch design with bigger and more expensive motors than necessary, leading to T_{req} being below T_{S1} for all velocities, including low velocities that are not suitable for continuous operation. This is not necessary, because it is possible to operate above the nominal S1 duty rating for shorter periods of time. Thus, In addition to reducing costs, this choice can improve weight and performance, since bigger motors will have higher mass and inertia. However, it must be kept in mind that after operating above the S1 duty cycle for a short period of time, the motors must be allowed to operate *below* S1 to avoid overheating.

We adopt the same cost function as in Bye et al. (2017) to achieve winch designs that adhere to our guidelines:

$$f_{\mathrm{cost}} = \sum_{k=1}^{N_{\mathrm{r}}} a \cdot \mathrm{R}_k + (1-a) \cdot \mathrm{S}_k \tag{1}$$

where

$$\mathrm{R}_k = \begin{cases} \delta \mathrm{R}_k^2 \text{ for } \delta \mathrm{R}_k > 0 \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

$$\mathrm{S}_k = \delta \mathrm{S}_k^2 \tag{3}$$

$$\delta \mathrm{R}_k = T_{\mathrm{req}}(v_k) - T_{\mathrm{max}}(v_k) \tag{4}$$

$$\delta \mathrm{S}_k = T_{\mathrm{req}}(v_k) - T_{\mathrm{S1}}(v_k) \tag{5}$$

$$a = 0.5 \tag{6}$$

and v_k is the kth sample of the wire velocity. For algorithms implemented using fitness functions such as the GA described previously, we simply negate the cost function:

$$f_{\mathrm{fitness}} = -f_{\mathrm{cost}} \tag{7}$$

Results

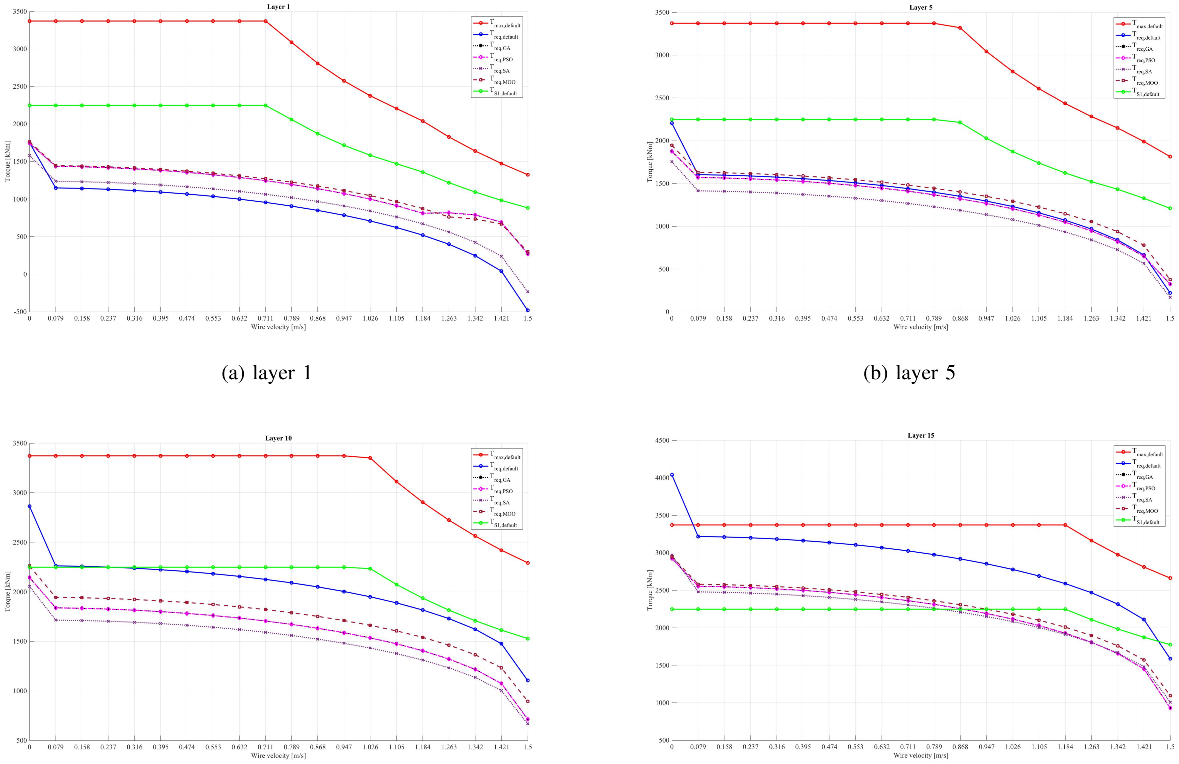
Figure 3 shows the torque profiles for T_{req} (blue), T_{S1} (green), and T_{max} (red) for winch layers 1, 5, 10, and 15, before optimisation (i.e., default design values) and after optimisation using GA ($T_{req,GA}$), PSO ($T_{req,PSO}$), SA ($T_{req,SA}$), and MOOGA ($T_{req,MOOGA}$), respectively.

[5] One of eight duty cycle classifications (S1–S8) provided by the International Electrotechnical Commission in the IEC 60034-1 standard.

[6] As per information provided by Seaonics AS.

subset	p	name	units	default	min	max	GA	PSO	SA	MOOGA
general	1	wavePeriodTime	s	10	-	-				
	2	load	kN	1500	-	-				
	3	iPlanetary	-	0.50	-	-				
	4	desiredWireVelocity	m/s	1.50	-	-				
drum	5	r	m	2.20	-	-				
	6	**innerDiameter**	**m**	**2.927**	**2.70**	**2.99**	**2.7007**	**2.7000**	**2.9702**	**2.8792**
	7	mass	kg	130000	-	-				
electric motor	**8**	**gearRatio**	-	**189.00**	**170**	**200**	**199.90**	**200.00**	**200.00**	**192.53**
	9	**motorQuantity**	-	**3**	**1**	**5**	**5**	**5**	**5**	**5**
	10	fieldWeakeningSpeed	rev/min	1440	-	-				
	11	maxSpeed	rev/min	2000	-	-				
	12	maxSpeedPower	kW	157.0	-	-				
	13	nominalPower	kW	210	-	-				
	14	nominalSpeed	rad/s	93.410	-	-				
	15	i	kg·m^2	9.00	-	-				
	16	nominalTorque	N·m	2248.15	-	-				
hydraulic motor	**17**	**gearRatio**	-	**159.16**	**150**	**190**	**189.99**	**190.00**	**158.58**	**188.83**
	18	**motorQuantity**	-	**4**	**1**	**5**	**3**	**3**	**4**	**3**
	19	friction	-	0.950	-	-				
	20	staticEfficiency	-	0.789	-	-				
	21	dynamicEfficiency	-	0.916	-	-				
	22	maxSpeed	rev/min	1600	-	-				
	23	displ	cm^3	1000	-	-				
	24	pressureDrop	bar	280	-	-				
	25	i	kg·m^2	0.5500	-	-				
	26	nominalTorque	N·m	4457.71	-	-				
wire	27	diameter	N·m	77.0	-	-				
	28	diameterReductionFactor	-	0.8660	-	-				
cost value	-	-	-	876900	-	-	784123	783757	805600	{0, 876900}

Table 4: Default and optimised winch designs using GA, PSO, SA, and MOOGA. Empty fields signify default values.



(a) layer 1

(b) layer 5

(c) layer 10

(d) layer 15

Figure 3: Torque profiles for T_{req} (blue), T_{S1} (green), and T_{max} (red) *before* optimisation and required torque after optimisation using GA ($T_{req,GA}$), PSO ($T_{req,PSO}$), SA ($T_{req,SA}$), and MOOGA ($T_{req,MOOGA}$), respectively, for winch layers 1, 5, 10, and 15.

Compared with the unsatisfactory default design in Figure 3, all optimisation methods result in very similar torque profiles for winch layer 15 that closely adhere to the guidelines presented previously, where T_{req} should lie below $T_{\max}$ and above T_{S1} for low wire velocities, and below T_{S1} (and $T_{\max}$) for high velocities. Moreover, for winch layer 10, all optimised torque profiles are highly satisfactory, with T_{req} below T_{S1} (and $T_{\max}$) for all wire velocities, which is a significant improvement from the default design. Finally, for winch layers 1 and 5, the torque profiles are similar to those of the default design, except that T_{req} is lower than T_{S1} for zero velocity. This is an improvement but not a requirement, since T_{req} is allowed to be higher than T_{S1} for short periods of time.

Whilst the qualitative examination of the torque profiles show that the desired relationships between T_{req}, T_{S1}, and $T_{\max}$ are met in a highly similar manner for all the optimisation methods, investigation of Table 4 indicate that there are several combinations of optimised parameter values that achieve the same desired performance. For example, using GA and PSO, the inner diameter of the drum is optimised to a value very close to the lower boundary value of 2.70 m, whereas using SA and MOOGA results in values of 2.97 (close to the upper boundary of 2.99) and 2.88, respectively.

Similarly, the gear ratio of the electric motors is optimised to a value very close to the upper boundary of 200 for GA, PSO, and SA, whereas MOOGA found the value of 192.53. For hydraulic motors, on the other hand, the gear ratio was optimised towards a value very close to the upper boundary of 190 for GA, PSO, and MOOGA, whereas the optimised value found by SA was 158.58, which is actually quite close to the default value of 159.16. This latter result may be due to the SA optimising the quantity of hydraulic motors to 4, whereas GA, PSO and MOOGA only used 3 (the default design used 4). Thus, the optimised design found by SA had one extra hydraulic motor compared with GA, PSO, and MOOGA, and therefore had less need for a high gear ratio.

Finally, all optimisation methods optimised the quantity of electric motors to 5, two more than the default design that only used 3 (which is probably a major contributor to the default design being unsatisfactory).

It is unsurprising that having extra electric motors help reducing the required torque T_{req} relative to the S1 duty cycle torque T_{S1}. Notably, however, all optimisation methods apart from SA simultaneously reduced the number of hydraulic motors to 3 from the default quantity of 4, to avoid excess torque capabilities as laid out in the guidelines.

Both GA and PSO found an optimised inner diameter of the drum very close to the lower constraint of 2.70 m. This makes intuitive sense, since a smaller diameter has lower torque requirements. But why does the SA and MOOGA have higher optimised diameter values towards the upper constraint then? We believe the answer may be twofold. First, increasing the inner diameter of the drum may ensure that the default speed requirements of the motors are satisfied, because the wire velocity increases with the diameter of the drum for the same rotational speed. Second, SA uses one extra hydraulic motor, whereas MOOGA has a very high gear ratio for the hydraulic motors. Both parameter values will improve the torque profiles, thus possibly dominating, or voiding the need for a low inner diameter of the drum.

Finally, we observe that GA and PSO has the greatest reduction in cost function evaluation when compared with the default design (bottom line in Table 4. Moreover, we did not get a Pareto optimal set because the objective functions are not competing.

We conclude from these results that the evolutionary optimisation algorithms have found different combinations of optimised parameter values (winch designs) that all yield very similar torque profiles. The optimised parameter values improve an unsatisfactory default winch design to yield designs that are satisfactory, but not excessively good, and adhere to the design guidelines we set out previously.

DISCUSSION

This paper extends our case study on winch optimisation (Bye et al., 2017) with a set of four evolutionary optimisation algorithms; the GA, PSO, SA, and MOOGA. The algorithms are implemented in a Matlab winch optimisation client, the MWOC, that connects to an online winch prototyping tool, the WPT, thus demonstrating practical use of our software framework for intelligent product design.

For the case study, we set out to improve a default winch design that did not satisfy the desired performance, or KPI, as described qualitatively in a set of design guidelines. These guidelines describe the desired relationship between the required torque T_{req}, the S1 duty cycle torque T_{S1}, and the maximum torque $T_{\max}$, which all are functions of the wire velocity. In Bye et al. (2017), we devised a cost function derived from these guidelines with the intention of using the aforementioned evolutionary algorithms to determine optimised parameter values that should result in satisfactory winch designs. Obviously, simply adding a large number of motors would result in highly proficient torque profiles. However, doing so would result in winch designs that are better than needed and also are very costly. Intelligently, the evolutionary optimisation algorithms we have employed here are able to strike a sweet spot between the weak default winch design and too powerful winch designs that massively exceed the desired torque performance.

Our work demonstrates that employing evolutionary algorithms in our product optimisation system can be useful for CAutoD of maritime equipment such as offshore winches and cranes but also for any other product for which the design can be parameterised and the design goal is to determine a suitable set of parameter values for which the resulting design satisfy some desired design criteria, or KPIs.

Future Work

A limitation of our work is that we have assumed that the parameters chosen for optimisation can take on any value within some pre-defined boundaries (minimum and maximum values). However, when our industrial partner Seaonics AS are doing their winch design, they need to pick suitable *combinations* of components (sets of fixed parameter values) that are commercially available and meet the clients' needs. Seaonics AS is currently in the process of providing us with an extensive set of real-world manufactured components that we could use to build up this library. After this, we will update our system and let Seaonics's professional winch designers test our software. During this process, we will try to identify and correct possible shortcomings with the potential of subsequently adopting the software for common use, such as developing new objective functions for MOO, adding user's guides, developing new application-dependent design guidelines with corresponding objective functions for optimisation, improving the web graphical user interface (GUI) of the WPT, and more. For more details about our work and future research, we refer to our accompanying paper submitted in parallel (Bye et al., 2017).

ACKNOWLEDGEMENTS

The SoftICE lab at NTNU in Ålesund wishes to thank ICD Software AS for their contribution in the software development process, and Seaonics AS for providing documentation and insight into the design and manufacturing process of offshore cranes. We are also grateful for the support provided by Regionalt Forskningsfond (RFF) Midt-Norge and the Research Council of Norway through the VRI research projects *Artificial Intelligence for Crane Design (Kunstig intelligens for krandesign (KIK))*, grant no. 241238, and *Artificial Intelligence for Winch Design (Kunstig intelligens for vinsjdesign (KIV))*, grant no. 249171.

REFERENCES

Bye, R. T., Hameed, I. A., Pedersen, B. S. and Osen, O. L. (2017), An intelligent winch prototyping tool, Proceedings of the 31st European Conference on Modelling and Simulation (ECMS '17). Under review.

Bye, R. T., Osen, O. L., Pedersen, B. S., Hameed, I. A. and Schaathun, H. G. (2016), A software framework for intelligent computer-automated product design, Proceedings of the 30th European Conference on Modelling and Simulation (ECMS '16), pp. 534–543.

Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Professional.

Hameed, I. A., Bye, R. T. and Osen, O. L. (2016*a*), A Comparison between Optimization Algorithms Applied to Offshore Crane Design using an Online Crane Prototyping Tool, Proceedings of the 2nd International Conference on Advanced Intelligent Systems and Informatics (AISI '16), Vol. 533 of *Advances in Intelligent Systems and Computing (AISC)*, pp. 266–276.

Hameed, I. A., Bye, R. T. and Osen, O. L. (2016*b*), Grey wolf optimizer (GWO) for automated offshore crane design, Proceedings of the IEEE Symposium Series on Computational Intelligence (IEEE SSCI '16), pp. 1–6.

Hameed, I. A., Bye, R. T., Osen, O. L., Pedersen, B. S. and Schaathun, H. G. (2016), Intelligent computer-automated crane design using an online crane prototyping tool, Proceedings of the 30th European Conference on Modelling and Simulation (ECMS '16), pp. 564–573.

Haupt, R. L. and Haupt, S. E. (2004), *Practical Genetic Algorithms*, 2nd edn, Wiley.

Holland, J. H. (1975), *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*, University of Michigan Press, Oxford, England.

Kennedy, J. and Eberhart, R. C. (1995), Particle swarm optimization, Proceedings of the IEEE International Conference on Neural Networks, Vol. IV, IEEE, IEEE service center, Piscataway, NJ, pp. 1942–1948.

Kirkpatrick, S., Gelatt, C. D. and Vecchi, M. P. (1983), Optimization by Simulated Annealing, *Science* 220(4598), 671–680.

Mathworks, Inc. (2015), *MATLAB Global Optimization Toolbox, R2015b*, The Mathworks, Inc., Natick, Massachusetts.

Pearlman, S. M., Gordon, D. R. and Pearlman, M. D. (2017), Winch Technology — Past Present and Future: A Summary of Winch Design Principles and Developments, Technical report, InterOcean Systems LLC. Accessed on 14 February 2017 from http://www.interoceansystems.com/winch_article.htm.

Simon, D. (2013), *Evolutionary Optimization Algorithms: Biologically Inspired and Population-Based Approaches to Computer Intelligence*, John Wiley & Sons, Inc., Hoboken, New Jersey.

Zhang, J., Zhan, Z., Lin, Y., Chen, N., jiao Gong, Y., hui Zhong, J., Chung, H., Li, Y. and Shi, Y. (2011), Evolutionary computation meets machine learning: A survey, *IEEE Computational Intelligence Magazine* 6(4), 68–75.

AUTHOR BIOGRAPHIES

IBRAHIM A. HAMEED is an IEEE Senior Member and has a BSc and a MSc in Industrial Electronics and Control Engineering, Menofia University, Egypt, a PhD in Industrial Systems and Information Engineering from Korea University, S. Korea, and a PhD in Mechanical Engineering, Aarhus University, Denmark. He has been working as an associate professor at NTNU Ålesund since 2015. His research interests includes artificial intelligence, optimisation, control systems, and robotics.

ROBIN T. BYE[7] is an IEEE Senior Member who graduated from the University of New South Wales, Sydney with a BE (Hons 1), MEngSc, and a PhD, all in electrical engineering. Working at NTNU in Ålesund since 2008, Dr. Bye is an associate professor in automation engineering. His research interests belong to the fields of artificial intelligence, cybernetics, neuroengineering, and education.

BIRGER SKOGENG PEDERSEN graduated from NTNU Ålesund with a BE in automation engineering and is a former employee at ICD Software AS. He is currently a MSc student of simulation and visualisation and employed as a research assistant in the Mechatronics Laboratory at NTNU Ålesund.

OTTAR L. OSEN is MSc in Engineering Cybernetics from the Norwegian Institute of Technology in 1991. He is the head of R&D at ICD Software AS and an assistant professor at NTNU Ålesund.

[7] www.robinbye.com

SHADE MUTATION STRATEGY ANALYSIS VIA DYNAMIC SIMULATION IN COMPLEX NETWORK

Adam Viktorin
Roman Senkerik
Michal Pluhacek
Tomas Kadavy
Tomas Bata University in Zlin, Faculty of Applied Informatics
Nam T.G. Masaryka 5555, 760 01 Zlin, Czech Republic
{aviktorin, senkerik, pluhacek, kadavy}@fai.utb.cz

KEYWORDS
Differential Evolution, SHADE, Complex Network, Mutation

ABSTRACT

This paper presents a novel approach to visualizing Evolutionary Algorithm (EA) dynamic in complex network and analyses the greediness of "current-to-*p*best/1" mutation strategy used in state-of-art Differential Evolution (DE) algorithm – Success-History based Adaptive DE (SHADE) on CEC2015 benchmark set of test functions. Provided analysis suggests that the greediness might not be the optimal approach for guiding the evolution.

INTRODUCTION

Differential Evolution (DE) is an algorithm for numerical optimization, which was introduced to the world by Storn and Price in 1995 (Storn and Price 1995). Its ingenuity and simplicity made it an Evolutionary Algorithm (EA) with one of the largest research community and therefore, continuous research provided various improvements to the original algorithm that led to one of the best performing EA branches. DE research was summarized in (Neri and Tirronen 2010; Das et al. 2016). Over the last decade, variants of the DE algorithm have won numerous numerical optimization competitions (Brest et al. 2006; Qin et al. 2009; Das et al. 2009; Mininno et al. 2011; Mallipeddi et al. 2011; Brest et al. 2013; Tanabe and Fukunaga 2014) and the common denominator since 2013 is Success-History based Adaptive DE (SHADE), which is an algorithm developed by Tanabe and Fukunaga (Tanabe and Fukunaga 2013). SHADE is based on the JADE algorithm by Zhang and Sanderson (Zhang and Sanderson 2009) and belongs to the family of adaptive DE algorithms. Adaptive algorithms are algorithms that adapt their control parameters during the optimization process and do not require user to set them at the beginning, thus these algorithms are more robust and do not require fine-tuning of the control parameters for given problem.

SHADE algorithm uses the "current-to-*p*best/1" mutation strategy inherited from JADE, which combines four individuals in the population. One of them is selected greedily from the smaller subset of best individuals, in terms of objective function value, in the current population.

The dynamic of the SHADE algorithm is in this paper transformed into the complex network and the greedy behavior of the mutation strategy is analyzed with the help of complex network feature – node centrality value. Whereas in previous work (Viktorin et al. 2016; Pluhacek et al. 2016) edges in complex network were undirected and unweighted, this paper proposes an approach with weighted and directed edges to better capture the dynamic of the heuristic. Weights are based on the individual donations in mutation and crossover and directed from the donator to the beneficiary. This approach is tested on CEC2015 benchmark set of 15 test functions and the results of the analysis are provided and discussed.

The remainder of the paper is structured as follows: Next two sections describe DE and SHADE algorithms, section that follows is about complex network design, experiment setting and results are provided in two following sections and the paper is concluded in the last section.

DIFFERENTIAL EVOLUTION

The DE algorithm is initialized with a random population of individuals $\boldsymbol{P}$, that represent solutions of the optimization problem. The population size NP is set by the user along with other control parameters – scaling factor F and crossover rate CR. In continuous optimization, each individual is composed of a vector $\boldsymbol{x}$ of length D, which is a dimensionality (number of optimized attributes) of the problem, where each vector component represents a value of the corresponding attribute, and the individual also contains the objective function value $f(\boldsymbol{x})$. For each individual in a population, three mutually different individuals are selected for mutation of vectors and resulting mutated vector $\boldsymbol{v}$ is combined with the original vector $\boldsymbol{x}$ in crossover step. The objective function value $f(\boldsymbol{u})$ of the resulting trial vector $\boldsymbol{u}$ is evaluated and compared to that of the original individual. When the quality (objective function value) of the trial individual is better, it is placed into

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

the next generation, otherwise, the original individual is placed there. This step is called selection. The process is repeated until the stopping criterion is met (e.g. the maximum number of objective function evaluations, the maximum number of generations, the low bound for diversity between objective function values in population).

Initialization

As aforementioned, the initial population $\boldsymbol{P}$, of size *NP*, of individuals is randomly generated. For this purpose, the individual vector $\boldsymbol{x}_i$ components are generated by Random Number Generator (RNG) with uniform distribution from the range which is specified for the problem by ***lower*** and ***upper*** bounds (1).

$$x_{j,i} = U\left[lower_j, upper_j\right] \text{ for } j = 1, \dots, D \tag{1}$$

where *i* is the index of a current individual, *j* is the index of current attribute and *D* is the dimensionality of the problem.

In the initialization phase, the scaling factor value *F* and the crossover rate value *CR* has to be assigned as well. The typical range for *F* value is [0, 2] and for *CR*, it is [0, 1].

Mutation

In the mutation step, three mutually different individuals $\boldsymbol{x}_{r1}$, $\boldsymbol{x}_{r2}$, $\boldsymbol{x}_{r3}$ are randomly selected from a population and combined in mutation according to the mutation strategy. The original mutation strategy of canonical DE is “rand/1” and is depicted in (2).

$$\boldsymbol{v}_i = \boldsymbol{x}_{r1} + F(\boldsymbol{x}_{r2} - \boldsymbol{x}_{r3}) \tag{2}$$

where *r1* ≠ *r2* ≠ *r3* ≠ *i*, *F* is the scaling factor value and $\boldsymbol{v}_i$ is the resulting mutated vector.

Crossover

In the crossover step, mutated vector $\boldsymbol{v}_i$ is combined with the original vector $\boldsymbol{x}_i$ and they produce trial vector $\boldsymbol{u}_i$. The binomial crossover (3) is used in canonical DE.

$$u_{j,i} = \begin{cases} v_{j,i} & \text{if } U[0,1] \leq CR \text{ or } j = j_{rand} \\ x_{j,i} & \text{otherwise} \end{cases} \tag{3}$$

where *CR* is the used crossover rate value and j_{rand} is an index of an attribute that has to be from the mutated vector $\boldsymbol{v}_i$ (ensures generation of a vector with at least one new component).

Selection

The selection step ensures, that the optimization progress will lead to better solutions because it allows only individuals of better or at least equal objective function value to proceed into the next generation *G+1* (4).

$$\boldsymbol{x}_{i,G+1} = \begin{cases} \boldsymbol{u}_{i,G} & \text{if } f(\boldsymbol{u}_{i,G}) \leq f(\boldsymbol{x}_{i,G}) \\ \boldsymbol{x}_{i,G} & \text{otherwise} \end{cases} \tag{4}$$

where *G* is the index of current generation.

The whole DE algorithm is depicted in pseudo-code below.

Algorithm pseudo-code 1: DE

```
Set $NP$, $CR$, $F$ and stopping criterion;
$G = 0$, $\boldsymbol{x}_{best} = \{\}$;
Randomly initialize (1) population $\boldsymbol{P} = (\boldsymbol{x}_{1,G}, \dots, \boldsymbol{x}_{NP,G})$;
$\boldsymbol{P}_{new} = \{\}$, $\boldsymbol{x}_{best}$ = best from population $\boldsymbol{P}$;
while stopping criterion not met
  for $i = 1$ to $NP$ do
    $\boldsymbol{x}_{i,G} = \boldsymbol{P}[i]$;
    $\boldsymbol{v}_{i,G}$ by mutation (2);
    $\boldsymbol{u}_{i,G}$ by crossover (3);
    if $f(\boldsymbol{u}_{i,G}) < f(\boldsymbol{x}_{i,G})$ then
      $\boldsymbol{x}_{i,G+1} = \boldsymbol{u}_{i,G}$;
    else
      $\boldsymbol{x}_{i,G+1} = \boldsymbol{x}_{i,G}$;
    end
    $\boldsymbol{x}_{i,G+1} \rightarrow \boldsymbol{P}_{new}$;
  end
  $\boldsymbol{P} = \boldsymbol{P}_{new}$, $\boldsymbol{P}_{new} = \{\}$, $\boldsymbol{x}_{best}$ = best from population $\boldsymbol{P}$;
end
return $\boldsymbol{x}_{best}$ as the best found solution
```

SUCCESS-HISTORY BASED ADAPTIVE DIFFERENTIAL EVOLUTION

In SHADE algorithm, the only control parameter that can be set by the user is the population size *NP*. Other two (*F, CR*) are adapted to the given optimization task, and a new parameter *H* is introduced, which determines the size of *F* and *CR* value memories. The initialization step of the SHADE is, therefore, similar to DE. Mutation, however, is completely different because of the used strategy “current-to-*p*best/1” and the fact, that it uses different scaling factor value F_i for each individual. Crossover is still binomial, but similarly to the mutation and scaling factor values, crossover rate value CR_i is also different for each individual. The selection step is the same and therefore following sections describe only different aspects of initialization, mutation and crossover steps.

Initialization

As aforementioned, initial population $\boldsymbol{P}$ is randomly generated as in DE, but additional memories for *F* and *CR* values are initialized as well. Both memories have the same size *H* and are equally initialized. The memory for *CR* values is titled $\boldsymbol{M}_{CR}$ and the memory for *F* is titled $\boldsymbol{M}_F$. Their initialization is depicted in (5).

$$M_{CR,i} = M_{F,i} = 0.5 \text{ for } i = 1, \dots, H \tag{5}$$

Also, the external archive of inferior solutions $\boldsymbol{A}$ is initialized. Since there are no solutions so far, it is initialized empty $\boldsymbol{A} = \emptyset$ and its maximum size is set to *NP*.

Mutation

Mutation strategy "current-to-*p*best/1" was introduced in (Zhang and Sanderson 2009) and unlike "rand/1", it combines four mutually different vectors, therefore $pbest \neq r1 \neq r2 \neq i$ (6).

$$\boldsymbol{v}_i = \boldsymbol{x}_i + F_i(\boldsymbol{x}_{pbest} - \boldsymbol{x}_i) + F_i(\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2}) \tag{6}$$

where $\boldsymbol{x}_{pbest}$ is randomly selected from the best $NP \times p$ best individuals in the current population. The *p* value is randomly generated for each mutation by RNG with uniform distribution from the range $[p_{min}, 0.2]$ (Tanabe and Fukunaga, 2013), where $p_{min} = 2/NP$. Vector $\boldsymbol{x}_{r1}$ is randomly selected from the current population and vector $\boldsymbol{x}_{r2}$ is randomly selected from the union of current population $\boldsymbol{P}$ and archive $\boldsymbol{A}$. The scaling factor value F_i is given by (7).

$$F_i = C[M_{F,r}, 0.1] \tag{7}$$

where $M_{F,r}$ is a randomly selected value (by index *r*) from $\boldsymbol{M}_F$ memory and *C* stands for Cauchy distribution. Therefore, the F_i value is generated from the Cauchy distribution with location parameter value $M_{F,r}$ and scale parameter value 0.1. If the generated value $F_i > 1$, it is truncated to 1 and if it is $F_i \leq 0$, it is generated again by (7).

Crossover

Crossover is the same as in (3), but the *CR* value is changed to CR_i, which is generated separately for each individual (8). The value is generated from the Gaussian distribution with mean parameter value of $M_{CR,r}$, which is randomly selected (by the same index *r* as in mutation) from $\boldsymbol{M}_{CR}$ memory and standard deviation value of 0.1.

$$CR_i = N[M_{CR,r}, 0.1] \tag{8}$$

Historical Memory Updates

Historical memories $\boldsymbol{M}_F$ and $\boldsymbol{M}_{CR}$ are initialized according to (5), but their components change during the evolution. These memories serve to hold successful values of *F* and *CR* used in mutation and crossover steps. Successful in terms of producing trial individual better than the original individual. During one generation, these successful values are stored in corresponding arrays $\boldsymbol{S}_F$ and $\boldsymbol{S}_{CR}$. After each generation, one cell of $\boldsymbol{M}_F$ and $\boldsymbol{M}_{CR}$ memories is updated. This cell is given by the index *k*, which is initialized to 1 and increases by 1 after each generation. When *k* overflows the size limit of memories *H*, it is reset to 1. The new value of *k*-th cell for $\boldsymbol{M}_F$ is calculated by (9) and for $\boldsymbol{M}_{CR}$ by (10).

$$M_{F,k} = \begin{cases} \text{mean}_{WL}(\boldsymbol{S}_F) & \text{if } \boldsymbol{S}_F \neq \emptyset \\ M_{F,k} & \text{otherwise} \end{cases} \tag{9}$$

$$M_{CR,k} = \begin{cases} \text{mean}_{WA}(\boldsymbol{S}_{CR}) & \text{if } \boldsymbol{S}_{CR} \neq \emptyset \\ M_{CR,k} & \text{otherwise} \end{cases} \tag{10}$$

where mean$_{WL}$() and mean$_{WA}$() are weighted Lehmer (11) and weighted arithmetic (12) means correspondingly.

$$\text{mean}_{WL}(\boldsymbol{S}_F) = \frac{\sum_{k=1}^{|\boldsymbol{S}_F|} w_k \cdot S_{F,k}^2}{\sum_{k=1}^{|\boldsymbol{S}_F|} w_k \cdot S_{F,k}} \tag{11}$$

$$\text{mean}_{WA}(\boldsymbol{S}_{CR}) = \sum_{k=1}^{|\boldsymbol{S}_{CR}|} w_k \cdot S_{CR,k} \tag{12}$$

where the weight vector $\boldsymbol{w}$ is given by (13) and is based on the improvement in objective function value between trial and original individuals.

$$w_k = \frac{\text{abs}\left(f(\boldsymbol{u}_{k,G}) - f(\boldsymbol{x}_{k,G})\right)}{\sum_{m=1}^{|\boldsymbol{S}_{CR}|} \text{abs}\left(f(\boldsymbol{u}_{m,G}) - f(\boldsymbol{x}_{m,G})\right)} \tag{13}$$

And since both arrays $\boldsymbol{S}_F$ and $\boldsymbol{S}_{CR}$ have the same size, it is arbitrary which size will be used for the upper boundary for *m* in (13). Complete SHADE algorithm is depicted in pseudo-code below.

Algorithm pseudo-code 2: SHADE

```
Set $NP$, $H$ and stopping criterion;
$G = 0$, $x_{best} = \{\}$, $k = 1$, $p_{min} = 2/NP$, $\boldsymbol{A} = \emptyset$;
Randomly initialize (1) population $\boldsymbol{P} = (\boldsymbol{x}_{1,G}, \dots, \boldsymbol{x}_{NP,G})$;
Set $\boldsymbol{M}_F$ and $\boldsymbol{M}_{CR}$ according to (5);
$\boldsymbol{P}_{new} = \{\}$, $\boldsymbol{x}_{best}$ = best from population $\boldsymbol{P}$;
while stopping criterion not met
  $\boldsymbol{S}_F = \emptyset$, $\boldsymbol{S}_{CR} = \emptyset$;
  for $i = 1$ to $NP$ do
    $\boldsymbol{x}_{i,G} = \boldsymbol{P}[i]$;
    $r = U[1, H]$, $p_i = U[p_{min}, 0.2]$;
    Set $F_i$ by (7) and $CR_i$ by (8);
    $\boldsymbol{v}_{i,G}$ by mutation (6);
    $\boldsymbol{u}_{i,G}$ by crossover (3);
    if $f(\boldsymbol{u}_{i,G}) < f(\boldsymbol{x}_{i,G})$ then
      $\boldsymbol{x}_{i,G+1} = \boldsymbol{u}_{i,G}$;
      $\boldsymbol{x}_{i,G} \rightarrow \boldsymbol{A}$;
      $F_i \rightarrow \boldsymbol{S}_F$, $CR_i \rightarrow \boldsymbol{S}_{CR}$;
    else
      $\boldsymbol{x}_{i,G+1} = \boldsymbol{x}_{i,G}$;
    end
```

```
        if $|A| > NP$ then randomly delete an ind. from $A$;
        $x_{i,G+1} \rightarrow P_{new}$;
    end
    if $S_F \neq \emptyset$ and $S_{CR} \neq \emptyset$ then
        Update $M_{F,k}$ (9) and $M_{CR,k}$ (10), $k$++;
        if $k > H$ then $k$ = 1, end;
    end
    $P = P_{new}$, $P_{new}$ = {}, $x_{best}$ = best from population $P$;
end
return $x_{best}$ as the best found solution
```

NETWORK DESIGN

The network is created for each generation of the SHADE algorithm, where the key factors are mutation and crossover steps. Individuals in population are nodes in the network and edges between them are created if the trial individual produced in mutation and crossover succeeds in selection (has better objective function value than the original individual). Each individual has its own ID, trial individual $\boldsymbol{u}_i$ inherits ID of the original individual $\boldsymbol{x}_i$ and therefore, the original individual and trial individual are represented in the network by the same node. Since there are four individuals acting in mutation and crossover ($\boldsymbol{x}_i$, $\boldsymbol{x}_{pbest}$, $\boldsymbol{x}_{r1}$ and $\boldsymbol{x}_{r2}$), four edges ($e_{i,i}$, $e_{pbest,i}$, $e_{r1,i}$, $e_{r2,i}$) are created for each successful selection. Each edge is directed from source to target (donor to beneficiary) and denoted by $e_{source,target}$. First edge is a self-loop. Each edge has also its weight w, which is denoted w_{source} and is based on the ratios of donations to the trial individual. If the edge already exists in the network, new weight is added to the current one. Edges with their weights are depicted in Figure 1.

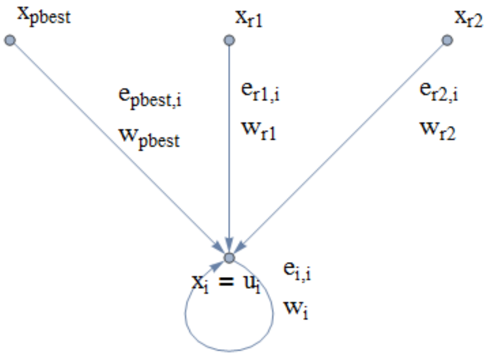


Figure 1: Edges Created for One Successful Evolution of an Individual

Original individual $\boldsymbol{x}_i$ donates during crossover and also in mutation. The number of donated attributes in crossover by $\boldsymbol{x}_i$ is divided by dimension of the problem D and the final ratio CR_r (real crossover) is subtracted from 1. This is done, because the sum of all four weights should be 1, therefore, the cumulative value left for the mutation donations is CRr. This value is divided between four individuals according to their ratio in (6). Ratios for each individual in mutation are:

- $\boldsymbol{x}_i \rightarrow 1 - F_i$
  $$\boldsymbol{v}_i = \boldsymbol{x}_i + F_i(\boldsymbol{x}_{pbest} - \boldsymbol{x}_i) + F_i(\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2})$$
- $\boldsymbol{x}_{pbest} \rightarrow F_i$
  $$\boldsymbol{v}_i = \boldsymbol{x}_i + F_i(\boldsymbol{x}_{pbest} - \boldsymbol{x}_i) + F_i(\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2})$$
- $\boldsymbol{x}_{r1} \rightarrow F_i$
  $$\boldsymbol{v}_i = \boldsymbol{x}_i + F_i(\boldsymbol{x}_{pbest} - \boldsymbol{x}_i) + F_i(\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2})$$
- $\boldsymbol{x}_{r2} \rightarrow -F_i$
  $$\boldsymbol{v}_i = \boldsymbol{x}_i + F_i(\boldsymbol{x}_{pbest} - \boldsymbol{x}_i) + F_i(\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2})$$

And each ratio is multiplied by CR_r and divided by the sum of ratios, which is 1 to obtain the proportion of CR_r as a weight. Resulting weights are depicted in (14, 15, 16 and 17).

$$w_i = (1 - CR_r) + CR_r * (1 - F_i) \quad (14)$$

$$w_{pbest} = CR_r * F_i \quad (15)$$

$$w_{r1} = CR_r * F_i \quad (16)$$

$$w_{r2} = CR_r * (-F_i) \quad (17)$$

where w_i sums two components – crossover donation and mutation donation. For example, if the dimensionality of the problem D = 10, scaling factor for this mutation F_i = 0.7 and 2 attributes are taken from the original individual $\boldsymbol{x}_i$ in crossover, then:

- $CR_r = 2 / 10 = 0.2$
- $w_i = (1 – 0.2) + 0.2 * (1 – 0.7) = 0.86$
- $w_{pbest} = 0.2 * 0.7 = 0.14$
- $w_{r1} = 0.2 * 0.7 = 0.14$
- $w_{r2} = 0.2 * (-0.7) = -0.14$
- $\sum w = 0.86 + 0.14 + 0.14 – 0.14 = 1$

EXPERIMENT SETTING

In order to obtain an analysis of the greedy behavior of the mutation in SHADE algorithm, complex networks were created for each generation in a SHADE run and node centrality value (sum of weights of outgoing edges from a node) for each individual in a population was recorded. This was done for 15 test functions in CEC2015 benchmark set and each test function was run 51 times with random initialization in 10D. The stopping criterion was set to 10,000×D objective function evaluations. The population size was set to NP = 100 and historical memory size was set to H = 10.

The basic assumption is that the nodes, that communicate the most (have high node centrality value) are the ones who lead the evolution towards the global optima. The greediness in “current-to-*p*best/1” mutation strategy is represented by $\boldsymbol{x}_{pbest}$ individual, which is selected from a subset of best individuals in the population *pbest* (the size of this subset varies from 2 individuals to a maximum of 20% of the population size). In theory, the *pbest* subset should contain individuals leading the optimization, therefore these individuals should correspond to the most active ones in the network. In order to test that, a new metric *centrRank* was proposed. An auxiliary variable *centrPosition* is introduced and corresponds to the number of individuals in the population that have worse node centrality value than the current individual (e.g. If the node has the highest centrality, than the *centrPosition* value would be 99 as the number of individuals is 100. On the other hand, if it is an individual with worst centrality value, *centrPosition* will be 0). The *centrRank* value is rescaled *centrPosition* to the range <0, 1>, where *centrPosition* = 99 translates to *centrRank* = 1.

Values of *centrRank* were calculated for the subset of 20 best individuals (maximum size of *pbest* subset) from the complex network created in each generation (the complex network is pruned after that and next generation creates a new one) and the results are depicted in the next section.

RESULTS

This section depicts representatives of typical average *centrRank* history among test functions in CEC2015 benchmark set. In the optimal scenario, the subset of 20 best individuals according to their objective function value would be the same as the subset of 20 best individuals according to their *centrRank* values and average *centrRank* value in one generation would be around 0.9. This behavior was not obtained from either of the test functions. The closest were functions 1, 2 and 15 where there is a fast convergence to the global (functions 1 and 2) or local optima (function 15). The average *centrRank* history behavior on function 1 is depicted in Figure 2 and convergence graph for this function is in Figure 3.

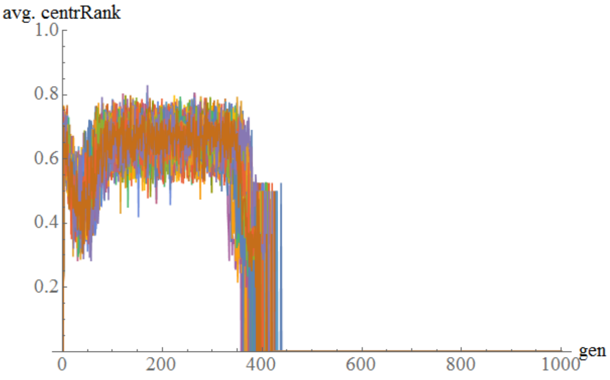


Figure 2: *CentrRank* History Graph of 51 Runs on f1 in 10*D*

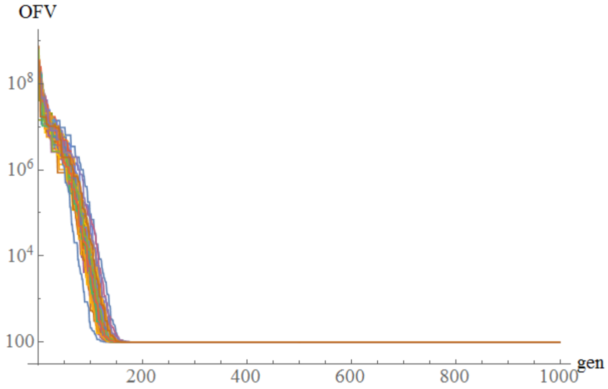


Figure 3: Convergence Graph of 51 Runs on f1 in 10*D*

The second obtained behavior is depicted in Figure 4, which represents average *centrRank* history of the function 4 with convergence graph in Figure 5. This behavior is common for functions 3, 4, 5, 7, 9, 12, 13. In this case, the average *centrRank* value is mostly lower and suggests that individuals with best objective function values are not the ones who lead the evolution.

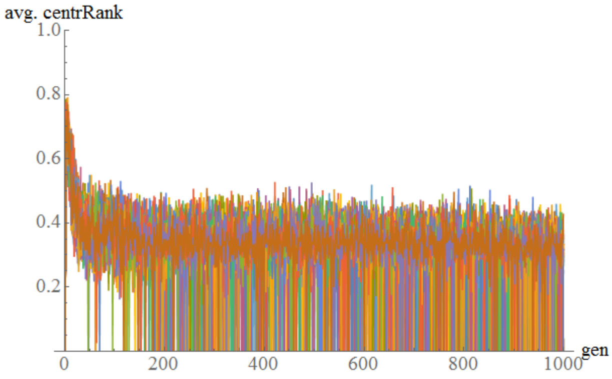


Figure 4: *CentrRank* History Graph of 51 Runs on f4 in 10*D*

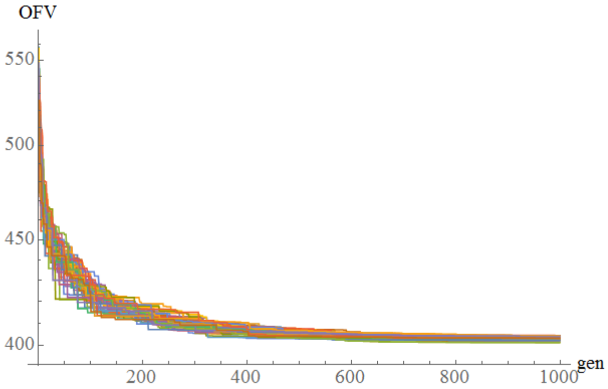


Figure 1: Convergence Graph of 51 Runs on f4 in 10*D*

The third behavior is the most interesting one. Figure 6 depicts the average *centrRank* history of function 10, where the average *centrRank* value fluctuates during the optimization process and suggests that the communication in network changes over time. This behavior is common for functions 6, 8, 10, 11, 14 and provides a lot of material for future research.

Convergence graph for function 10 is depicted in Figure 7.

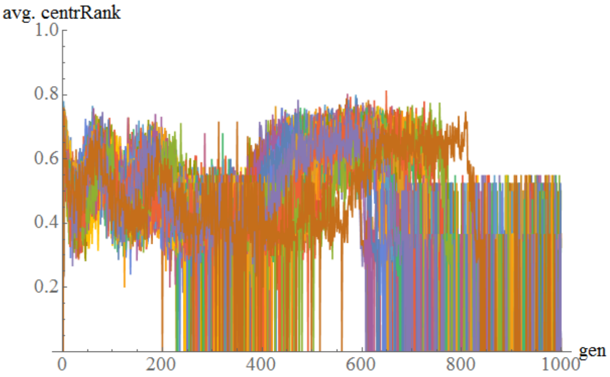


Figure 6: *CentrRank* History Graph of 51 Runs on f10 in 10*D*

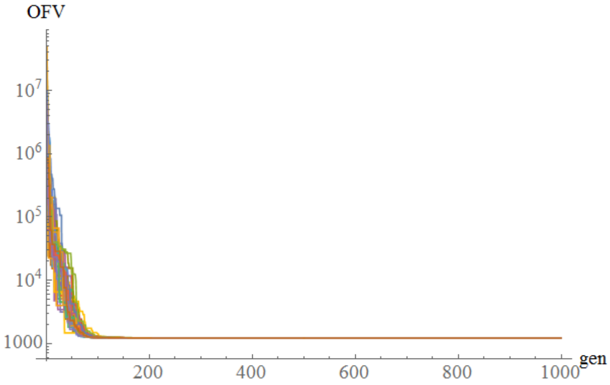


Figure 7: Convergence Graph of 51 Runs on f10 in 10*D*

Overall, behaviors 2 and 3 are not according to the basic assumption and suggest that there might be a possibility of adapting the mutation strategy to the given problem on the basis of information from the complex network.

All average *centrRank* history and convergence graphs along with numerical results can be found here: https://owncloud.cesnet.cz/index.php/s/d59pVbT5gqXbSrW

CONCLUSION

This work presented a novel approach for capturing SHADE optimization process in directed and weighted complex network, which should provide more accurate information about the heuristic dynamic. In the experimental part, the basic network feature – node centrality value was used for the analysis of greedy behavior of the "current-to-*p*best/1" mutation strategy. This analysis provided evidence that the best individuals in population are not the ones who communicate the most and therefore lead the evolution. This area of research should be exploited and therefore, the future research will be aimed at that direction. Complex network features will be examined and used for adaptation of the mutation strategy in order to improve the performance of given algorithm.

ACKNOWLEDGEMENT

This work was supported by Grant Agency of the Czech Republic – GACR P103/15/06700S, further by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme Project no. LO1303 (MSMT-7778/2014). Also by the European Regional Development Fund under the Project CEBIA-Tech no. CZ.1.05/2.1.00/03.0089 and by Internal Grant Agency of Tomas Bata University under the Projects no. IGA/CebiaTech/2017/004.

REFERENCES

Brest, J., Greiner, S., Boskovic, B., Mernik, M., & Zumer, V. (2006). Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark prob-lems. IEEE transactions on evolutionary computation, 10(6), 646-657.

Brest, J., Korošec, P., Šilc, J., Zamuda, A., Bošković, B., & Maučec, M. S. (2013). Differen-tial evolution and differential ant-stigmergy on dynamic optimisation problems. International Journal of Systems Science, 44(4), 663-679.

Das, S., Abraham, A., Chakraborty, U. K., & Konar, A. (2009). Differential evolution using a neighborhood-based mutation operator. IEEE Transactions on Evolutionary Computa-tion, 13(3), 526-553.

Das, S., Mullick, S. S., & Suganthan, P. N. (2016). Recent advances in differential evolution–An updated survey. Swarm and Evolutionary Computation, 27, 1-30.

Mallipeddi, R., Suganthan, P. N., Pan, Q. K., & Tasgetiren, M. F. (2011). Differential evolu-tion algorithm with ensemble of parameters and mutation strategies. Applied Soft Compu-ting, 11(2), 1679-1696.

Mininno, E., Neri, F., Cupertino, F., & Naso, D. (2011). Compact differential evolution. IEEE Transactions on Evolutionary Computation, 15(1), 32-54.

Neri, F., & Tirronen, V. (2010). Recent advances in differential evolution: a survey and exper-imental analysis. Artificial Intelligence Review, 33(1-2), 61-106.

Pluhacek, M., Janostik, J., Senkerik, R., & Zelinka, I. (2016). Converting PSO dynamics into complex network-Initial study. In T. Simos, & C. Tsitouras (Eds.), AIP Conference Proceedings (Vol. 1738, No. 1, p. 120021). AIP Publishing.

Qin, A. K., Huang, V. L., & Suganthan, P. N. (2009). Differential evolution algorithm with strategy adaptation for global numerical optimization. IEEE transactions on Evolutionary Computation, 13(2), 398-417.

Storn, R., & Price, K. (1995). Differential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces (Vol. 3). Berkeley: ICSI.

Tanabe, R., & Fukunaga, A. (2013). Success-history based parameter adaptation for differential evolution. In Evolutionary Computation (CEC), 2013 IEEE Congress on (pp. 71-78). IEEE.

Tanabe, R., & Fukunaga, A. S. (2014). Improving the search performance of SHADE using linear population size reduction. In Evolutionary Computation (CEC), 2014 IEEE Con-gress on (pp. 1658-1665). IEEE.

Viktorin, A., Pluhacek, M., & Senkerik, R. (2016). Network Based Linear Population Size Reduction in SHADE. In Intelligent Networking and Collaborative Systems (INCoS), 2016 International Conference on (pp. 86-93). IEEE.

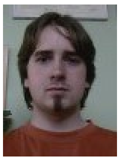
Zhang, J., & Sanderson, A. C. (2009). JADE: adaptive differential evolution with optional external archive. Evolutionary Computation, IEEE Transactions on, 13(5), 945-958.

AUTHOR BIOGRAPHIES

ADAM VIKTORIN was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Computer and Communication Systems and obtained his MSc degree in 2015. He is studying his Ph.D. at the same university and the fields of his studies are: Artificial intelligence, data mining and evolutionary algorithms. His email address is: aviktorin@fai.utb.cz

ROMAN SENKERIK was born in the Czech Republic, and went to the Tomas Bata University in Zlin, where he studied Technical Cybernetics and obtained his MSc degree in 2004, Ph.D. degree in Technical Cybernetics in 2008 and Assoc. prof. in 2013 (Informatics). He is now an Assoc. prof. at the same university (research and courses in: Evolutionary Computation, Applied Informatics, Cryptology, Artificial Intelligence, Mathematical Informatics). His email address is: senkerik@fai.utb.cz

MICHAL PLUHACEK was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Information Technologies and obtained his MSc degree in 2011 and Ph.D. in 2016 with the dissertation topic: Modern method of development and modifications of evolutionary computational techniques. He now works as a researcher at the same university. His email address is: pluhacek@fai.utb.cz



TOMAS KADAVY was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Information Technologies and obtained his MSc degree in 2016. He is studying his Ph.D. at the same university and the fields of his studies are: Artificial intelligence and evolutionary algorithms. His email address is: kadavy@fai.utb.cz

UNCOVERING COMMUNICATION DENSITY IN PSO USING COMPLEX NETWORK

Michal Pluhacek, Roman Senkerik, Adam Viktorin and Tomas Kadavy
Tomas Bata University in Zlin , Faculty of Applied Informatics
Nam T.G. Masaryka 5555, 760 01 Zlin, Czech Republic
{pluhacek, senkerik, aviktorin, kadavy}@fai.utb.cz

KEYWORDS

Swarm Intelligence, Particle Swarm Optimization, Complex Network, Swarm communication.

ABSTRACT

In this study, we investigate the communication in particle swarm optimization (PSO) by the means of network visualization. We measure the communication density of PSO optimizing four different benchmark functions. It is presented that the communication density varies over different fitness landscapes and in different phases of the optimizing process. We analyze the results in terms of use for future research.

INTRODUCTION

The Particle Swarm Optimization algorithm (PSO) (Kennedy, Eberhart 1995, Shi, Eberhart, 1998, Kennedy 1997, Nickabadi et al., 2011) is known as one of the leading metaheuristic optimizers. Heuristic methods are widely used for solving industrial problems (Volná, Kotyrba, 2016). In the past decades the inner dynamic of the PSO algorithm has been studied in detail and many modifications were proposed to tackle the known weaknesses of the method (e.g. premature convergence).

Recently the interconnection between metaheuristics and complex networks (CNs) has been (Zelinka 2011a, 2011b, 2013, Senkerik et al., 2016) with interesting results (Davendra, 2014a, 2014b).

We take inspiration in above mentioned examples of interconnection of metaheuristics and CNs and use the network-style visualization to uncover the density of communication in the PSO. A network structure is constructed from the inner communication of the swarm and afterwards analyzed.

The rest of the paper is structured as follows: The PSO is described in the next section, following is the description of network construction process. The experiment design is presented in the next section followed by the results discussion. The paper concludes with suggestion for future research.

PARTICLE SWARM OPTIMIZATION

The Particle Swarm Optimization algorithm (PSO) is an evolutionary optimization algorithm based on the natural behavior of birds. It was introduced by R. Eberhart and J. Kennedy in 1995 (Kennedy, Eberhart 1995).

In the PSO algorithm the particles (representing candidate solutions) fly in the multidimensional space of possible solutions. The new position of the particle in the next iteration is obtained as a sum of its actual position and velocity. The velocity calculation follows two natural tendencies of the particle: To move to the best solution found so far by the particular particle (personal best: *pBest*). And to move to the overall best solution found in the swarm (global best: *gBest*).

In the original PSO the new position of particle is altered by the velocity given by Eq. 1:

$$v_{ij}^{t+1} = w \cdot v_{ij}^{t} + c_1 \cdot Rand \cdot (pBest_{ij} - x_{ij}^{t}) + c_2 \cdot Rand \cdot (gBest_j - x_{ij}^{t}) \tag{1}$$

Where:
v_i^{t+1} - New velocity of the *i*th particle in iteration *t*+1.
w – Inertia weight value.
v_i^t - Current velocity of the *i*th particle in iteration *t*.
c_1, c_2 - Priority factors.
$pBest_i$ – Local (personal) best solution found by the *i*th particle.
gBest - Best solution found in a population.
x_{ij}^t - Current position of the *i*th particle (component *j* of the dimension *D*) in iteration *t*.
$Rand_{1j}$, $Rand_{2j}$ – Pseudo random numbers, interval (0, 1).

The maximum velocity of particles in the PSO is typically limited to 0.2 times the range of the optimization problem and this pattern was followed in this study. The new position of a particle is then given by Eq. 2, where x_i^{t+1} is the new particle position:

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{2}$$

Finally the linear decreasing inertia weight (Nickabadi et al., 2011) is used in this study. Its purpose is to slow the particles over time and improve the local search capability in the later phase of the optimization. The inertia weight has two control parameters w_{start} and w_{end}. A new *w* for each iteration is given by Eq. 3, where *t* stands for current iteration number and *n* stands for the total number of iterations.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

$$w = w_{start} - \frac{\left(\left(w_{start} - w_{end}\right) \cdot t\right)}{n} \qquad (3)$$

NETWORK CONSTRUCTION

In this study we use the network structure as a tool to help use represent the communication in the swarm. The nodes in the network represent the particles in different time points (Particle ID with iteration code). This means that the theoretical maximal number of nodes in the network is the number of particles times the number of iterations. However a new node in the network is created only when a particle manages to find a new personal best solution (*pBest*).

When a node is created, two links are also crated. First link is between the newly created node and previous node with the same particle ID (but different iteration code). This represents the information from *pBest* according to (1). Similarly the information from *gBest* represented by a link between the newly created node and a node that represents the last update of *gBest*.

THE EXPERIMENT

The following four well known test functions were used in this study: Sphere function, Rosenbrock function, Rastrigin function, Schwefel function with dimension setting 10 and 100.

In the experiment the PSO was set in the following way:

Iterations: 1000;

Population size: 20;

c_1,c_2: 2;

w_{start}: 0.9;

w_{end}: 0.4;

During the run of the algorithm the communication network was constructed according to the rules presented in the previous section.

Following is the visualization of the final networks. In the network visualizations a color coding is used to differentiate the phases of the run as percentage of the final number of cost functions evaluations (CFE). (The first 20% of CFE are represented by red color, magenta represents the 20-40% of CFE, green is the 40-60% CFE., 60-80% CFE is represented by yellow color and finally the 80-100% CFE is represented as cyan).

The network visualizations for Sphere function are presented in Fig. 1 (dim =10) and Fig. 2 (dim = 100) alongside the *gBest* development in Fig. 3 and Fig. 4.

Similarly the network visualizations and *gBest* history are presented in Fig. 5 – 8 for Rosenbrock function, Fig. 9 – 12 for Schwefel function in Fig 13 – 16 for Rastrigin functions.

It is clear from the visualizations that the number of newly created links in different phases of the algorithm varies. The numerical representation of newly created links is presented in Table 1 – 4.

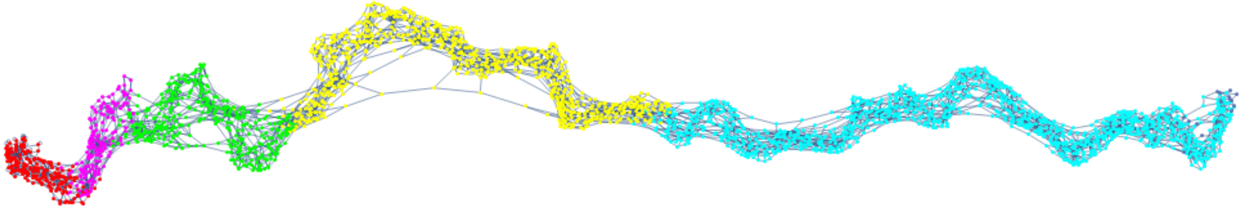


Figure 1: Network visualization with highlighted phases - Sphere function; dim = 10

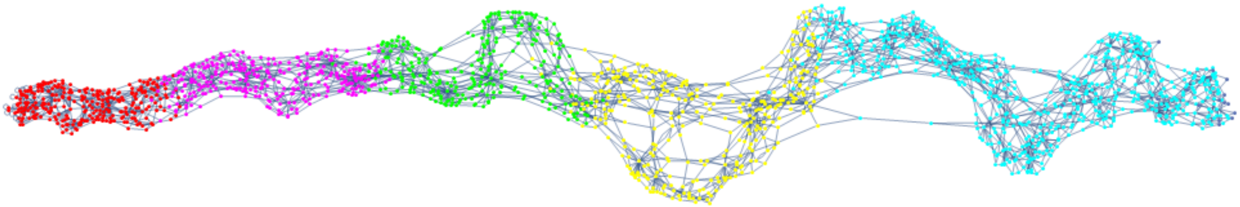


Figure 2: Network visualization with highlighted phases - Sphere function; dim = 100

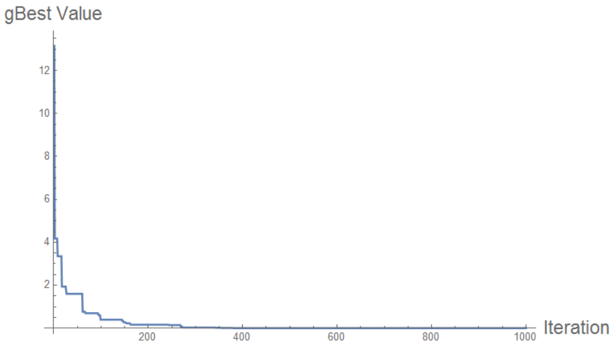


Figure 3: gBest history - Sphere function; dim = 10

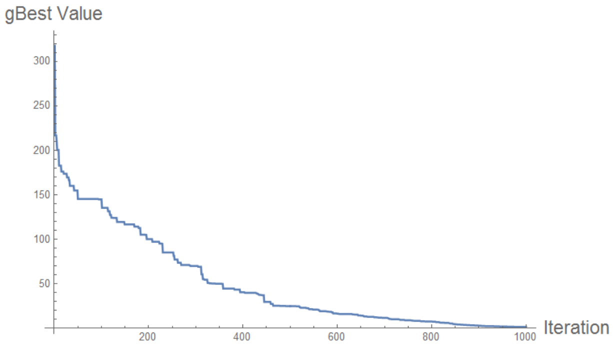


Figure 4: gBest history - Sphere function; dim = 100

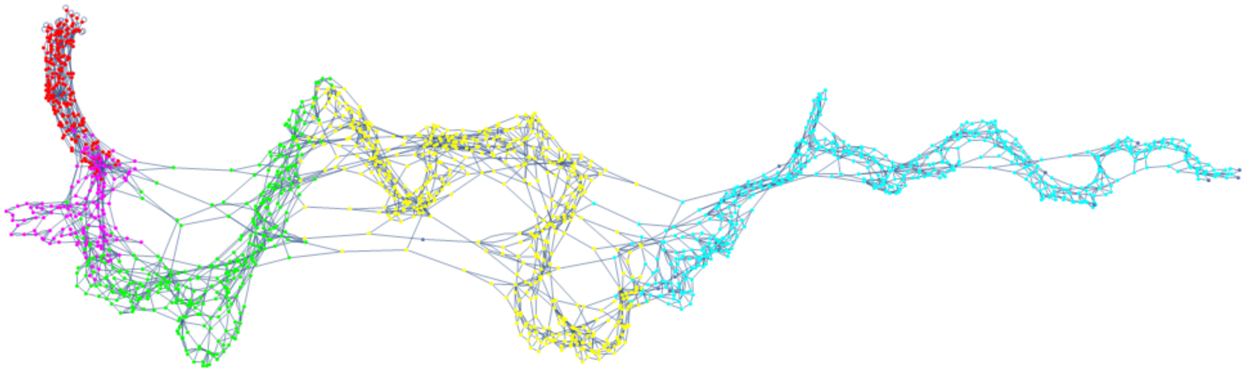


Figure 5: Network visualization with highlighted phases - Rosenbrock function; dim = 10

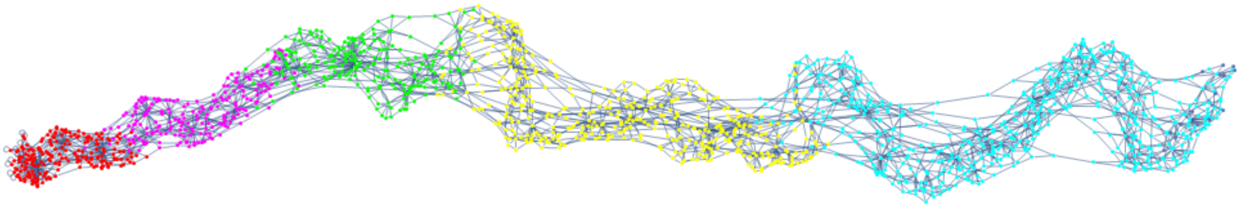


Figure 6: Network visualization with highlighted phases - Rosenbrock function; dim = 100

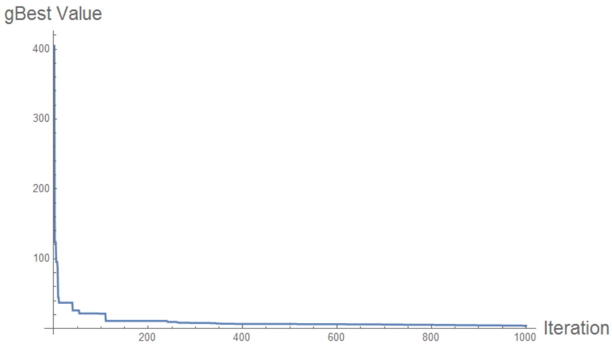


Figure 7: gBest history - Rosebrock function; dim = 10

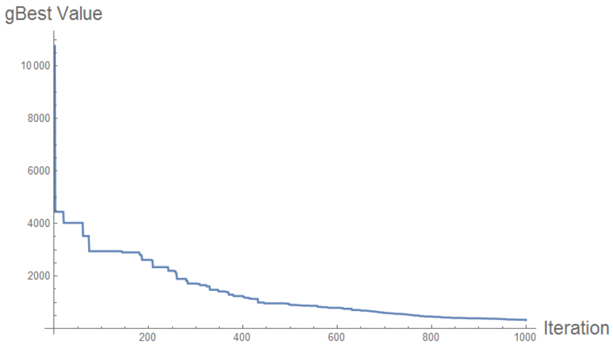


Figure 8: gBest history - Rosebrock function; dim = 100

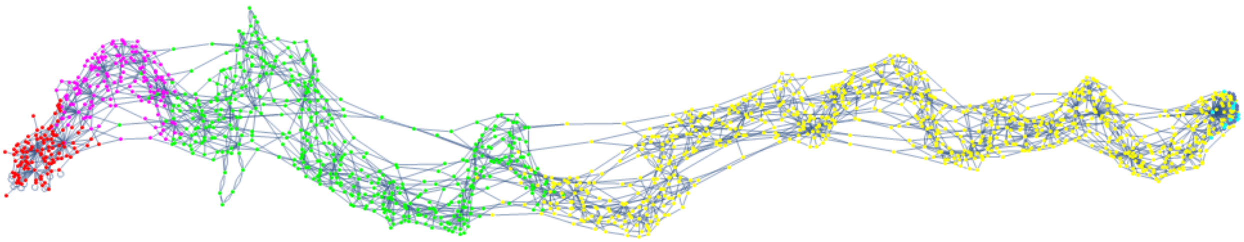


Figure 9: Network visualization with highlighted phases - Schwefel function; dim = 10

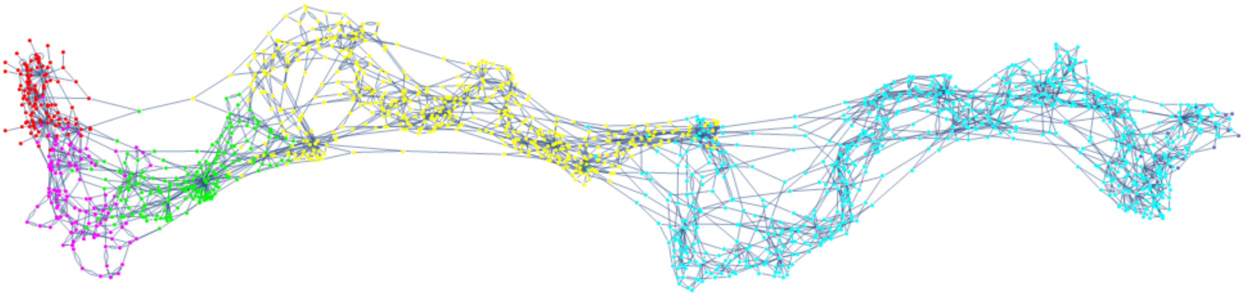


Figure 10: Network visualization with highlighted phases - Schwefel function; dim = 100

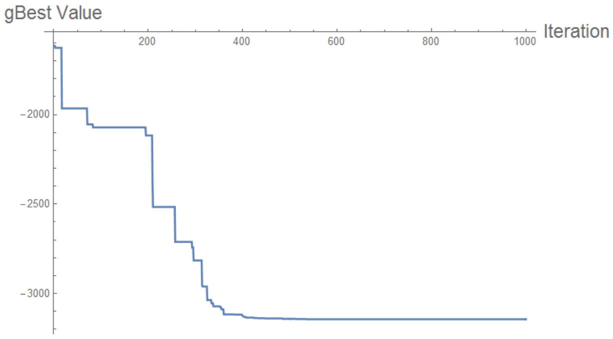


Figure 11: gBest history - Sschwefel function; dim = 10

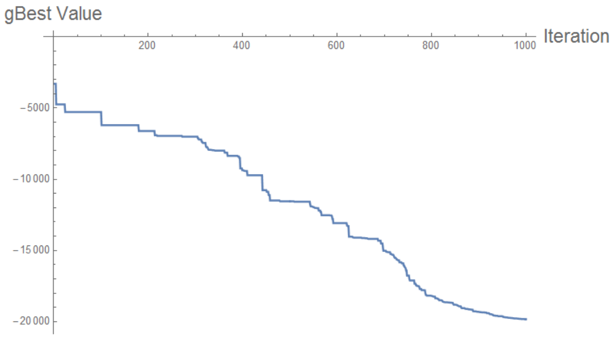


Figure 12: gBest history - Sschwefel function; dim = 100

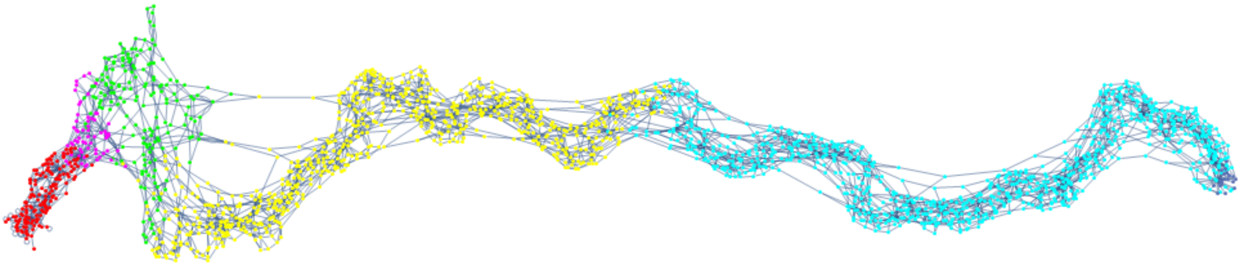


Figure 13: Network visualization with highlighted phases – Rastrigin function; dim = 10

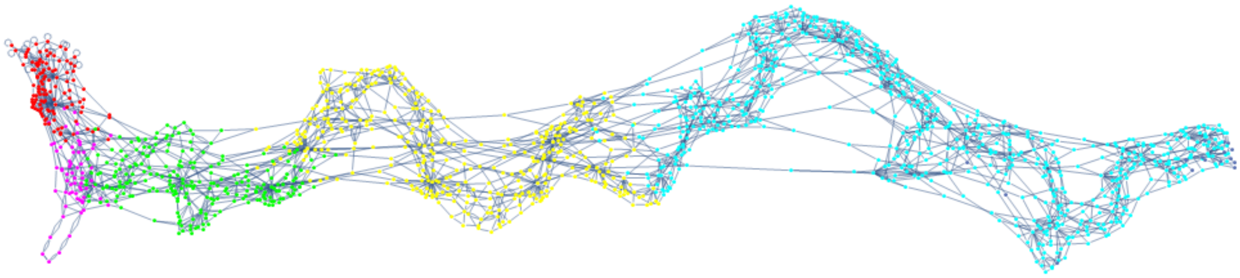


Figure 14: Network visualization with highlighted phases – Rastrigin function; dim = 100

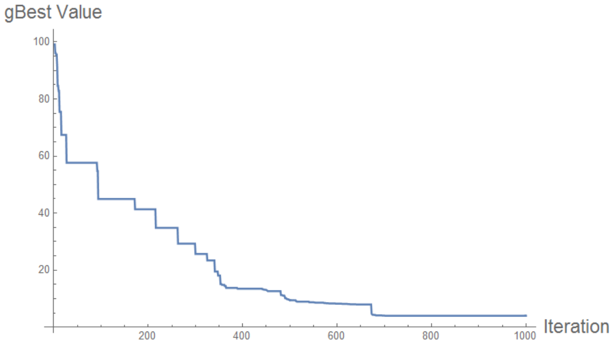


Figure 15: gBest history - Rastrigin function; dim = 10

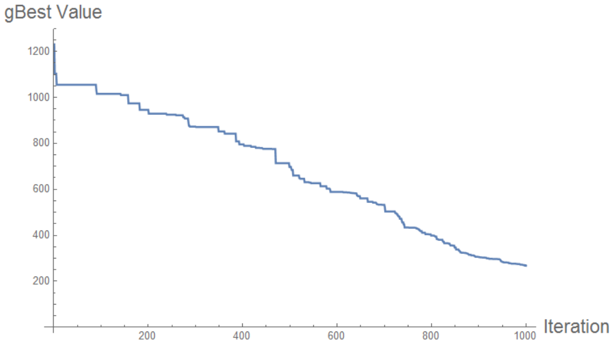


Figure 16: gBest history - Rastrigin function; dim = 100

Table 1: Newly created link overview – Sphere function

	Newly created links by CFE %				
dim	0 - 20	20 - 40	40 - 60	60 - 80	80 - 100
10	448	334	730	1520	2396
100	584	604	550	678	1280

Table 2: Newly created link overview – Rosenbrock function

	Newly created links by CFE %				
dim	0 - 20	20 - 40	40 - 60	60 - 80	80 - 100
10	424	298	684	1030	1030
100	440	404	452	816	1156

Table 3: Newly created link overview – Schwefel function

	Newly created links by CFE %				
dim	0 - 20	20 - 40	40 - 60	60 - 80	80 - 100
10	330	126	350	1408	1810
100	268	128	466	754	1378

Table 4: Newly created link overview – Rastrigin function

	Newly created links by CFE %				
dim	0 - 20	20 - 40	40 - 60	60 - 80	80 - 100
10	248	294	942	1666	22
100	230	222	374	938	1472

RESULTS DISSCUSION

Firstly, according to the results presented in the previous section it is clear that the number of newly created nodes in different phases of the algorithm varies. However when put into context with the history of *gBest*, often the smallest part of newly created nodes represents the most dramatic improvement of *gBest* value and vice versa.

Secondly, the shape of the network and number of newly created nodes seems to be affected by the fitness landscapes in terms of modality and complexity.

Further, there seems to be a tendency for some particles to improve after a very long time window without improvement (possibly escaping local optima), this trend can be observed namely in Figs. 5, 10 and 13.

In most cases the majority of the newly created nodes is crated in the last phases of the optimization. This is most likely due to the decreasing inertia weight and small velocities of the particles.

CONCLUSION

In this study we have presented the possible use of network visualization to highlight the trends in communication density in the particle swarm optimization. We have concluded that the communication density varies significantly in different phases of the optimization and also varies based on the fitness landscape.

There are two main directions for our future research. First is employment of more advanced network analysis for classification of various fitness landscapes and second is a feedback-loop style control of the swarm based on the number of newly created links in a specified time window.

ACKNOWLEDGEMENT

This work was supported by Grant Agency of the Czech Republic – GACR P103/15/06700S, further by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme Project no. LO1303 (MSMT-7778/2014. Also by the European Regional Development Fund under the Project CEBIA-Tech no. CZ.1.05/2.1.00/03.0089 and by Internal Grant Agency of Tomas Bata University under the Projects no. IGA/CebiaTech/2017/004

REFERENCES

Davendra, D., Zelinka, I, Metlicka, M., Senkerik, R., Pluhacek, M., "Complex network analysis of differential evolution algorithm applied to flowshop with no-wait problem," Differential Evolution (SDE), 2014 IEEE Symposium on , pp.1,8, 9-12 Dec. 2014

Davendra, D., Zelinka, I., Senkerik, R. and Pluhacek, M. Complex Network Analysis of Discrete Self-organising Migrating Algorithm, in: Zelinka, I. and Suganthan, P. and Chen, G. and Snasel, V. and Abraham, A. and Rossler, O. (Eds.) Nostradamus 2014: Prediction, Modeling and Analysis of Complex Systems, Advances in Intelligent Systems and Com-puting, Springer Berlin Heidelberg, pp. 161–174 (2014).

Kennedy J. and Eberhart R., “Particle swarm optimization,” in Proceedings of the IEEE International Conference on Neural Networks, 1995, pp. 1942–1948.

Kennedy J., “The particle swarm: social adaptation of knowledge,” in Proceedings of the IEEE International Conference on Evolutionary Computation, 1997, pp. 303–308.¨

Nickabadi A., Ebadzadeh M. M., Safabakhsh R., A novel particle swarm optimization algorithm with adaptive inertia weight, Applied Soft Computing, Volume 11, Issue 4, June 2011, Pages 3658-3670, ISSN 1568-4946

Senkerik, R., Viktorin, A., Pluhacek, M., Janostik, J. Oplatkova, Z. K. (2016). Study on the time development of complex network for metaheuristic. In Artificial Intelligence Perspectives in Intelligent Systems (pp. 525-533). Springer International Publishing.

Shi Y. and Eberhart R., “A modified particle swarm optimizer,” in Proceedings of the IEEE International Conference on Evolutionary Computation (IEEE World Congress on Computational Intelligence), 1998, pp. 69–73.I. S.

Volná, E. and Kotyrba, M. Unconventional heuristics for vehicle routing problems. Journal of Numerical Analysis, Industrial and Applied Mathematics. 2016, vol. 9-10, pp. 57-67. ISSN 1790-8140.

Zelinka, I. Investigation on relationship between complex network and evolutionary algo-rithms dynamics, AIP Conference Proceedings 1389 (1) 1011–1014 2011a.

Zelinka, I., Davendra, D., Enkek, R., Jaek, R.: Do Evolutionary Algorithm Dynamics Create Complex Network Structures? Complex Systems 2, 0891–2513, 20, 127–140, 2011b

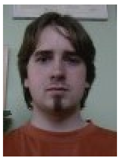
Zelinka, I., Davendra, D.D., Chadli, M., Senkerik, R., Dao, T.T., Skanderova, L.:Evolutionary Dynamics as The Structure of Complex Networks. In: Zelinka, I.,Snasel, V., Abraham, A. (eds.) Handbook of Optimization. ISRL, vol. 38, pp. 215–243. Springer, Heidelberg (2013)

AUTHOR BIOGRAPHIES

MICHAL PLUHACEK was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Information Technologies and obtained his MSc degree in 2011 and Ph.D. in 2016 with the dissertation topic: Modern method of development and modifications of evolutionary computational techniques. He now works as a researcher at the same university. His email address is: pluhacek@fai.utb.cz

ROMAN SENKERIK was born in the Czech Republic, and went to the Tomas Bata University in Zlin, where he studied Technical Cybernetics and obtained his MSc degree in 2004, Ph.D. degree in Technical Cybernetics in 2008 and Assoc. prof. in 2013 (Informatics). He is now an Assoc. prof. at the same university (research and courses in: Evolutionary Computation, Applied Informatics, Cryptology, Artificial Intelligence, Mathematical Informatics). His email address is: senkerik@fai.utb.cz

ADAM VIKTORIN was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Computer and Communication Systems and obtained his MSc degree in 2015. He is studying his Ph.D. at the same university and the field of his studies are: Artificial intelligence, data mining and evolutionary algorithms. His email address is: aviktorin@fai.utb.cz



TOMAS KADAVY was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Information Technologies and obtained his MSc degree in 2016. He is studying his Ph.D. at the same university and the fields of his studies are: Artificial intelligence and evolutionary algorithms. His email address is: kadavy@fai.utb.cz

FIREWORK ALGORITHM DYNAMICS SIMULATED AND ANALYZED WITH THE AID OF COMPLEX NETWORK

Tomas Kadavy
Michal Pluhacek
Adam Viktorin
Roman Senkerik
Tomas Bata University in Zlin, Faculty of Applied Informatics
Nam T.G. Masaryka 5555, 760 01 Zlin, Czech Republic
{kadavy, pluhacek, aviktorin, senkerik }@fai.utb.cz

KEYWORDS
Firework Algorithm, FWA, Complex Network, Surface Analysis

ABSTRACT

In this paper, we are presenting a visualization of Firework Algorithm (FWA) inner dynamics as an evolving complex network. Recent research in unconventional controlling and simulation of metaheuristic dynamics shows that this kind of visualization technique has been utilized only for algorithms with some social communication or behavior leading to sharing information across the population. Our simulation experiment presents the original approach for analyzing the complex dynamics of an algorithm based mostly on random/local search engines. Provided analysis suggests that the built network can be used for identification of test function surfaces types.

INTRODUCTION

The Firework Algorithm (FWA) is an algorithm for numerical optimization, which has been introduced in 2010 by authors Tan and Zhu (Tan, Zhu 2010). The algorithm is based on fireworks explosions in the sky. This algorithm has similar characteristics like some scatter search algorithms (Laguna, Marti 2003) or tabu search algorithm (Glover 1986). This FWA can be described as non-bio-inspired algorithm like a water drop algorithm (Shah-Hosseini 2009), brain storm optimization (Shi 2011) or magnetic optimization algorithms (Tayarani, Akbarzadeh-T 2008).

In this paper, the inner complex dynamic of FWA is transformed into the evolving complex network (Barrat et al. 2008). The population is visualised as an evolving complex network that usually exhibits non-trivial features – e.g. degree distribution, clustering, centralities and in between. These features offer a clear description of the population under evaluation and can be utilised for adaptive population as well as parameter control during the metaheuristic run. Such a complex networks can be more detailed analyzed using dedicated techniques (Otte et al. 2002; Kudelka et al. 2015). Recent research shows, that those analyses have been made for Swarm Intelligence (SI) algorithms (Pluhacek et al. 2016a; Pluhacek et al. 2016b; Senkerik et al. 2016a; Senkerik et al. 2016b), which exhibits some social behavior or communication across particles (individual solutions).

Our simulation experiment presents the original approach for analyzing the complex dynamics of an algorithm based not on social behavior, but mostly on random/local search engines. The research tasks can be summarized as follows:

- Could the similar behavior or communication be observed using the complex network approach also for FWA?
- Can the network features be used to observe and identify the surface of tested function?

The paper is structured as follows. The FWA is described in details within the next section. The complex network design and used test functions follow afterwards. Last two sections discuss the experiment setting and results.

FIREWORK ALGORITHM

The FWA is an algorithm that is inspired by fireworks explosion in a night sky. This algorithm is initialized with a random population of fireworks $\boldsymbol{X}$. The x_i firework position is represented as coordinates in n-dimensional space of solutions. These coordinates are parameters of the optimized problem. The number of the fireworks is defined by parameter *NP*; this parameter is set by the user. Moreover, the user defines the parameters like number of iterations of the algorithm (terminal condition), Gaussian mutation $\widehat{m}$, number of sparks m, parameters a and b and constant $\hat{A}$. This algorithm consists of four parts: explosion operator, mutation operator, mapping rule and selection strategy. These parts and adjustable parameters are more explained in next sections.

The realization of FWA is as follows:

1. Randomly generate *NP* fireworks in the n-dimensional search space.
2. Obtain fitness values of these generated fireworks by fitness function.
3. Calculate the number of generated sparks and their amplitude for each firework by explosion operator.
4. Use Gaussian mutation to generate new random sparks by mutation operator.
5. Apply mapping rule to all generated sparks.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

6. Calculate fitness values of sparks and by applying selection strategy pick the selected sparks as new fireworks.
7. If the terminal conditions are met, stop the algorithm. Otherwise, continue the iteration process from 3.

There can be more or different terminal conditions defined by the user. For example, a number of fitness evaluation (*FE*) instead of a number of iterations of the algorithm.

Initialization

The initial *NP*, the number of fireworks, fireworks ***X*** are randomly generated with uniform distribution from the range which is specified for the problem by lower and upper bounds defined by the optimized problem with dimensionality *dim*.
In the initialization phase, the adjustable parameters mentioned before has to be defined as well.

Explosion Operator

The number of sparks generated from each firework is determined by the firework fitness value. The firework with better fitness value produces more sparks (the lower cost function *f(x)*, the better fitness value). This number of sparks is calculated by explosion strength in (1).

$$S_i = m \cdot \frac{Y_{max} - f(x_i) + \varepsilon}{\sum_{i=1}^{NP}\left(Y_{max} - f(x_i)\right) + \varepsilon} \tag{1}$$

where S_i is the number of sparks for firework *i*, *m* is the total number of sparks defined by the user. Y_{max} means the fitness value of the worst individual (firework). Function $f(x_i)$ is the fitness value for the individual firework *i*. The last parameter ε is used to prevent the denominator from becoming zero and it should be the smallest possible number.
There is also a limitation of the number of generated spark defined as (2).

$$\widehat{s_i} = \begin{cases} round(a \cdot m), & if\ s_i < a \cdot m \\ round(b \cdot m), & if\ s_i > b \cdot m \\ round(s_i), & otherwise \end{cases} \tag{2}$$

where *a* and *b* are constants defined by the user (these constants has to be *a<b<1*), $\widehat{s_i}$ is the limitation of the number of sparks and *round()* is the rounding function.
The amplitude for generated sparks is then calculated by explosion amplitude in (3). Like the previous, explosion strength, the amplitude of explosion is defined by firework fitness function. The better fitness value is, the smaller is the amplitude of explosion and vice versa.

$$A_i = \hat{A} \cdot \frac{f(x_i) - Y_{min} + \varepsilon}{\sum_{i=1}^{NP}[f(x_i - Y_{min})] + \varepsilon} \tag{3}$$

where A_i is the amplitude of *i* firework. $\hat{A}$ is a constant defined by the user and means the sum of all amplitudes. Y_{min} means the fitness value of the best firework.
The new sparks are generated in randomly chosen dimensions ***z*** and the position is calculated in (4).

$$\hat{x}_j^k = x_i^k + U(-A_i, A_i) \tag{4}$$

where $\hat{x}_j^k$ is spark *j* in dimension *k* ($k \in \boldsymbol{z}$) generated from firework x_i. *U* is a random number from a uniform distribution in the range of the explosion amplitude of *i* firework.

Mutation Operator

To maintain the diversity of the population, some mutation operator is needed. For FWA, the Gaussian mutation is used. The sparks are generated as follows:

1. Choose random firework *i*.
2. Compute new spark using formula (5).
3. If the number of generated spark by Gaussian mutation reaches the value $\widehat{m}$, stop generating next sparks

$$\hat{x}_j^k = x_i^k \cdot N(1,1) \tag{5}$$

where $\hat{x}_j^k$ is spark *j* in dimension *k* ($k \in \boldsymbol{z}$) generated from firework x_i. Vector ***z*** are randomly chosen dimensions like in section Explosion Operator. *N* is a random number from normal (Gaussian) distribution with mean 1 and variance 1.

Mapping Rule

This rule ensures, that all previously generated sparks are in feasible space. If any spark lies outside of the available search space, its mapped back to allowed space. This mapping rule defined as (6).

$$\hat{x}_i^k = B_L^k + \hat{x}_i^k mod(B_U^k - B_L^k) \tag{6}$$

where $\hat{x}_i^k$ is *i* particle in *k* dimension, B_L^k and B_U^k are lower and upper boundaries of the available search space in *k* dimension. The *mod* represents modular operation.

Selection Strategy

Some of the generated sparks need to be selected and passed into the new iteration. These selected sparks will become new fireworks. For this selection, the distance-based strategy is used to maintain the diversity of the population. The spark that is farther from the others has the greater chance to be selected than those sparks near the other sparks. The first chosen spark is always the one with the best fitness value. Others (*NP*-1) individuals are chosen by roulette method. The possibility of choosing the spark into next iteration is calculated in (7).

$$p_i = \frac{R_i}{\sum_{j=1}^{K} R_j} \tag{7}$$

where p_i is the possibility of the *i* spark, R_i is the sum of distances of the *i* spark, *K* is the number of all generated sparks. The Euclidean distance is used to compute the R_i in formula (8).

$$R_i = \sum_{j=1}^{K} d(\hat{x}_i, \hat{x}_j) = \sum_{j=1}^{K} \left\| \hat{x}_i - \hat{x}_j \right\| \tag{8}$$

where K is the number of all sparks, $\hat{x}_i$is the spark for which the R_i is computed and $\hat{x}_j$ are others sparks where $j \in K$.
The whole FWA is depicted in the pseudo-code below.

Algorithm pseudo-code 1: FWA

```
Randomly initialize NP fireworks
while terminal condition not met
  count fireworks fitness values
  for $i$ = 1 to $NP$ do
    calculate $S_i$
    calculate $A_i$
    generate sparks of $i$ firework
  end
  for $j$ = 1 to $\hat{m}$ do
    Gaussian mutation
  end
  selection strategy for new
  fireworks
end
```

NETWORK DESIGN

The network is created as a history of contributions. In each iteration, there are *NP* fireworks. These fireworks create *K* sparks. Some of these sparks are transferred into a new iteration as new fireworks. Fireworks are then represented as the nodes in the network. These nodes are labelled 1…*NP* for each iteration. The nodes (fireworks) are sorted by their fitness values before labelling so that the best node (smallest fitness value) gets number 1 and the worst node gets number *NP*. The edge between nodes represents spark that creates a new firework in next iteration. The initial node of the edge represents the firework from which the spark is created. The terminal node is the firework in the next iteration created by the spark. With that rule, the initial node from *t* iteration can have from 0 to *NP* edges and terminal node can only have one edge as input.
An example of the network with five fireworks in four iterations is shown in Figure 1. Blue edges indicate the spark with the best fitness function value. The blue edge direction can only be towards the node number one. The first iteration is on left side of the figure, and the last iteration is on the right side. From the first iteration, four sparks create new fireworks in the second iteration and one of them contributes to improving the solution.

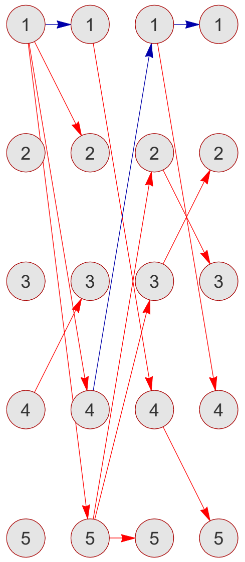


Figure 1: The example of FWA network.

TEST FUNCTIONS

For the simulation experiment, a set of 5 classic functions were selected. The set consists of unimodal and multimodal functions:

- Sphere function (f_1) (9),
- Rosenbrock function (f_2) (10),
- Rastrigin function (f_3) (11),
- Schwefel function (f_4) (12),
- Egg holder function (f_5) (13).

$$f(x)_1 = \sum_{i=1}^{dim} x_i^2 \tag{9}$$

$$f(x)_2 = \sum_{i=1}^{dim-1} \left[100 \cdot \left(x_{i+1} - x_i^2\right)^2 + (1 - x_i)^2\right] \tag{10}$$

$$f(x)_3 = 10 \cdot dim + \sum_{i=1}^{dim} \left[x_i^2 - 10 \cdot cos(2\pi x_i)\right] \tag{11}$$

$$f(x)_4 = \sum_{i=1}^{dim} \left[-x_i \cdot sin(|x_i|^{0.5})\right] \tag{12}$$

$$f(x)_5 = \sum_{i=1}^{dim-1} \left[-(x_{i+1} + 47) \cdot sin\left(\sqrt{\left|x_{i+1} + \frac{x_i}{2} + 47\right|}\right) - x_i \cdot sin\left(\sqrt{|x_i - (x_{i+1} + 47)|}\right)\right] \tag{13}$$

EXPERIMENT SETTING

The experiments were performed for test functions dimensions 2, 10 and 30. The number of iterations was set to 45. The control FVA parameters were set accordingly to (Tan, Zhu 2010). The number of fireworks, population size (*NP*), was set to 5 for all dimensions. The number of sparks (*m*) was set as 50. Parameters *a* and *b* were set as 0.8 and 0.04. Other constant settings were following: $\hat{A}$ = 40 and $\hat{m}$ = 5.
The basic logical assumption was that the *longest path of steady improvement* in the network (i.e. the path between

nodes labeled 1 and joined with blue edges) would be observable mostly for the unimodal function (e.g. f_1).

RESULTS

The results for the aforementioned longest paths of the stable improvement are given in Table 1.

Table 1: Longest paths of steady improvement in the network.

Function	Dimension		
	2	*10*	*30*
f_1	10	8	15
f_2	20	10	8
f_3	11	6	7
f_4	9	8	7
f_5	5	5	15

Results depicted in Table 1 confirm the anticipated logical assumption made in the previous section. For the unimodal functions (f_1 and f_2), the observed path is quite longer compared to the results for the multimodal functions. The differences are decreasing with the higher dimension setting. These trends are graphically confirmed also in Figures 2 - 7. Nonetheless, the data shown in Figure 8 as well as in Table 1, indicate an exception to the primary logical assumption. Detailed analysis reveals that the path of steady improvement is present at the beginning of the captured evolved network (optimization process) and this may be caused by premature stagnation in the local optimum of tested function. For unimodal functions, the path of steady improvement seems to be present more often at the end of the recorded optimization process.

CONCLUSION

In this paper, the possibility of simulation and simple analysis of complex network evolvement for firework algorithm inner dynamics is present. Our novel approach was tested on the set of five simple classical benchmark functions.

The preliminary results lend weight to the argument that the ability of a network to identify the surface type of optimised function seems to be present. Nevertheless, more and detailed in-depth study is required to be performed in this field.

Another phenomenon has been discovered. The network seems to have a lack of any other usable information. The results of this simple simulation study will be further used in future research to suggest possible improvements to building complex networks for the family of algorithms based on the local/random search techniques without accessible direct social/communication interactions.

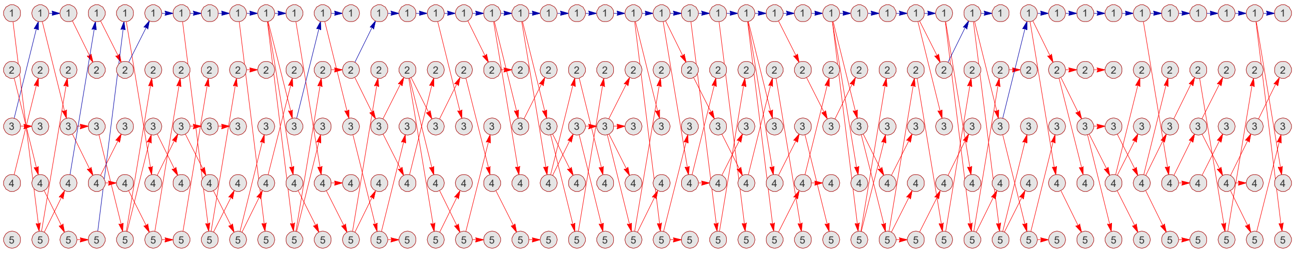


Figure 2: Network of f_2 for *dim* 2.

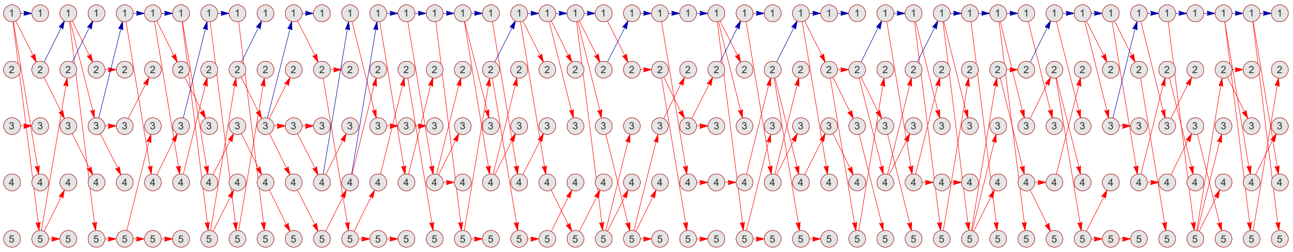


Figure 3: Network of f_5 for *dim* 2.

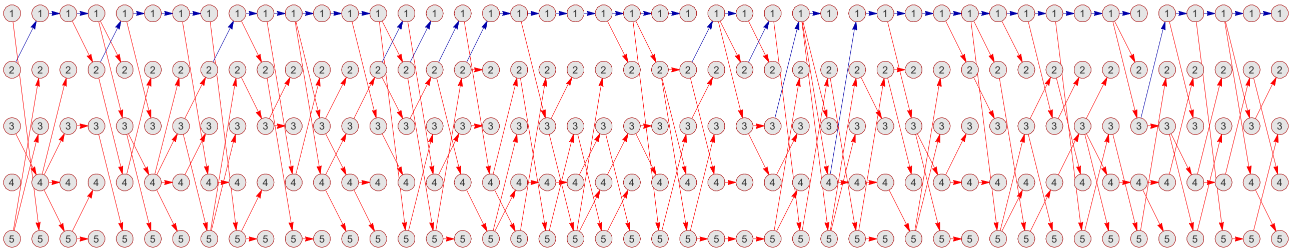


Figure 4: Network of f_2 for *dim* 10.

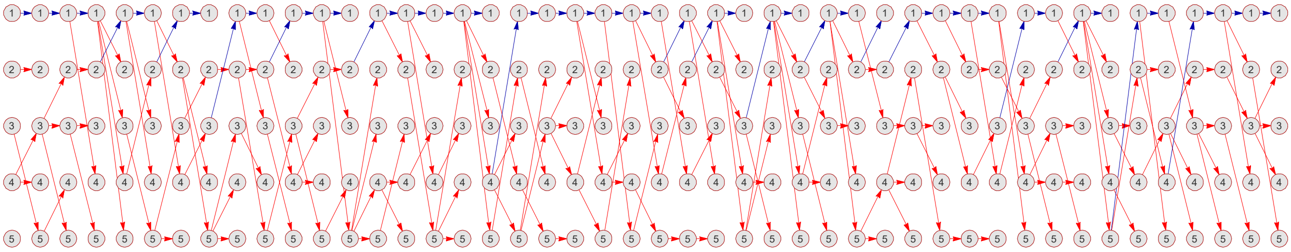


Figure 5: Network of *f_5* for *dim* 10.

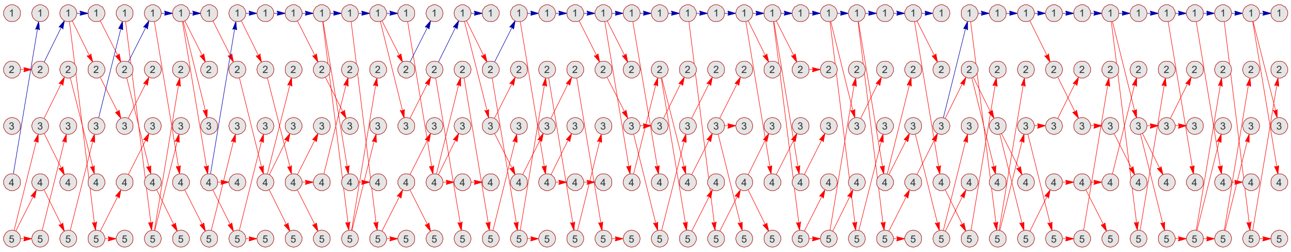


Figure 6: Network of *f_1* for *dim* 30.

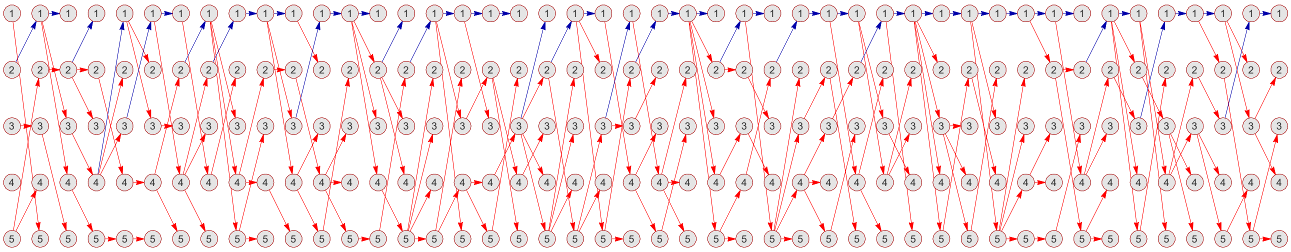


Figure 7: Network of *f_4* for *dim* 30.

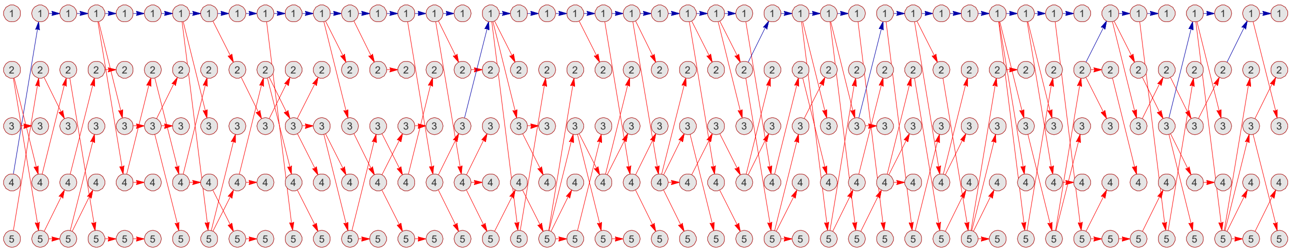


Figure 8: Network of *f_5* for *dim* 30.

ACKNOWLEDGEMENT

This work was supported by Grant Agency of the Czech Republic – GACR P103/15/06700S, further by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme Project no. LO1303 (MSMT-7778/2014). Also by the European Regional Development Fund under the Project CEBIA-Tech no. CZ.1.05/2.1.00/03.0089 and by Internal Grant Agency of Tomas Bata University under the Projects no. IGA/CebiaTech/2017/004.

REFERENCES

Tan Y., Zhu Y. (2010) Fireworks Algorithm for Optimization. In: Tan Y., Shi Y., Tan K.C. (eds) Advances in Swarm Intelligence. ICSI 2010. Lecture Notes in Computer Science, vol 6145. Springer, Berlin, Heidelberg

Laguna M., Marti R. Scatter search: methodology and implementations in C. Boston: Kluwer Academic Publishers, c2003. ISBN 9781402073762.

Glover F., Future paths for integer programming and links to artificial intelligence, Computers & Operations Research, Volume 13, Issue 5, 1986, Pages 533-549

Shah-Hosseini H., the intelligent water drops algorithm: a nature-inspired swarm-based optimization algorithm. Int. J. Bio-Inspir. Comput. 1(1), 71-79 (2009)

Shi Y., Brain storm optimization algorithm, in Advances in Swarm intelligence (Springer, Berlin, 2011), pp. 303-309

Tayarani N.M.H., Akbarzadeh-T M.R., Magnetic optimization algorithms a new synthesis, in 2008 IEEE World Congress on Compunational Intelligence Evolutionary Computation (CEC) (IEEE, 2008), pp. 2659-2664

Barrat, A., Barthelemy M., Vespignani A. Dynamical processes on complex networks. New York: Cambridge University Press, 2008. ISBN 9780521879507.

Otte, Evelien; Rousseau, Ronald (2002). "Social network analysis: a powerful strategy, also for the information sciences". Journal of Information Science. 28 (6): 441–453

Kudělka, M., Zehnalová, Š., Horák, Z., Krömer, P., & Snášel, V. (2015). Local dependency in networks. International Journal of Applied Mathematics and Computer Science, 25(2), 281-293.

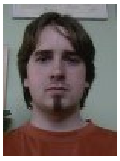
Pluhacek, M., Janostik, J., Senkerik, R., & Zelinka, I. (2016a). Converting PSO dynamics into complex network-Initial study. In T. Simos, & C. Tsitouras (Eds.), AIP Conference Proceedings (Vol. 1738, No. 1, p. 120021). AIP Publishing.

Pluhacek, M., Senkerik, R., Janostik, J., Viktorin, A., & Zelinka, I. (2016b). Study on swarm dynamics converted into complex network. In Proceedings-30th European Conference on Modelling and Simulation, ECMS 2016. European Council for Modelling and Simulation (ECMS).

Senkerik, R., Viktorin, A., Pluhacek, M., Janostik, J., & Davendra, D. (2016a). On the Influence of Different Randomization and Complex Network Analysis for Differential Evolution. In 2016 IEEE Congress on Evolutionary Computation (CEC) (pp. 3346-3353). IEEE.

Senkerik, R., Viktorin, A., Pluhacek, M., Janostik, J., & Oplatkova, Z. K. (2016b). Study on the Time Development of Complex Network for Metaheuristic. In Artificial Intelligence Perspectives in Intelligent Systems (pp. 525-533). Springer International Publishing.

AUTHOR BIOGRAPHIES



TOMAS KADAVY was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Information Technologies and obtained his MSc degree in 2016. He is studying his Ph.D. at the same university and the fields of his studies are: Artificial intelligence and evolutionary algorithms. His email address is: kadavy@fai.utb.cz

MICHAL PLUHACEK was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Information Technologies and obtained his MSc degree in 2011 and Ph.D. in 2016 with the dissertation topic: Modern method of development and modifications of evolutionary computational techniques. He now works as a researcher at the same university. His email address is: pluhacek@fai.utb.cz

ADAM VIKTORIN was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Computer and Communication Systems and obtained his MSc degree in 2015. He is studying his Ph.D. at the same university and the fields of his studies are: Artificial intelligence, data mining and evolutionary algorithms. His email address is: aviktorin@fai.utb.cz

ROMAN SENKERIK was born in the Czech Republic, and went to the Tomas Bata University in Zlin, where he studied Technical Cybernetics and obtained his MSc degree in 2004, Ph.D. degree in Technical Cybernetics in 2008 and Assoc. prof. in 2013 (Informatics). He is now an Assoc. prof. at the same university (research and courses in: Evolutionary Computation, Applied Informatics, Cryptology, Artificial Intelligence, Mathematical Informatics). His email address is: senkerik@fai.utb.cz

SIMULATION OF CHAOTIC DYNAMICS FOR CHAOS BASED OPTIMIZATION – AN EXTENDED STUDY

Roman Senkerik, Michal Pluhacek, Adam Viktorin, Zuzana Kominkova Oplatkova and Tomas Kadavy

Tomas Bata University in Zlin , Faculty of Applied Informatics
Nam T.G. Masaryka 5555, 760 01 Zlin, Czech Republic
{senkerik, oplatkova , pluhacek , aviktorin , kadavy}@fai.utb.cz

KEYWORDS

Deterministic chaos; Chaotic oscillators; Heuristic; Chaotic Optimization; Chaotic Pseudo Random Number Generators

ABSTRACT

This paper discuss the utilization of the complex chaotic dynamics given by the selected time-continuous chaotic systems as well as by the discrete chaotic maps, as the chaotic pseudo random number generators and driving maps for the chaos based optimization. Such an optimization concept is utilizing direct output iterations of chaotic system transferred into the required numerical range or it uses the chaotic dynamics for mapping the search space mostly within the local search techniques. This paper shows totally three groups of complex chaotic dynamics given by chaotic flows, oscillators and discrete maps. Simulations of examples of chaotic dynamics mapped to the search space were performed and related issues like parametric plots, distributions of such a systems, periodicity, and dependency on internal accessible parameters are briefly discussed in this paper.

INTRODUCTION

Generally speaking, the term “chaos” can denote anything that cannot be predicted deterministically. When the term “chaos” is combined with an attribute such as “deterministic”, a specific type of chaotic paradigm is involved, having their specific laws, mathematical apparatus, and a physical origin.

Recently, the chaos has been observed in many of various systems (including evolutionary one). Systems exhibiting deterministic chaos include, for instance, weather, biological systems, many electronic circuits (Chua’s circuit), mechanical systems, such as double pendulum, magnetic pendulum, or so called billiard problem.

The deterministic chaos is not based on the presence of randomness or any stochastic effects. It is clear from the mathematic definition and structure of the equations (see the section “*Chaotic Optimization”*), that no mathematical term expressing randomness is present. The seeming randomness in deterministic chaos is related to the extreme sensitivity to the initial conditions (Celikovsky and Zelinka 2010).

The idea of using chaotic dynamics as a replacement of classical pseudo-number generators – (PRNGs) has been presented in several research papers and in many applications with promising results (Lee and Chang 1996; Wu and Wang, 1999).

Another research joining deterministic chaos and pseudorandom number generator has been done for example in (Lozi 2012). Possibility of generation of pure random or pseudorandom numbers by use of the ultra weak multidimensional coupling of 1-dimensional dynamical systems is discussed there.

Another paper (Persohn and Povinelli 2012), deeply investigate well-known logistic map as a possible pseudorandom number generator and is compared with contemporary pseudo-random number generators. A comparison of logistic map results is made with conventional methods of generating pseudorandom numbers. The approach used to determine the number, delay, and period of the orbits of the logistic map at varying degrees of precision (3 to 23 bits). Another paper (Wang and Qin 2012) proposed an algorithm of generating pseudorandom number generator, which is called (couple map lattice based on discrete chaotic iteration) and combine the couple map lattice and chaotic iteration. Authors also tested this algorithm in NIST 800-22 statistical test suits and for future utilization in image encryption. In (Narendra et al. 2010) authors exploit interesting properties of chaotic systems to design a random bit generator, called CCCBG, in which two chaotic systems are cross-coupled with each other. A new binary stream-cipher algorithm based on dual one-dimensional chaotic maps is proposed in (Yang and Wang 2012) with statistic proprieties showing that the sequence is of high randomness. Similar studies are also done in (Bucolo et al. 2002).

MOTIVATION

This paper represents the extension of preliminary suggestions described in (Senkerik et al. 2016). Recently the deterministic chaos has been frequently used either as a replacement of (mostly uniform distribution based) pseudo-number generators (PRGNs) in metaheuristic algorithms or for simple mapping of solutions/iterations within intelligent local search engines. The metaheuristic chaotic approach generally uses the chaotic system in the place of a pseudo random number generator (Aydin et al. 2010). This causes the

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

heuristic to map search regions based on unique sequencing and periodicity of transferred chaotic dynamics, thus simulating the dynamical alternations of several subpopulations. The task is then to select a very good chaotic system (either discrete or time-continuous) as the pseudo random number generator (Caponetto et al. 2003).

Recently, the concept of embedding of chaotic dynamics into the evolutionary algorithms has been studied intensively. The self-adaptive chaos differential evolution (SACDE) (Zhenyu et al. 2006) was followed by the implementation of chaos into the simple not-adaptive differential evolution (Davendra et al. 2010), (Senkerik et al. 2014); chaotic searching algorithm for the very same metaheuristic was introduced in (Liang et al. 2011). Also the PSO (Particle Swarm Optimization) algorithm with elements of chaos was introduced as CPSO (Coelho and Mariani 2009). Many other works focusing on the hybridization of the swarm and chaotic movement have been published afterwards (Pluhacek et al. 2013), (Pluhacek et al. 2014). Later on, the utilization of chaotic sequences became to be popular in many interdisciplinary applications and techniques. The question of impact and importance of different randomization within heuristic search was intensively studied in (Zamuda and Brest 2015)

The primary aim of this work is to try, analyze and compare the implementation of different natural chaotic dynamic as the mapping procedure for the optimization/searching process. This paper presents the discussion about the usability of such systems, periodicity, and dependency on internal accessible parameters; thus the usability for local search or metaheuristic based optimization techniques.

CHAOTIC OPTIMIZATION

Generally, there exist three possible utilizations of chaotic dynamics in optimization tasks.

Firstly, as aforementioned in previous section, the direct output simulation iterations of chaotic system are transferred into the required numerical range (as simple CPRNG). The idea of CPRNG is to replace the default system PRNG with the chaotic system. As the chaotic system is a set of equations with a static start position (See the next section), we created a random start position of the system, in order to have different start position for different experiments. Once the start position of the chaotic system has been obtained, the system generates the next sequence using its current position. Subsequently, simple techniques as to how to deal with the negative numbers as well as with the scaling of the wide range of the numbers given by the chaotic systems into the typical range 0 – 1

Secondly, the complexity of chaotic systems and its movement in the space is used for dynamical mapping of the search space mostly within the local search techniques (Hamaizia and Lozi 2011).

Finally, the hybridization of searching/optimization process and chaotic systems is represented by chaos based random walk technique.

CHAOTIC SYSTEMS

This section contains the description of three different groups of chaotic dynamics: time-continuous chaotic systems (flows and oscillators), and the discrete chaotic maps.

Time-continuous Chaotic Systems

In this research, following chaotic systems were used: Lorenz system (1) and Rossler system (2) as two examples from chaotic flows; further unmodified UEDA oscillator (3); and Driven Van der Pol Oscillator (4) as chaotic oscillators (Sprott 2003).

The Lorenz system (1) is a 3-dimensional dynamical flow, which exhibits chaotic behavior. It was introduced by Edward Lorenz in 1963, who derived it from the simplified equations of convection rolls arising in the equations of the atmosphere.

The Rossler system (2) exhibits chaotic dynamics associated with the fractal properties of the attractor. It was originally introduced as an example of very simple chaotic flow containing chaos similarly to the Lorenz attractor. This attractor has some similarities to the Lorenz attractor, but is simpler

UEDA oscillator (3) is the simple example of driven pendulums, which represent some of the most significant examples of chaos and regularity.

UEDA system can be simply considered as a special case of intensively studied Duffing oscillator that has both a linear and cubic restoring force. Ueda oscillator represents the both biologically and physically important dynamical model exhibiting chaotic motion.

Finally, The Van der Pol oscillator (4) is the simple example of the limit cycles and chaotic behavior in electrical circuits employing vacuum tubes. Similarly to the UEDA oscillator, it can be used to explore physical (unstable) behaviour in biological sciences. (Bharti and Yuasa 2010).

The equations, which describe the chaotic systems, have parameter settings for Lorenz system: $a = 3.0$, $b = 26.5$; Rössler system: $a = 0.2$, $b = 0.2$, and $c = 5.7$; UEDA oscillator: $a = 1.0$ $b = 0.05$, $c = 7.5$ and $\omega = 1.0$; and Van der Pol oscillator : $\mu = 0.2$ $\gamma = 8.0$, $a = 0.35$ and $\omega = 1.02$.

$$\begin{aligned}\frac{dx}{dt} &= -a(x-y)\\ \frac{dy}{dt} &= x(b-z)-y\\ \frac{dz}{dt} &= xy-z\end{aligned} \tag{1}$$

$$\begin{aligned}\frac{dx}{dt} &= -y-z\\ \frac{dy}{dt} &= x+ay\\ \frac{dz}{dt} &= b+z(x-c)\end{aligned} \tag{2}$$

$$\begin{aligned}\frac{dx}{dt} &= y \\ \frac{dy}{dt} &= -ax^3 - by + c\sin\omega t\end{aligned} \quad . \tag{3}$$

$$\begin{aligned}\frac{dx}{dt} &= y \\ \frac{dy}{dt} &= \mu\left(1-\gamma x^2\right)y - x^3 + a\sin\omega t\end{aligned} \tag{4}$$

The parametric plots of the chaotic systems are depicted in Figure 1. The Figure 2 show the example of dynamical sequencing during the generating of pseudo number numbers transferred into the range <0 - 1> by means of particular studied CPRNGs and with the sampling rate of 0.5s. The dependency of sequencing and periodicity on the sampling rate is discussed in details in (Senkerik et al. 2015)

Discrete Chaotic Maps

The examples of chaotic maps are following: Arnold Cat map (5), Dissipative Standard map (6), Ikeda map (7), and Lozi map (8). Map equations and parameters values are given in Table 1. Parametric plots are depicted in Figure 3, whereas the examples of dynamical sequencing are given in Figure 4.

Table 1: Used discrete chaotic systems as CPRNG and parameters set up.

Chaotic system	Notation	Parameters values
Arnold Cat Map	$X_{n+1} = X_n + Y_n \pmod 1$ $Y_{n+1} = X_n + kY_n \pmod 1$	$k = 2.0$
Dissipative Standard Map	$X_{n+1} = X_n + Y_{n+1} (\bmod 2\pi)$ $Y_{n+1} = bY_n + k\sin X_n (\bmod 2\pi)$	$b = 0.6$ and $k = 8.8$
Ikeda Map	$X_{n+1} = \gamma + \mu(X_n\cos\phi + Y_n\sin\phi)$ $Y_{n+1} = \mu(X_n\sin\phi + Y_n\cos\phi)$ $\phi = \beta - \alpha/\left(1 + X_n^2 + Y_n^2\right)$	α = 6, β = 0.4, γ = 1 and μ = 0.9
Lozi Map	$X_{n+1} = 1 - a\lvert X_n\rvert + bY_n$ $Y_{n+1} = X_n$	$a = 1.7$ and $b = 0.5$

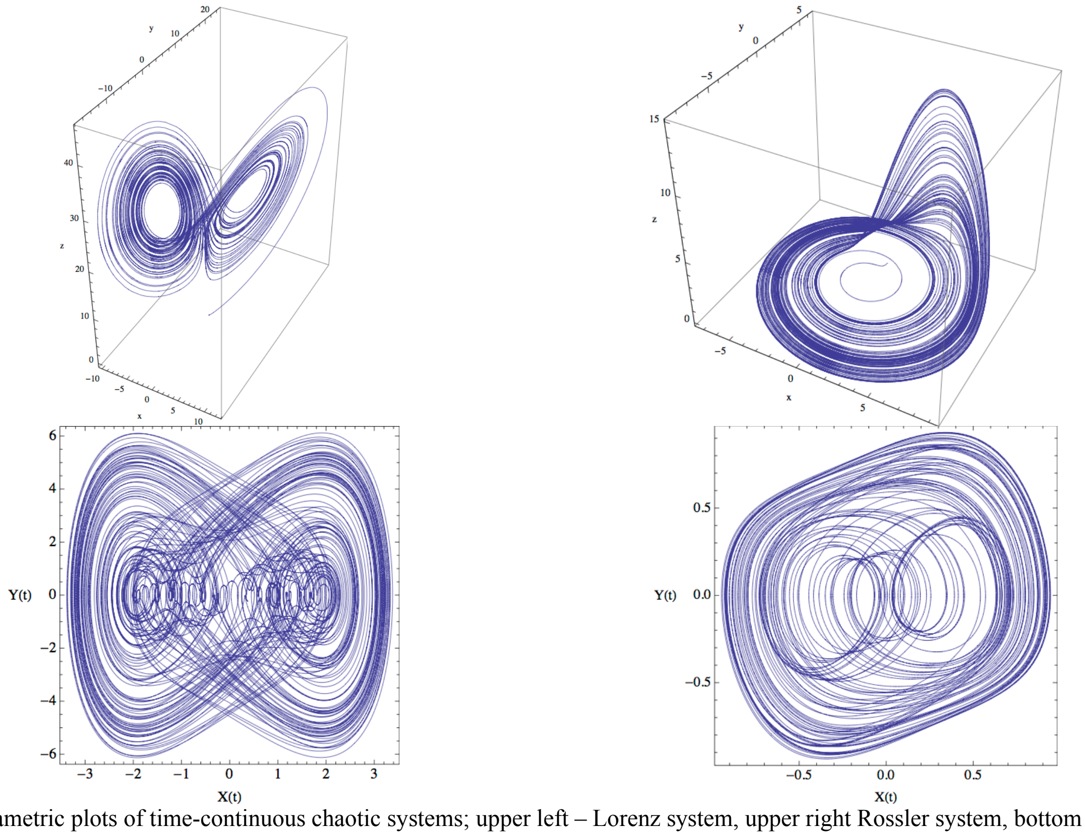


Fig. 1: Parametric plots of time-continuous chaotic systems; upper left – Lorenz system, upper right Rossler system, bottom left UEDA oscillator, bottom right Van der Pol Oscillator.

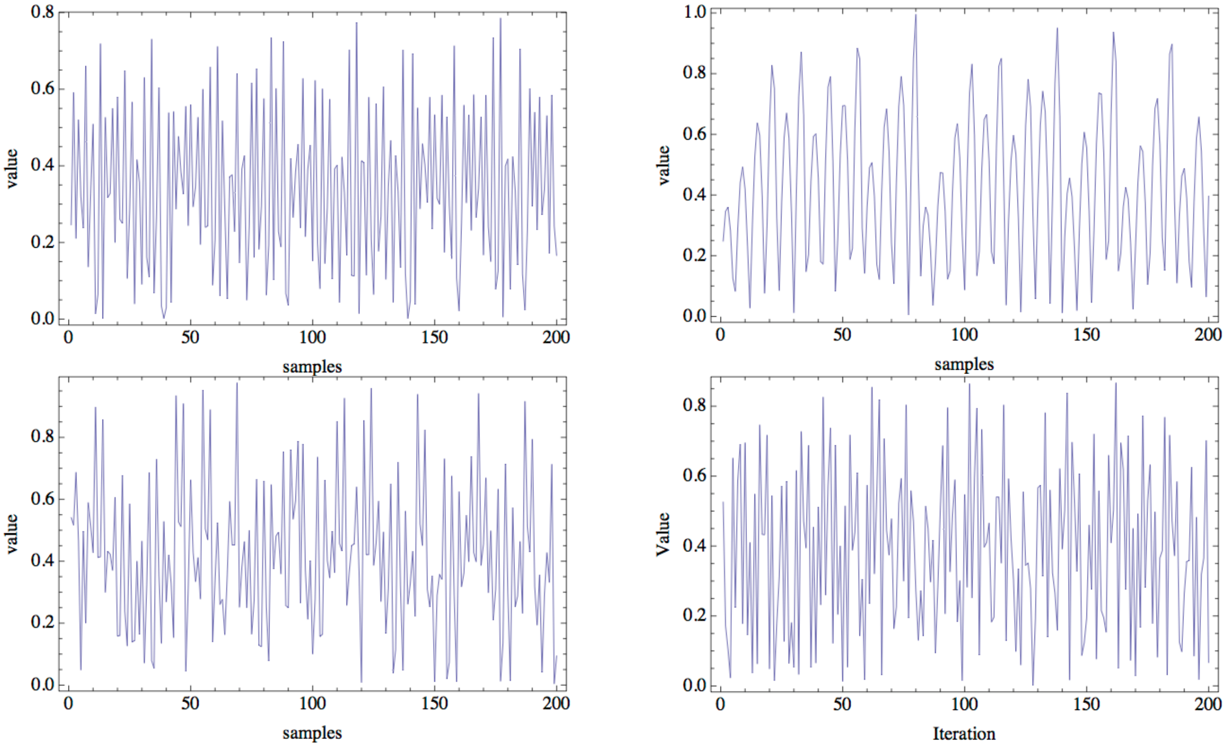


Fig. 2: Comparison of dynamical sequencing of pseudo random real numbers transferred into the range <0 - 1>; upper left – Lorenz system, upper right Rossler system, bottom left UEDA oscillator, bottom right Van der Pol Oscillator.

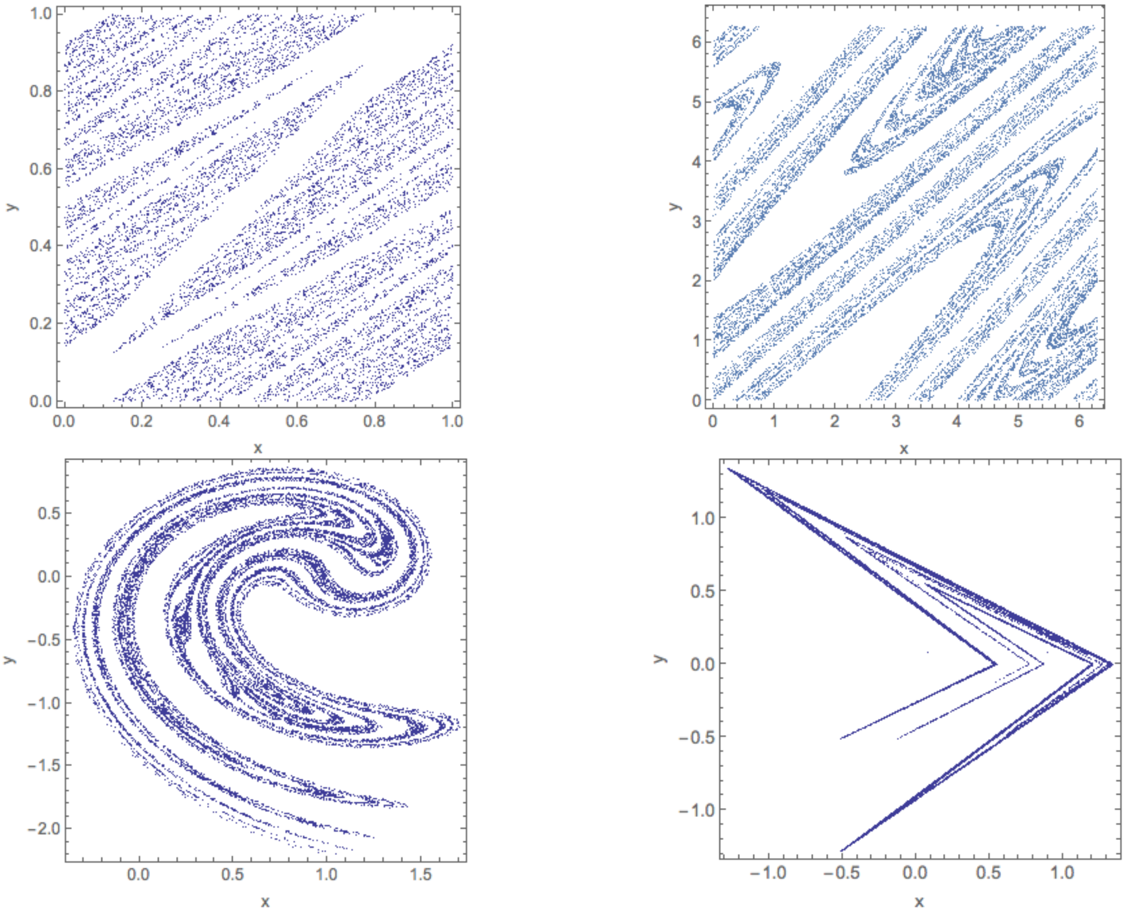


Fig. 3: Parametric plots of discrete chaotic maps; upper left – Arnold Cat map, upper right Dissipative standard map, bottom left Ikeda map, bottom right Lozi map.

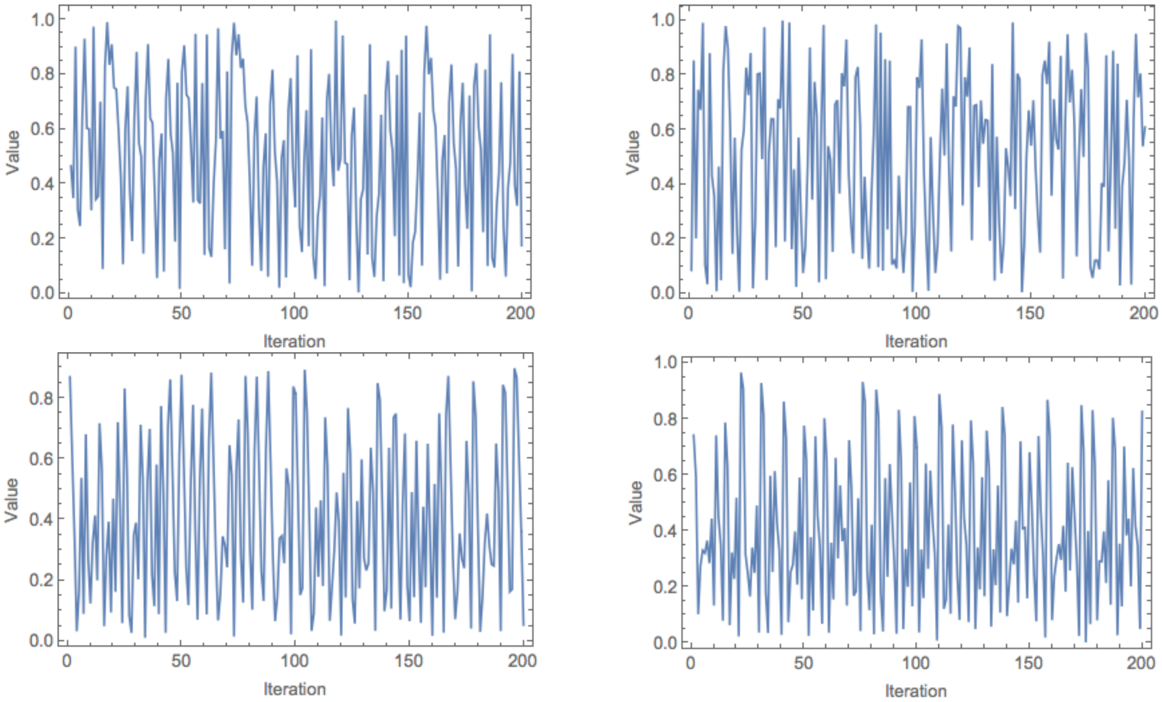


Fig. 4: Comparison of dynamical sequencing of pseudo random real numbers transferred into the range <0 - 1>; upper left – Arnold Cat map, upper right Dissipative standard map, bottom left Ikeda map, bottom right Lozi map.

EXPERIMENT DISCUSSION

The main aim of this research is to try, test, analyze the usability and compare the implementation of different natural chaotic dynamic as the mapping procedure for the optimization/searching process. Eight different complex chaotic time-continuous flows/oscillators and discrete chaotic maps have been simulated and the output behaviors have been transferred into the pseudo random number sequences. Findings can be summarized as follows:

- In many research works, it was proven that chaos based optimization is very sensitive to the hidden chaotic dynamics driving the CPRNG/mapping the search space. Such a chaotic dynamics can be significantly changed by the selection of sampling time in the case of the time-continuous systems (both flows and oscillators). Very small sampling rate of approx. 0.1s - 0.5s keeps the information about the chaotic dynamics inside the generated pseudo-random sequence. By changing or simple learning adaptation of such sampling rate, we can fully keep, partially suppress or even fully remove the chaotic information from CPRNG sequence. There is no such a control possibility for discrete maps.
- Oscillators are giving more dynamical pseudo random sequences with unique quasi-periodical sequencing in comparison with chaotic flows (See Figs 1 and 2). Therefore chaotic oscillators are more suitable for CPRNG.
- When comparing the time continuous systems and discrete maps – the first group is suitable only for direct CPRNG purposes (and preferably oscillators as stated in the previous point). Mapping of those systems to the optimization search space will lead only to covering a limited area, where the chaotic attractor is moving (See Figure 1). Again, in case of chaotic flows, these areas will be restricted only close to the cycles of the attractor. Oscillators are showing better coverage of the space. Selected discrete chaotic maps are depicted in graphics grid (Figure 3) and sorted from the highest density of coverage to the least. The graphical data in Figure 3 lends weight to the argument, that some chaotic maps support the basic claim and feature of deterministic chaos. This feature is called density of periodic orbits, assuming that chaotic attractor (system) will visit most of the points in the space.
- Even though the coverage of search space is lower (Ikeda map) or very limited (Lozi map), these two chaotic maps can be combined together within some hybrid multidimensional mapping.
- Furthermore presented chaotic systems have additional accessible parameters, which can by tuned. This issue opens up the possibility of examining the impact of these parameters to generation of random numbers, and thus influence on chaos based optimization (including adaptive switching between chaotic systems or sampling rates). The impact of parameter changing for chaotic map is demonstrated in Figures 5 – 10. Very small change of original settings for Dissipative map resulted in increasing of the search space mapping density (See Figure 7 compared to the Figure 3). The change of chaotic dynamics is also transferred to different search trajectories (Figures 5, 8 and 8) and distributions (Figures 6 and 9).

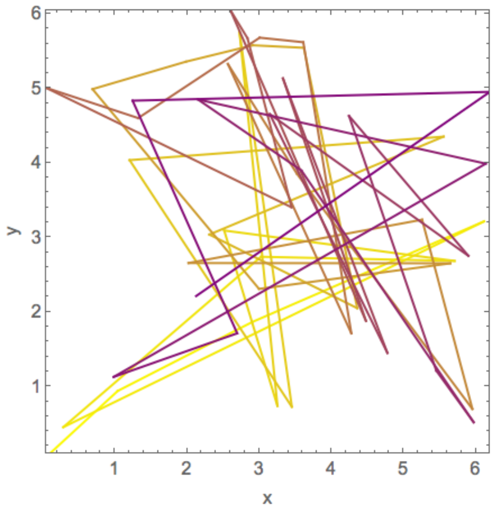


Fig. 5: Mapping of chaotic dynamics to the search space: Dissipative standard map – with original parameters b = 0.6 and k = 8.8. Coloring of path – from yellow to purple.

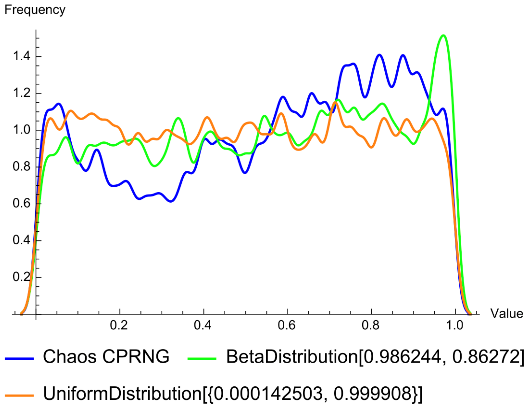


Fig. 6: Original chaotic CPRNG and identified distributions for Dissipative standard map – with original parameters b = 0.6 and k = 8.8 (5000 samples).

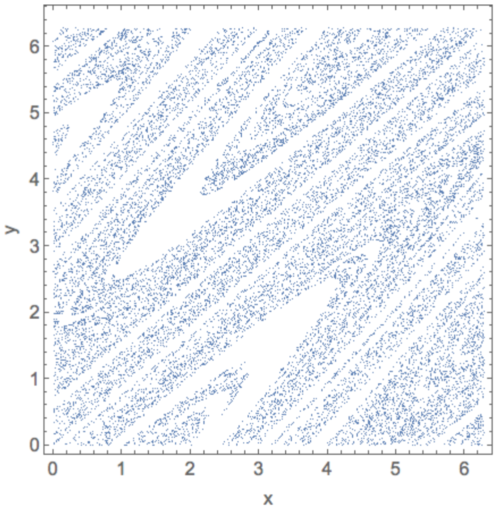


Fig. 7: Parametric plots of Dissipative standard map – with changed parameters b = 0.8 and k = 8.8

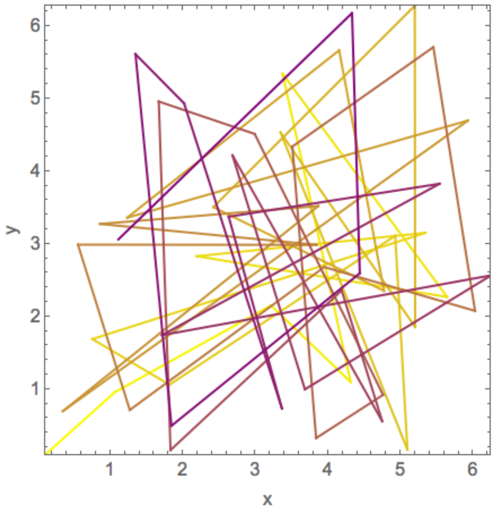


Fig. 8: Mapping of chaotic dynamics to the search space: Dissipative standard map – with changed parameters b = 0.8 and k = 8.8. Coloring of path – from yellow to purple.

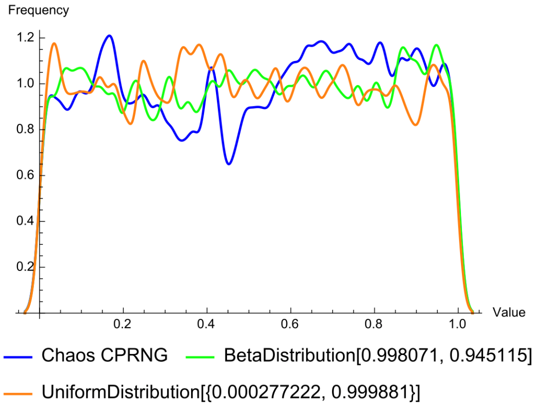


Fig. 9: Original chaotic CPRNG and identified distributions for Dissipative standard map – with changed parameters b = 0.8 and k = 8.8 (5000 samples).

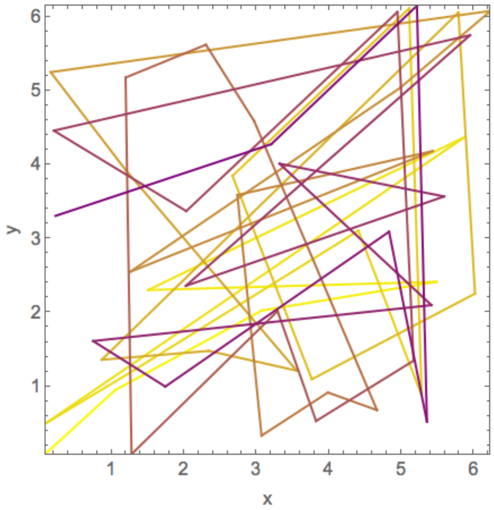


Fig. 10: Mapping of chaotic dynamics to the search space: Dissipative standard map – with changed parameters b = 0.9 and k = 8.6. Coloring of path – from yellow to purple.

CONCLUSION

The novelty of this research represents investigation on the utilization of the complex chaotic dynamics given by the several selected time-continuous chaotic systems, as the chaotic pseudo random number generators and driving maps for the chaos based optimization (mapping of chaotic movement to the optimization/search space). This paper showed three groups of complex chaotic dynamics given by chaotic flows, oscillators and discrete chaotic maps.

Future plans are including the testing of combination of wider set of chaotic systems as well as the adaptive switching between systems, adaptive or chaos-like sequencing sampling rates and obtaining a large number of results to perform statistical tests.

ACKNOWLEDGEMENT

This work was supported by Grant Agency of the Czech Republic - GACR P103/15/06700S, further by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme project No. LO1303 (MSMT-7778/2014) and also by the European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089, and by Internal Grant Agency of Tomas Bata University, No. IGA/CebiaTech/2017/004.

REFERENCES

Aydin, I., Karakose, M. and Akin, E. (2010) 'Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection', *Expert Systems with Applications,* 37(7), 5285-5294.

Bharti, L. and Yuasa, (2010) M. Energy Variability and Chaos in Ueda Oscillator. Available: http://www.rist.kindai.ac.jp/no.23/yuasa-EVCUO.pdf

Bucolo, M., Caponetto, R., Fortuna, L., Frasca, M. and Rizzo, A. (2002) 'Does chaos work better than noise?', *Circuits and Systems Magazine, IEEE,* 2(3), 4-19.

Caponetto, R., Fortuna, L., Fazzino, S. and Xibilia, M. G. (2003) 'Chaotic sequences to improve the performance of evolutionary algorithms', *IEEE Transactions on Evolutionary Computation,* 7(3), 289-304.

Celikovsky, S. and Zelinka, I. (2010) 'Chaos Theory for Evolutionary Algorithms Researchers' in Zelinka, I., Celikovsky, S., Richter, H. and Chen, G., eds., *Evolutionary Algorithms and Chaotic Systems*, Springer Berlin Heidelberg, 89-143.

Coelho, L. d. S. and Mariani, V. C. (2009) 'A novel chaotic particle swarm optimization approach using Hénon map and implicit filtering local search for economic load dispatch', *Chaos, Solitons & Fractals,* 39(2), 510-518.

Davendra, D., Zelinka, I. and Senkerik, R. (2010) 'Chaos driven evolutionary algorithms for the task of PID control', *Computers & Mathematics with Applications,* 60(4), 1088-1104.

Hamaizia, T. and Lozi, R. (2011) *Improving Chaotic Optimization Algorithm using a new global locally averaged strategy,* translated by pp. 17-20.

Lee, J. S. and Chang, K. S. (1996) 'Applications of chaos and fractals in process systems engineering', *Journal of Process Control,* 6(2–3), 71-87.

Liang, W., Zhang, L. and Wang, M. (2011) 'The chaos differential evolution optimization algorithm and its application to support vector regression machine', *Journal of Software,* 6(7), 1297- 1304.

Lozi, R. (2012) 'Emergence of Randomness from Chaos', *International Journal of Bifurcation and Chaos,* 22(02), 1250021.

Narendra, K. P., Vinod, P. and Krishan, K. S. (2010) 'A Random Bit Generator Using Chaotic Maps', *International Journal of Network Security,* 10, 32 - 38.

Persohn, K. J. and Povinelli, R. J. (2012) 'Analyzing logistic map pseudorandom number generators for periodicity induced by finite precision floating-point representation', *Chaos, Solitons & Fractals,* 45(3), 238-245.

Pluhacek, M., Senkerik, R. and Zelinka, I. (2014) 'Multiple Choice Strategy Based PSO Algorithm with Chaotic Decision Making – A Preliminary Study' in Herrero, Á., Baruque, B., Klett, F., Abraham, A., Snášel, V., Carvalho, A. C. P. L. F., Bringas, P. G., Zelinka, I., Quintián, H. and Corchado, E., eds., *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*, Springer, 21-30.

Pluhacek, M., Senkerik, R., Davendra, D., Kominkova Oplatkova, Z. and Zelinka, I. (2013) 'On the behavior and performance of chaos driven PSO algorithm with inertia weight', *Computers & Mathematics with Applications,* 66(2), 122-134.

Senkerik, R., Pluhacek, M., Zelinka, I., Oplatkova, Z., Vala, R. and Jasek, R. (2014) 'Performance of Chaos Driven Differential Evolution on Shifted Benchmark Functions Set' in Herrero, Á., Baruque, B., Klett, F., Abraham, A., Snášel, V., Carvalho, A. C. P. L. F., Bringas, P. G., Zelinka, I., Quintián, H. and Corchado, E., eds., *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*, Springer International Publishing, 41-50.

Senkerik R., Pluhacek M., Davendra D., Zelinka I., Kominkova Oplatkova Z. (2015). Simulation Of Time-Continuous Chaotic UEDA Oscillator As The Generator Of Random Numbers For Heuristic, ECMS 2015 Proceedings edited by: Valeri M. Mladenov, Petia Georgieva, Grisha Spasov, Galidiya Petrova European Council for Modeling and Simulation. doi:10.7148/2015-0543

Senkerik R., Pluhacek M., Viktorin A, Kominkova OplatkovaZ. (2016). On The Simulation Of Complex Chaotic Dynamics For Chaos Based Optimization, ECMS 2016 Proceedings edited by: Thorsten Claus, Frank Herrmann, Michael Manitz, Oliver Rose European Council for Modeling and Simulation. doi:10.7148/2016-0258.

Sprott, J. C. (2003) *Chaos and Time-Series Analysis,* Oxford University Press.

Wang, X.-y. and Qin, X. (2012) 'A new pseudo-random number generator based on CML and chaotic iteration', *Nonlinear Dynamics,* 70(2), 1589-1592.

Wu, J., Lu, J. and Wang, J. (2009) 'Application of chaos and fractal models to water quality time series prediction', *Environmental Modelling & Software,* 24(5), 632-636.

Yang, L. and Wang, X.-Y. (2012) 'Design of Pseudo-random Bit Generator Based on Chaotic Maps', *International Journal of Modern Physics B,* 26(32), 1250208.

Zamuda, A., Brest, J. (2015) Self-adaptive control parameters′ randomization frequency and propagations in differential evolution. Swarm and Evolutionary Computation 25, 72-99.

Zhenyu, G., Bo, C., Min, Y. and Binggang, C. (2006) 'Self-Adaptive Chaos Differential Evolution' in Jiao, L., Wang, L., Gao, X.-b., Liu, J. and Wu, F., eds., *Advances in Natural Computation*, Springer Berlin Heidelberg, 972-975.

DIFFERENT APPROACHES FOR CONSTANT ESTIMATION IN ANALYTIC PROGRAMMING

Zuzana Kominkova Oplatkova, Adam Viktorin, Roman Senkerik, Tomas Urbanek

Tomas Bata University in Zlin, Faculty of Applied Informatics
Nam T.G. Masaryka 5555, 760 01 Zlin, Czech Republic
{oplatkova, aviktorin, senkerik, turbanek}@fai.utb.cz

KEYWORDS
Analytic programming, Differential evolution, approximation, sextic, quintic.

ABSTRACT

This research deals with different approaches for constant estimation in analytic programming (AP). AP is a tool for symbolic regression tasks which enables to synthesise an analytical solution based on the required behaviour of the system. Some tasks do not need any constant estimation - AP is used in its basic version without any constant estimation handling. Compared to this, cases like data approximation need constants (coefficients) which are essential for the process of precise solution synthesis. This paper offers another strategy to already known and used by the AP from the very beginning and approaches published recently in 2016. This paper compares these procedures and the discussion also includes nonlinear fitting and metaevolutionary approach. As the main evolutionary algorithm, a differential algorithm (de/rand/1/bin) for the main process of AP is used.

INTRODUCTION

Analytic Programming (AP) (Zelinka et al., 2011) is a tool of symbolic regression which uses techniques from the area of evolutionary computation techniques (EVT). The basic case of a regression represents a process in which the measured data is fitted and a suitable mathematical formula is obtained in an analytical way. This process is widely known for mathematicians. They use this process when a need arises for a mathematical model of unknown data, i.e. the relation between input and output values. Classical regression usually requires to select an expected type of model in advance and a suitable method is applied for the coefficient estimation of the proposed model. Compared to that, symbolic regression in the context of EVT means to build a complex formula from basic operators defined by users. The final shape of the expression is managed to breed via evolutionary optimisation algorithms.

Initially, John Koza proposed the idea of symbolic regression done by means of a computer in Genetic Programming (GP) (Back et al., 1997), (Koza, 1998), (Koza, 1999). The other approaches are e.g. Grammatical Evolution (GE) developed by Conor Ryan (O'Neill et al., 2003) and some others included Analytic Programming (Zelinka et al., 2011). The symbolic regression can be used for different tasks: data approximation, design of electronic circuits, optimal trajectory for robots, classical neural networks and pseudo neural networks synthesis (Oplatkova, 2016) and many other applications (Back et al., 1997), (Koza, 1998), (Koza, 1999), (O'Neill et al., 2003), (Zelinka et al., 2011), (Oplatkova, 2009), (Varacha et al., 2006), (Volna et al., 2013). The results and usage depend on the user-defined set of operators and their possible combinations and nesting into themselves.

This paper deals with strategies and their comparison for constants (coefficients) estimation - nonlinear fitting (Zelinka et al., 2011), metaevolutionary approach (Zelinka et al., 2011) and direct encoding in the individuals (extended individual (Viktorin et al, 2016) or a special handling with an individual (Urbanek et al., 2016) and a stance proposed in this paper.

ANALYTIC PROGRAMMING

Basic principles of the AP were developed in 2001 (Zelinka et al., 2005), (Zelinka et al., 2008), (Zelinka et al., 2011).

The core of AP is based on a special set of mathematical objects and operations. The collection of mathematical objects is the set of functions, operators and terminals, which are usually constants or independent variables. Various functions and terminals can be mixed in this set. This set is called general functional set (GFS) due to its variability of the content. The structure of GFS is created by subsets of functions according to the number of their arguments. For example, GFSall is a set of all functions, operators and terminals, GFS3arg is a subset containing functions with only three arguments, GFS0arg represents only terminals, etc. The subset structure presence in GFS is of vital importance for AP. It is used to avoid synthesis of pathological programs, i.e. programs containing functions without arguments, etc. The content of GFS is dependent only on the user (Zelinka et al., 2005), (Zelinka et al., 2008), (Oplatkova, 2009).

The second part of the AP core is a sequence of mathematical operations, which are used for the program synthesis. These operations are used to transform an individual of a population into a suitable program. Mathematically stated, it is a mapping from an individual domain into a program domain. This

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

mapping consists of two main parts. The first part is called discrete set handling (DSH) (See Figure 1) (Zelinka et al., 2005), (Lampinen and Zelinka, 1999) and the second one stands for security procedures which do not allow synthesising pathological programs. The method of DSH, when used, allows handling arbitrary objects including nonnumerical objects like linguistic terms {hot, cold, dark…}, logic terms (True, False) or other user defined functions. In the AP DSH is used to map an individual into GFS and together with security procedures creates the mapping mentioned above which transforms the arbitrary individual into a program.

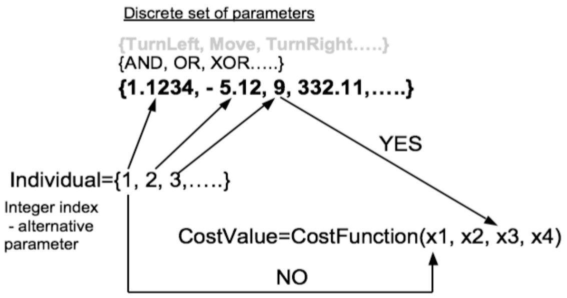

Discrete set of parameters
{TurnLeft, Move, TurnRight.....}
{AND, OR, XOR.....}
{1.1234, - 5.12, 9, 332.11,.....}
YES
Individual={1, 2, 3,.....}
Integer index - alternative parameter
CostValue=CostFunction(x1, x2, x3, x4)
NO

Figure 1: Discrete set handling

AP needs some evolutionary algorithm (Zelinka, 2004) that consists of a population of individuals for its run. Individuals in the population consist of integer parameters, i.e. an individual is an integer index pointing into GFS. The creation of the program can be schematically observed in Fig. 2.

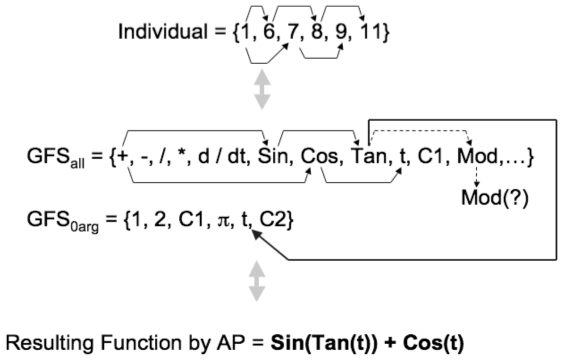

Individual = {1, 6, 7, 8, 9, 11}
GFSall = {+, -, /, *, d / dt, Sin, Cos, Tan, t, C1, Mod,…}
Mod(?)
GFS0arg = {1, 2, C1, π, t, C2}
Resulting Function by AP = Sin(Tan(t)) + Cos(t)

Figure 2: Main principles of AP

An example of the process of the final complex formula synthesis (according to the Fig. 2) follows.

The number 1 in the position of the first parameter means that the operator plus (+) from GFS$_{all}$ is used (the end of the individual is far enough). Because the operator + must have at least two arguments, the next two index pointers 6 (sin from GFS) and 7 (cos from GFS) are dedicated to this operator as its arguments. The two functions, sin and cos, are one-argument functions; therefore the next unused pointers 8 (tan from GFS) and 9 (*t* from GFS) are dedicated to the sin and cos functions. As an argument of cos, the variable *t* is used, and this part of the resulting function is closed (*t* has zero arguments) in its AP development. The one-argument function tan remains, and there is one unused pointer 11, which stands for Mod in GFS$_{all}$. The modulo operator needs two arguments but the individual in the example has no other indices (pointers, arguments). In this case, it is necessary to employ security procedures and jump to the subset with GFS$_{0arg}$. The function tan is mapped on *t* from GFS$_{0arg}$ which is in the 11th position, cyclically from the beginning. The detailed description is represented in (Zelinka et al., 2005), (Zelinka et al., 2008), (Oplatkova et al., 2009).

ANALYTIC PROGRAMMING - VERSIONS

The above-described version is the basic one AP$_{basic}$ (Zelinka et al., 2005) - without constant estimation. Such approach is used for tasks like logic circuit design where numerical coefficients are not usually used. It can also be applied to a pre-generated set of numerical values as in genetic programming where e.g. 4000 random numerical values of constants are selected. They are used as standard terminals like variable x.

When a constant estimation is necessary, e.g. in data approximation or pseudo neural network synthesis, etc., firstly general approach is applied which is different from genetic programming technique where all constants (e.g. 4000 random generated values) were part of the nonterminal and terminal sets.

AP uses the constant K (Zelinka et al., 2005) which is indexed during the evolution (1) - (3). The K is a terminal, i.e. GFS$_{0arg}$. So it is used as a standard terminal, e.g. similar to variable x in the evolutionary process (1). When K is needed, a proper index is assigned – K$_1$, K$_2$, ... K$_n$ (2). Numeric values of indexed Ks are estimated (3) via different techniques - AP$_{nf}$ (nonlinear fitting package in Mathematica) (Zelinka et al., 2005), AP$_{meta}$,(metaevolutionary approach with a second/slave evolutionary algorithm) (Zelinka et al., 2005, Oplatkova 2009) and three novel direct approaches AP$_{extend}$ (extended individual - a part of it for AP and the rest of it for constant estimation) (Viktorin et al., 2016), AP$_{direct1}$ (the part behind decimal point determines the K from the selected range) (Urbanek et al., 2016) and AP$_{direct2}$ (new proposed approach in this paper - the whole value determines the K from the selected range). The first 3 versions of AP from its very beginning have been extended by two other approaches in 2016 (Viktorin et al., 2016), (Urbanek et al., 2016).

$$\frac{x^2+K}{\pi^K} \tag{1}$$

$$\frac{x^2+K_1}{\pi^{K_2}} \tag{2}$$

$$\frac{x^2+3.156}{\pi^{90.78}} \tag{3}$$

AP$_{nf}$ - Nonlinear fitting version

The estimation of constants K has been done via a package for nonlinear fitting in Wolfram Mathematica environment (www.wolfram.com). The used function was FindFit, which includes different methods. Documentation refers to Conjugate Gradient, Gradient, Levenberg-Marquardt, Newton, NMinimize and Quasi-Newton. When one searches deeper, it can be found that NMinimize includes following techniques Nelder-Mead, Random Search, Simulated Annealing and Differential Evolution. Cost function evaluations for AP$_{nf}$ were not interpreted correctly in previous publications. The package from Mathematica environment contains techniques which belong to a group of iterative algorithms. Thus, each constant estimation of the particularly found model is not only one step evaluation but many iterations are needed. The method is selected automatically in Mathematica. Simple tests showed that mostly around 5000 iterations are necessary in the case of described sextic and quintic problems. The result does not mean necessarily that the found model is perfectly precise. For the suggested model, the nonlinear fitting tries to find the best constants (coefficients). The final cost value comes from the performed nonlinear fitting process - the error between required and actual obtained model.

On that account, authors think that nonlinear fitting case is a specific approach of AP$_{meta}$ (metaevolutionary approach), except the second slave algorithm does not need to be only an evolutionary algorithm. Above mentioned methods might be even faster and more precise for this particular task.

AP$_{meta}$ - metaevolutionary approach

Generally, metaevolution means the evolution of evolution. Several directions as the usage of an evolutionary algorithm for tuning or controlling of another evolutionary technique or the evolutionary design of evolutionary algorithms are discussed for instance in (Diosan, 2009), (Edmons, 2001), (Jones, 2002), (Oplatkova, 2009), (Kordik, 2010), Deugo, 2004), (Eiben, 2007).

In AP$_{meta}$, the metaevolution means that one evolutionary algorithm drives the main process of symbolic regression and the second is used for the constant estimation. This meta approach of analytic programming is used when the constants are not possible to estimate by AP$_{nf}$ because of the character of the problem. In data approximation tasks, a technique from non-linear fitting package can be used easily because the problem is designed so that the found constants (e.g. coefficients of polynomials) move the basic shape of the curve around the coordinate system. However, it is not possible to employ such a package in the case of the synthesis of more sophisticated problems, for instance, pseudo neural networks synthesis (Kominkova Oplatkova 2016). These applications do not use the found result as a model which could be adjusted to some “measured” values in the sense of interpolation but the obtained solution is used further as a part of the complex technique to find a quality of the solution and cost function estimation.

AP$_{meta}$ is a time-consuming process and the number of cost function evaluations, which is one of the comparable factors, is usually very high. This fact is given by two evolutionary procedures (Fig. 3).

$$EA_{master} \Rightarrow program \Rightarrow K_{indexing} \Rightarrow EA_{slave} \Rightarrow K_{estimation} \Rightarrow final \cdot solution$$

Figure 3: Schema of AP procedures

EA$_{master}$ is the main evolutionary algorithm for AP, EA$_{slave}$ is the second evolutionary algorithm inside AP. Thus, the number of cost function evaluation (CFE) is given by (4).

$$CFE = EA_{master} * EA_{slave} \tag{4}$$

As mentioned in the last paragraph of AP$_{nf}$ section, nonlinear fitting (NF) methods adopted in Mathematica environment are iterative processes. Thus, EA$_{slave}$ in the case of AP$_{nf}$ would be a number of iterations of used NF method.

The following three approaches were developed to find a suitable constant estimation which will decrease the number of cost function evaluations to (5).

$$CFE = EA_{master} \tag{5}$$

AP$_{extended}$ - extended individual

The constant handling technique with an extended individual was introduced in (Viktorin et al, 2016). The individual used in AP has an extended part which is used for the evolution of constant values.

The important task was to determine what the correct size of an extension is (6).

$$k = l - floor\big((l-1)/(\max\_\arg)\big), \tag{6}$$

where k is the maximum number of constants that can appear in the synthesised program (extension) of length l and *max_arg* is the maximum number of arguments needed by functions in GFS. Also, the *floor()* is a common floor round function. The final individual dimensionality (length) will be $k+l$ and the example might be:

- Program length $l = 10$
- GFS: {+, -, *, /, *sin*, *cos*, *x*, *k*}
- GFS maximum argument *max_arg = 2*
- Extension size $k = 10 - floor((10-1) / 2) = 6$
- Dimensionality of the extended individual $k+l = 16$

This means, that the EA will work with individuals of length 16, but only first 10 features will be used for indexing into the GFS and the rest will be used as constant values.

It is worthwhile to note that only features which are going to be mapped to GFS are rounded and the rest is omitted (not rounded). An example can be viewed in

Fig. 3. Individual features in bold are the constant values.

$$Individual = \left\{ \begin{matrix} 5.08, 1.64, 6.72, 1.09, 6.20, \\ 1.28, \mathbf{0.07}, \mathbf{3.99}, \mathbf{5.27}, \mathbf{2.64} \end{matrix} \right\}$$

$$Rounded\ individual = \{5, 2, 7, 1, 6, 1\}$$

$$GFS_{all} = \{+, -, *, /, sin, cos, x, k\}$$

$$\boldsymbol{Program}\text{: } \mathbf{cos}(\boldsymbol{k1} * (\boldsymbol{x} - \boldsymbol{k2}))$$

$$\boldsymbol{Replaced}\text{: } \mathbf{cos}(\mathbf{0.07} * (\boldsymbol{x} - \mathbf{3.99}))$$

Figure 3: Principles of AP$_{\text{extended}}$

***AP$_{direct1}$* - direct encoding of K in the individual 1**

This constant handling technique was introduced in (Urbanek et al, 2016). It works with a direct encoding in an individual and is based on a part behind a decimal point which as a proportional pointer determines the value from the selected range of K.
The part behind decimal point is obtained from (7).

$$ind_K = \left| ind - ind_f \right|, \tag{7}$$

where $ind = \{x_1, x_2, x_3 \ldots x_n\}$ and $ind_f = \{floor(x_1), floor(x_2), floor(x_3) \ldots . floor(x_n)\}$
The decimal values in ind$_K$ are in the interval <0,1>. The corresponding K is then computed easily from (8).

$$K = ind_K * \left| rangeK_{\max} - rangeK_{\min} \right| + rangeK_{\min} \tag{8}$$

The mapping is done in the standard procedure as in AP$_{\text{basic}}$ and general approach of K indexation. When K is needed, the value in the corresponding position from (8) is directly used.

***AP$_{direct2}$* - direct encoding of K in the individual 2**

Within a later analysis of AP$_{\text{direct1}}$ behaviour, authors found out some problematic issues connected with the neighbourhood of arguments which are responsible for K estimation. Since they are dependent only on the decimal part of the argument regardless the integer part of the value, two points placed on the opposite sides of the coordinate system can be neighbours from ind$_K$ point of view. It does not help the evolutionary optimisation process which expects for a successful performance that two points lie next to each other physically in the coordinate system.
This new approach is based on the previous and above-described AP$_{\text{direct1}}$ (Urbanek et al, 2016). The difference is in the different computation of ind$_K$ (9). It takes the value of the not rounded individual as the proportional part in respect of length of all components in GFS$_{\text{All}}$.

$$ind_K = \frac{ind}{Dim(GFS_{All})}, \tag{9}$$

where $ind = \{x_1, x_2, x_3 \ldots x_n\}$ and Dim(GFS$_{\text{All}}$) means the number of all non-terminals and terminals used in AP. For instance, if GFS$_{\text{All}}$={+, -, /, *, x, K}, the Dim(GFS$_{\text{All}}$) = 6 and the valid range for arguments in the individual is in the interval <1,6>.

USED EVOLUTIONARY ALGORITHM - DIFFERENTIAL EVOLUTION

As mentioned above, the Analytic Programming needs an evolutionary algorithm for the optimisation - finding the best shape of the complex formula. This research used Differential Evolution (Price, 2005) in its canonical version DE/Rand/1/Bin. Future research expects to use some other strategies as DE/Best/1/Bin or SHADE which was quite promising in (Viktorin et al., 2016).
DE is a population-based optimisation method that works on real-number-coded individuals (Price, 2005). For each individual $\vec{x}_{i,G}$ in the current generation G, DE generates a new trial individual $\vec{x}'_{i,G}$ by adding the weighted difference between two randomly selected individuals $\vec{x}_{r1,G}$ and $\vec{x}_{r2,G}$ to a randomly selected third individual $\vec{x}_{r3,G}$. The resulting individual $\vec{x}'_{i,G}$ is crossed-over with the original individual $\vec{x}_{i,G}$. The fitness of the resulting individual, referred to as a perturbed vector $\vec{u}_{i,G+1}$, is then compared with the fitness of $\vec{x}_{i,G}$. If the fitness of $\vec{u}_{i,G+1}$ is greater than the fitness of $\vec{x}_{i,G}$, then $\vec{x}_{i,G}$ is replaced with $\vec{u}_{i,G+1}$; otherwise, $\vec{x}_{i,G}$ remains in the population as $\vec{x}_{i,G+1}$. DE is quite robust, fast, and effective, with global optimisation ability. It does not require the objective function to be differentiable, and it works well even with noisy and time-dependent objective functions. Description of used DERand1Bin mutation strategy is presented in (10). Please refer to (Price and Storn 2001, Price 2005) for the description of all other strategies.

$$u_{i,G+1} = x_{r1,G} + F \bullet \left(x_{r2,G} - x_{r3,G} \right) \tag{10}$$

PROBLEM DESIGN

These above-mentioned strategies of AP$_{\text{extended}}$, AP$_{\text{direct1}}$ and AP$_{\text{direct2}}$ were applied on standard benchmark tests - approximation of polynomial expression - quintic (11) and sextic (12).

$$x^5 - 2x^3 + x \tag{11}$$

$$x^6 - 2x^4 + x^2 \tag{12}$$

RESULTS AND DISCUSSION

The paper will compare AP$_{\text{extended}}$, AP$_{\text{direct1}}$ and AP$_{\text{direct2}}$ strategies with differential evolution DE/Rand/1/Bin. The setting was based on some previous research in this field (Tab. 1.).

Table 1: DE settings

PopSize	50
F	0.5
CR	0.8
Generations	4000
Max. CF Evaluations (CFE)	200 000

The previously published results (Oplatkova, 2009) stated that the cost function evaluations for quintic and sextic problems were in the interval <500, 18 000>. Compared to these already published results with AP_{nf}, it seems that currently, we do not provide any improvement (Tab. 1.) in the sense of convergence speed. As already mentioned, AP_{nf} is a specific case of AP_{meta}. Therefore the correct statement of the cost function evaluations should be <500, 18 000> * cca 5000 iterations which is equal to <2 500 000, 90 000 000>. Thus, our setting in Tab.1. means the significant reduction of computation time.

All simulations were performed 30 times out. The results for the quintic problem are depicted in Fig. 4 - Fig. 6.

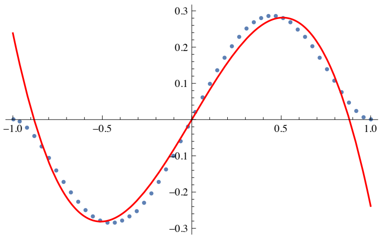


Figure 4: Quintic problem with $AP_{extended}$

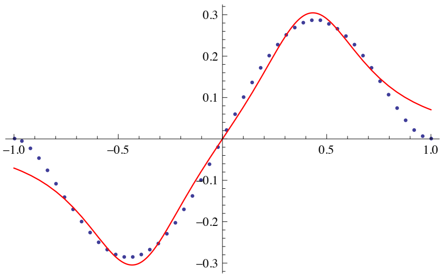


Figure 5: Quintic problem with $AP_{direct1}$

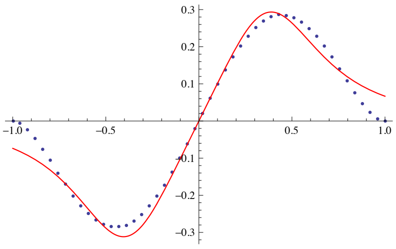


Figure 6: Quintic problem with $AP_{direct2}$

The results for the sextic problem are depicted in Fig. 7. - Fig. 8.

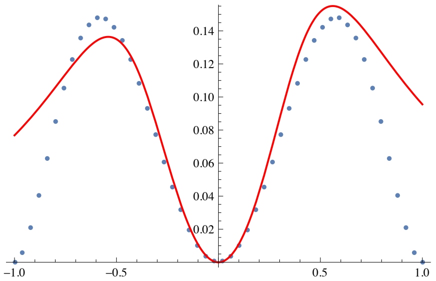


Figure 7: Sextic problem with $AP_{extended}$

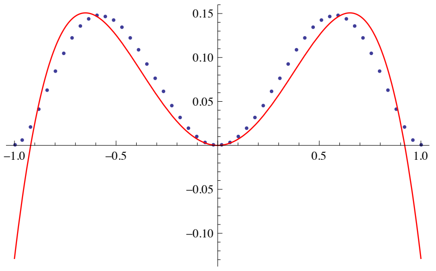


Figure 8: Sextic problem with $AP_{direct1}$

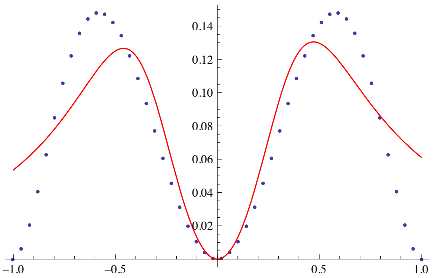


Figure 9: Sextic problem with $AP_{direct2}$

Tab. 2. and Tab. 3. show statistical measures from the performed simulations.

Table 2: Statistical results for quintic

	$AP_{extended}$	$AP_{direct1}$	$AP_{direct2}$
Min	1.84172	1.14635	1.2115
Max	3.05308	2.82948	2.41022
Avg	2.46638	1.90875	1.90549
Median	2.53626	2.07038	1.86441
St.Dev.	0.280835	0.508758	0.394415

Table 3: Statistical results for sextic

	$AP_{extended}$	$AP_{direct1}$	$AP_{direct2}$
Min	1.07682	1.02398	1.02347
Max	2.37171	2.08186	2.41894
Avg	1.86946	1.61992	1.675
Median	1.85952	1.71677	1.7273
St.Dev.	0.377488	0.269501	0.3595

The results showed that $AP_{extended}$, $AP_{direct1}$ and $AP_{direct2}$ are comparable in the achieved results.

The evolution process within $AP_{direct1}$ and $AP_{direct2}$ was carried out longer ($20x200000 = 4x10^6$ CFE) for possible further analysis. The results for the sextic problem can be found in Tab. 4.

Table 4: Statistical results for sextic - 4 000 000 CFE

	$AP_{direct1}$	$AP_{direct2}$
Min	0.471035	0.094396
Max	1.11581	1.0897
Avg	0.740252	0.76077
Median	0.73997	0.843761
St.Dev.	0.155372	0.2293

The post analysis showed that the process got often stuck in local optima for a long time, e.g., Fig.10 depicts the history of cost function evaluation on one example of one $AP_{direct2}$ run. However, some quality solutions (e.g. Fig. 11) were obtained after circa CFE equal to 2 000 000 which is still significantly less than AP_{nf}.

The results also proved that the assumption of authors which led to $AP_{direct2}$ proposal was wrong. The evolution can work even with the individuals who assume non physical neighbourhood.

The future plans include to leave the evolution in the process when the acceptable error will be reached. The final number of cost function evaluations will be compared.

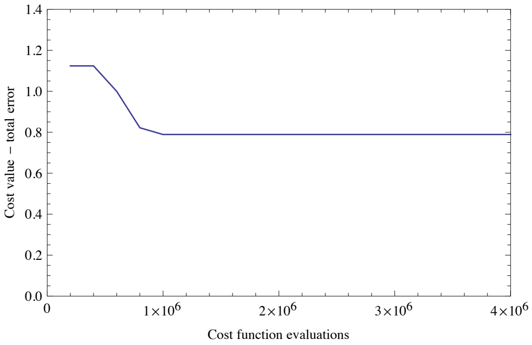


Figure 10: CFE history of one $AP_{direct2}$ run

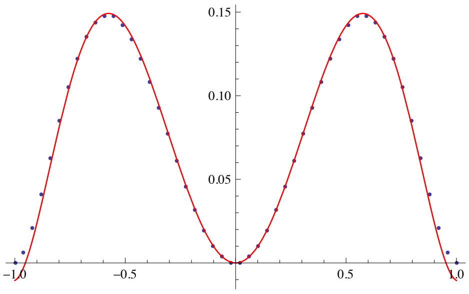


Figure 11: Sextic problem with $AP_{direct2}$ and DE

Since (Viktorin, 2016) presented better results with SHADE strategy of differential evolution for the same setting for population size and the number of generations (Tab. 1), future plans include usage of different strategies of DE and different evolutionary algorithms as Self-organizing migrating algorithm, Particle swarm algorithm and others. The best option for mentioned SHADE strategy was Min = 0.000139781 for the quintic (Fig. 12) and Min= 0.113134 for the sextic which secure very precise fitting.

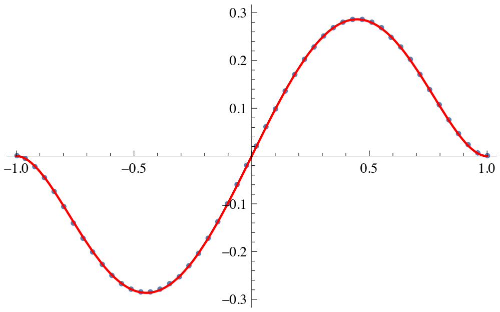


Figure 12: Quintic problem with $AP_{extended}$ and SHADE

CONCLUSION

This paper deals with Analytic programming and compares three novel approaches for constant estimation - $AP_{extended}$, $AP_{direct1}$ and $AP_{direct2}$. All simulations were performed with a DE/Rand/1/Bin strategy of differential evolution algorithm.

The results showed that all three approaches are comparable and use significantly less number of cost function evaluations than AP_{nf} or AP_{meta}.

Future plans include - comparison of cost function evaluations for these three mentioned approaches when an acceptable error is reached. Certainly, other evolutionary techniques as for instance SHADE will be employed for further testing.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme project No. LO1303 (MSMT-7778/2014) and also by the European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089, further it was supported by Grant Agency of the Czech Republic—GACR P103/15/06700S and by Internal Grant Agency of Tomas Bata University in Zlin under the project No. IGA/CebiaTech/2017/004.

REFERENCES

Back T., Fogel D. B., Michalewicz Z., *Handbook of evolutionary algorithms*, Oxford University Press, 1997, ISBN 0750303921

Deugo D., Ferguson D.: Evolution to the xtreme: Evolving evolutionary strategies using a meta-level approach, Proceedings of the 2004 IEEE congress on evolutionary computation, IEEE Press, Portland, Oregon, pp. 31–38, 2004

Dioşan, L., Oltean, M.: Evolutionary design of evolutionary algorithms. Genetic Programming and Evolvable Machines, Vol. 10, Issue 3, p. 263-306, 2009

Edmonds, B.: Meta-genetic programming: Co-evolving the operators of variation, Elektrik, Vol. 9, Issue 1, pp. 13-29, 2001

Eiben A.E., Michalewicz Z., Schoenauer M., Smith J.E.: Parameter control in evolutionary algorithms, pp. 19–46, Springer, 2007

Jones D.F., Mirrazavi S.K., Tamiz M.: Multi-objective meta-heuristics: An overview of the current state-of-the-art, European Journal of Operational Research, Volume 137, Issue 1, 16 February 2002, Pages 1-9, ISSN 0377-2217.

Kominkova Oplatkova Z., Senkerik R. (2013): Evolutionary Synthesis of Complex Structures - Pseudo Neural Networks for the Task of Iris Dataset Classification. In Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems. Heidelberg : Springer-Verlag Berlin, 2013, p. 211-220. ISSN 2194-5357. ISBN 978-3-319-00541-6.

Kominkova Oplatkova Z., Senkerik R. (2016): Control Law and Pseudo Neural Networks Synthesized by Evolutionary Symbolic Regression Technique, in Al-Begain K., Bargiela A.: Seminal Contributions to Modelling and Simulation - Part of the series Simulation Foundations, Methods and Applications, pp 91-113, doi: 10.1007/978-3-319-33786-9_9, ISBN: 978-3-319-33785-2.

Kordík P., Koutník J., Drchal J., Kovářík O., Čepek M., Šnorek M.: Meta-learning approach to neural network optimization, Neural Networks, Vol. 23, Issue 4, p. 568-582, 2010, ISSN 0893-6080.

Koza J. R. et al., *Genetic Programming III; Darwinian Invention and problem Solving*, Morgan Kaufmann Publisher, 1999, ISBN 1-55860-543-6

Koza J. R., *Genetic Programming,* MIT Press, 1998, ISBN 0-262-11189-6

Lampinen J., Zelinka I., 1999, "New Ideas in Optimization – Mechanical Engineering Design Optimization by Differential Evolution", Volume 1, London: McGraw-hill, 1999, 20 p., ISBN 007-709506-5.

O'Neill M., Ryan C., *Grammatical Evolution. Evolutionary Automatic Programming in an Arbitrary Language*, Kluwer Academic Publishers, 2003, ISBN 1402074441

Oplatkova Z.: Metaevolution: Synthesis of Optimization Algorithms by means of Symbolic Regression and Evolutionary Algorithms, Lambert Academic Publishing Saarbrücken, 2009, ISBN: 978-3-8383-1808-0

Price K., Storn R. M., Lampinen J. A., 2005, "Differential Evolution : A Practical Approach to Global Optimization", (Natural Computing Series)*,* Springer; 1 edition.

Price, K. and Storn, R. (2001), *Differential evolution homepage*, [Online]: http://www.icsi.berkeley.edu/~storn/code.html, [Accessed 29/02/2012].

Urbanek T., Prokopova Z., Silhavy R., Kuncar A.: New Approach of Constant Resolving of Analytical Programming. In *30th European Conference on Modelling and Simulation*, 2016, p. 231-236. ISBN 978-0-9932440-2-5.

Viktorin A., Pluhacek M., Kominkova Oplatkova Z., Senkerik R.: Analytical Programming with Extended Individuals. In *30th European Conference on Modelling and Simulation*, 2016, p. 237-244. ISBN 978-0-9932440-2-5.

Volna, E., Kotyrba, M., & Jarusek, R. (2013). Multi-classifier based on Elliott wave's recognition. Computers & Mathematics with Applications, 66(2), 213-225.

Zelinka et al.: Analytical Programming - a Novel Approach for Evolutionary Synthesis of Symbolic Structures, in Kita E.: Evolutionary Algorithms, InTech 2011, ISBN: 978-953-307-171-8

Zelinka I., Varacha P., Oplatkova Z., *Evolutionary Synthesis of Neural Network*, Mendel 2006 – 12th International Conference on Softcomputing, Brno, Czech Republic, 31 May – 2 June 2006, pages 25 – 31, ISBN 80-214-3195-4

Zelinka I.,Oplatkova Z, Nolle L., 2005. *Boolean Symmetry Function Synthesis by Means of Arbitrary Evolutionary Algorithms-Comparative Study*, International Journal of Simulation Systems, Science and Technology, Volume 6, Number 9, August 2005, pages 44 - 56, ISSN: 1473-8031.

AUTHOR BIOGRAPHIES

ZUZANA KOMINKOVA OPLATKOVA is an associate professor at Tomas Bata University in Zlin. Her research interests include artificial intelligence, soft computing, evolutionary techniques, symbolic regression, neural networks. She is an author of around 100 papers in journals, book chapters and conference proceedings. Her e-mail address is: oplatkova@fai.utb.cz

ADAM VIKTORIN was born in the Czech Republic, and went to the Faculty of Applied Informatics at Tomas Bata University in Zlín, where he studied Computer and Communication Systems and obtained his MSc degree in 2015. He is studying his Ph.D. at the same university and the field of his studies are: Artificial intelligence, data mining and evolutionary algorithms. His email address is: aviktorin@fai.utb.cz

ROMAN SENKERIK was born in the Czech Republic, and went to the Tomas Bata University in Zlin, where he studied Technical Cybernetics and obtained his MSc degree in 2004, Ph.D. degree in Technical Cybernetics in 2008 and Assoc. prof. in 2013 (Informatics). He is now an Assoc. prof. at the same university (research and courses in: Evolutionary Computation, Applied Informatics, Cryptology, Artificial Intelligence, Mathematical Informatics). His email address is: senkerik@fai.utb.cz

TOMAS URBANEK was born in Zlin in 1987. He received a B.Sc. (2009), M.Sc. (2011) in Information Technology from Faculty of Applied Informatics, Tomas Bata University in Zlin. He is a doctoral student at the Computer and Communication Systems Department. Major research interests are software engineering, effort estimation in software engineering and artificial intelligence.

Modelling, Simulation and Control of Technological Processes

MODELING OF CONTINUOUS ETHANOL FERMENTATION IN IDEAL MIXING COLUMN BIOREACTOR

Ivan Petelkov*, Rositsa Denkova**, Vesela Shopska* Georgi Kostov*, Zapryana Denkova***
*Department "Technology of wine and brewing" ** Department "Biochemistry and molecular biology"; *** Department "Microbiology"
University of Food Technologies, 4002, 26 Maritza blvd., Plovdiv, Bulgaria
E-mail: george_kostov2@abv.bg; vesi_nevelinova@abv.bg; zdenkova@abv.bg; rositsa_denkova@mail.bg; i_petelkov92@abv.bg;

Bogdan Goranov
LBLact, Plovdiv, Bulgaria
E-mail: goranov_chemistry@abv.bg

Vasil Iliev
Weissbiotech, Ascheberg, Germany
E-mail: illiev.vasil@gmail.com

KEYWORDS

ethanol, immobilized cells, modeling, bioreactor with ideal mixing

ABSTRACT

A method for formalization and analytical determination of the kinetic parameters of continuous alcohol fermentation in a reactor and in a cascade of reactors with ideal mixing was discussed in the present work. A fluidized bed bioreactor with immobilized yeast with elucidated structure of the flows was used. The method of formalization of the process kinetics involved two steps – determination of the optimal dilution rate of a continuous fermentation process, and determination of the total kinetic parameters of the cascade of reactors, including the minimum number of steps that can ensure production of maximum ethanol yield in the system.

INTRODUCTION

Ethanol is a fermentation product, which is widely used in food and chemical industry and as a bio-fuel. Liquid bio-fuels are divided into the following categories: (a) bio-alcohols; (b) vegetable oils and biodiesels; and (c) bio-crude and bio-synthetic oils. In EU several instruments encouraging bio-fuel and especially ethanol production are approved: The white book "Energy for the future: renewable energy sources" from 1997; The green book "Towards European strategy for energy supplies stability" from 2000; Directive 2003/30/EU for encouragement of bio-fuels and other renewable energy sources utilization in transportation from May 2003 (Berg 2004; Demirbas 2008; Directive 2003/30/EC; Kosaric and Vardar-Sukan 2001; Thomsen et al. 2003; Zaldivar et al. 2001).

Worldwide ethanol is obtained through fermentation of boiled raw cereals (wheat, barley, maize), potato, lignocellulosic materials, bagasse, etc., which are subjected to a fermentation process using yeast of the species *Saccharomyces cerevisiae*. The fermentation process is still a subject to a number of studies regarding the optimization of its parameters - time, temperature, process equipment (Berg 2004; Demirbas 2008; Directive 2003/30/EC; Kosaric and Vardar-Sukan 2001; Thomsen et al. 2003; Zaldivar et al. 2001).

Culturing of microorganisms (alcohol fermentation) occupies an important place in ethanol production. The consumption of sugars by yeast is accompanied by the accumulation of yeast biomass, ethanol and a number of secondary metabolites - esters, aldehydes and higher alcohols. The modelling of the fermentation process, its optimization and intensification are based on the formalization of the process kinetics after a number of assumptions. Usually formalization is done in terms of ethanol accumulation as it is considered to be the only product of yeast biomass metabolism. This formalization is based on one of the two assumptions: microorganisms simultaneously multiply and accumulate product or microorganisms do not multiply, thus acting as a biocatalyst. The second assumption is valid for systems with immobilized cells, in which the localization of cells in the matrix of the carrier limit their growth to a minimum, leading to their action only as a catalyst of the fermentation process (Kostov 2015; Malek and Fencl 1968; Pirt 1975; Willaert et al. 1996; Yarovenko 2002).

The equipment layout, mainly the type of fermentation system used - reactor with ideal mixing, reactor with ideal ejection or a real reactor, is important for the formalization of the alcohol fermentation kinetics. The structure of the flows in the apparatus largely determine the substrate consumption rate and the specific release rate of the target product (Willaert et al. 1996; Levenspiel 1999).

The purpose of the present work was to present an analytical method for the determination of kinetic parameters of a continuous fermentation process in a reactor with ideal mixing. A column bioreactor with immobilized biocatalyst in a fluidized bed, with a known flows structure was selected as a model reactor. In the fulfillment of this purpose two problems were solved – determination of the optimal dilution rate in one fermentation step and determination of certain kinetic parameters in a cascade of reactors with ideal mixing.

MATERIALS AND METHODS

Microorganisms, media and immobilization conditions.

The study was performed with the dry yeast *Saccharomyces cerevisiae* 46EDV supplied by the Company "Martin Vialatte OEnologie", France. Yeasts were stored under refrigeration conditions and were rehydrated according to manufacturer's recommendations prior to the survey.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

For the conduction of the fermentation process was used a medium with pre-optimized qualitative and quantitative composition (g/dm^3): glucose - 118.40; $(NH_4)_2SO_4$ – 2; KH_2PO_4 – 2,72; $MgSO_4x7H_2O$ – 0,5; yeast extract – 1 (Kostov, 2007; Kostov, 2015).

The cells were immobilized in a 4 % calcium alginate gel. After autoclaving the alginate solution for 20 min at 120 °C, the solution was mixed with the cell suspension to obtain a cell concentration of 10^8 cells/mL of gel. This suspension was forced through a syringe needle by means of peristaltic pump and dropped into 2 % (w/v) $CaCl_2$ solution. The resulting beads were approximately 2 mm in diameter. The beads were left for 30 min in calcium solution and then were washed with physiological solution (saline solution) (Kostov 2007).

Bioreactor and culturing conditions

The laboratory bioreactor (Fig. 1) with a fluidized bed "F2" was a glass cylinder (1) with a height of 420 mm, disposed between stainless steel flanges. The apparatus was equipped with a cylindrical phase separator (6) and a liquid degassing cylinder (10) with a height of 100 mm. The column was provided with a cover on which the electrodes for monitoring the fermentation process were placed.

The temperature of the culture medium was measured by RTD (16) and was controlled by the control unit by switching on the heater (13) and periodically passing the cooling water in the heat exchanger (17) through a magnetic valve (15).

The pH during fermentation was kept constant at the optimum value of pH=4.5 using 10% H_2SO_4 and 20% KOH with the help of peristaltic pumps, operated by the control unit. The pH was measured using a combined electrode (12).

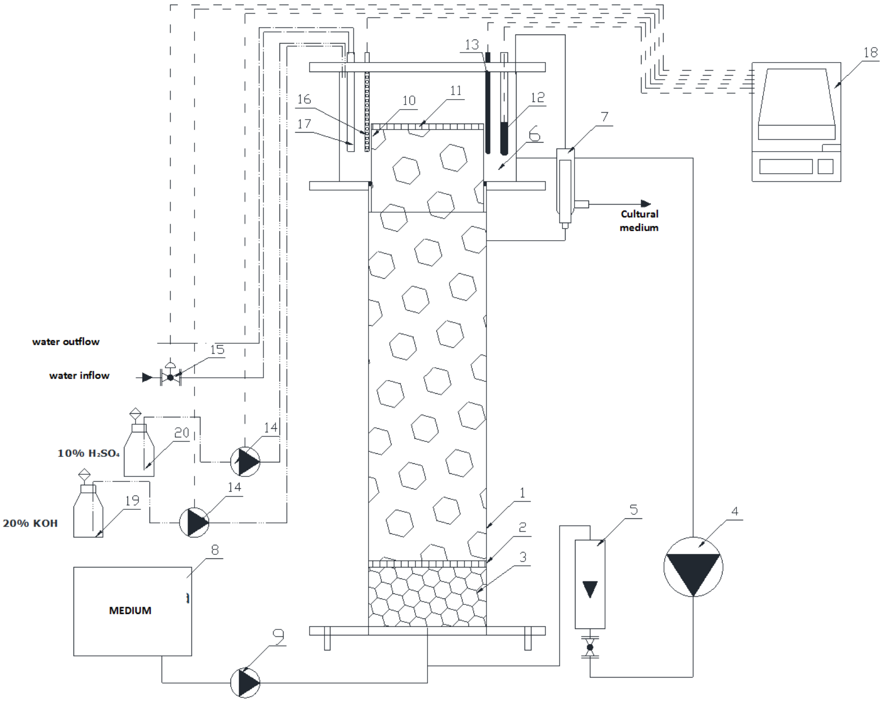


Figure 1: Scheme of a Laboratory Bioreactor With a Fluidized Bed

1 - column; 2 – grid; 3 - drainage; 4 - peristaltic pump; 5 - calibrated rotameter; 6 - phase separator; 7 - outgoing cell; 8 – medium reservoir; 9 - peristaltic pump for continuous operation; 10 – gas separator; 11 – grid; 12 - pH electrode; 13 - heater; 14 - pumps for pH correction; 15 - solenoid valve; 16 - RTD Pt 100; 17 - heat exchanger for submission of cold water; 18 - control device "Applikon"; 19,20 - banks for pH – reagents

300 g of the immobilized preparation and 1,8 dm^3 of the broth medium were placed in the apparatus. The preparation was brought in a fluidised state by means of a pump (4) and batch fermentation at 28 °C for 24 h, which provided ethanol yield of 50% in the reactor, was conducted after the establishment of a steady state of fluidization. The system was then put into a continuous mode, by turning peristaltic pump (9) on. The flow rate of the supplied fluid was determined by the required dilution rate. Constant fluid volume was maintained through spillway (7).

CO_2 formed during fermentation left the apparatus through a heat exchanger placed on the lid of the column in which cold cooling water was circulating. The alcohol vapors escaping from the apparatus were being condensed in the heat exchanger.

Analyses

The concentrations of ethanol and glucose in the medium were determined automatically using an Anton Paar DMA 4500 ", Austria.

The biomass concentration in the immobilized preparation was determined after its dissolution in sodium citrate according to the cell-free spectrophotometrical method at a wavelength of 620 nm (Zhou et. al. 1998).

RESULTS AND DISCUSSION

Structure of the flows in the reactor

The structure of the flows in the bioreactor was defined in a series of studies. From the analysis of the tracer concentration at the output of the system, it can be determined that the bioreactor with immobilized cells in a fluidized bed and recirculation of the culture medium may be referred to the group of reactors with ideal ejection (Fig. 2) (Kostov 2007).

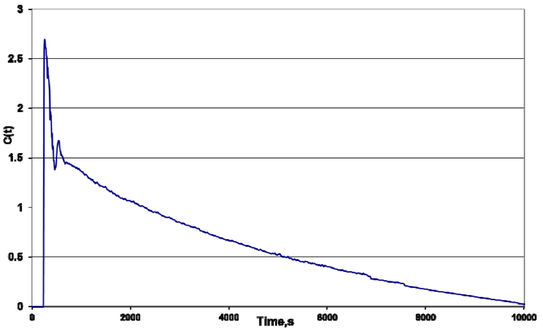


Figure 2: Changes in the Tracer Concentration at the Output of the System

Kinetics of continuous fermentation process in one fermentation step in a reactor with ideal mixing

A series of experiments to determine the optimal dilution rate of the fermentation process (D, h^{-1}) were conducted. The results are presented in Fig. 3 and Fig. 4. The optimum dilution rate was based on the system productivity, which was presented as the multiplication the product quantity at the output of the system and the dilution rate - Q = P.D, g / (dm^3.h). The concentration of the basic fermentation process parameters was determined as the average concentration at the system output at time corresponding to more than 3/D.

Unlike the preparation of biomass in continuous systems where D=μ, the specific rate of product formation was also influenced by the concentration of biomass and product in the medium. Thus, various combinations and relative freedom exist in choosing the appropriate conditions for continuous process in this case. The microorganism concentration itself did not uniquely determine product formation. It was only one of the variables that were crucial to the process, namely representing the total amount of active enzyme systems involved in the reaction. This amount varied depending on the physiological state of the culture, and was generally determined by the age, the culture conditions, etc. (Malek and Fencl 1968).

Another important condition for the development of the continuous fermentation model was the assumption that the process with immobilized cells could be described by the equations for free cells. This assumption was valid at zero and minimal impact of the internal diffusion resistances. Similar negligible influence was proven in the work of Kostov, 2007. Under these conditions, the fermentation process was described by the following system of differential equations:

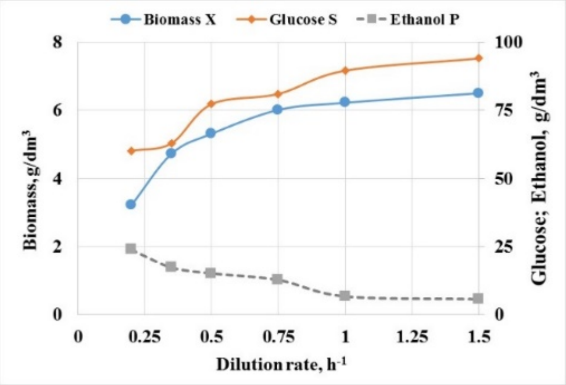


Figure 3: Changes in the Fermentation Process Parameters at Various Dilution Rates

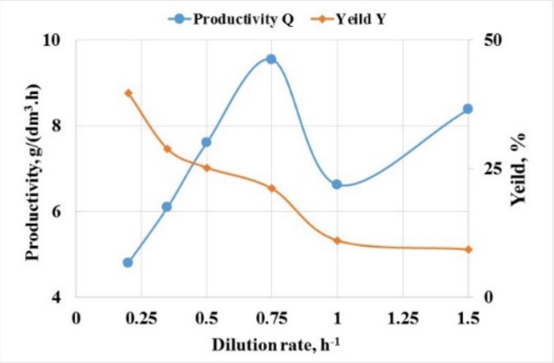


Figure 4: Changes in the Productivity of the System and the Ethanol Yield at Various Dilution Rates

$$\left|\begin{array}{l} \dfrac{dX}{d\tau} = \mu X - DX \\ \dfrac{dP}{d\tau} = q_p X - D.P \\ \dfrac{dS}{d\tau} = D\left(S_0 - S\right) - \dfrac{q_p X}{Y_{P/S}} \end{array}\right. \qquad (1)$$

wherein: X, S, P were the concentrations at the outlet of the system of biomass, substrate and product, g/dm^3; μ - specific growth rate of the biomass, h^{-1}; q_P - specific rate of ethanol accumulation, g/(dm^3.h); D - dilution rate, h^{-1}; $Y_{P/S}$ - product yield per unit substrate.

It is characteristic of the immobilized cell system that the biomass is confined within the matrix of the carrier and does not leave the volume of the apparatus. Assuming that condition, dX/dτ=0, but since all calculations of the kinetic characteristics included biomass growth, it was assumed that its concentration inside the carrier was determined by the first equation in the equation system (1).

The maximum biomass yield and the specific expense of the substrate to maintain the vital activity of the microorganisms were linked by the relationship:

$$q_S = \frac{D}{Y_{X/S}} = \frac{D}{Y^m_{X/S}} + m_S = aD + b \qquad (2)$$

wherein: a=1/Y^m^X/S and b=mS.

In order to determine the values of the parameters in equation (2), an equation to describe the specific growth rate in stationary continuous mode should be defined. In the absence of substrate inhibition it could be assumed that:

$$\mu = \mu_m \frac{S}{K_S + S} - m_S Y^m_{X/S} \qquad (3)$$

This equation is also known as Monod model for hemostat culturing (Pirt, 1975).

The relation between q_S and D could be determined at stationary mode at $dP/d\tau=0$ and $dS/d\tau=0$:

$$q_S = \frac{D}{\dfrac{Q}{q_{p0}(S_0 - S)}} \qquad (4)$$

The function $q_S(D)$ was plotted using equation (4) and the parameters in equation (2) were defined (Fig. 5). The parameter q_{p0} represented the specific rate of product accumulation in batch mode (Kostov, 2007). It had a value of q_{p0} = 2.052 g/(dm^3.h).

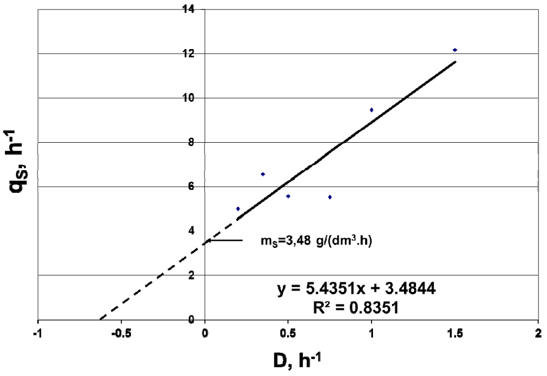


Figure 5: Graphical Definition of the Parameters in Equation (2)

The coefficient m_S=3,48 g/(dm^3.h) was equal to the specific consumption of substrate for maintenance of the vital activity of existing cells at a given time. It characterized the physiological state of the culture. This part of the total amount of consumed substrate was consumed in the synthesis of cellular structures, for maintaining ionic gradients and neutral molecules between the cell and the medium, and between the different structures of the cell (Pirt, 1979). The high value of m_S was most likely due to the large amount of viable cells in the volume of the medium and to the reduced living space in the pores of the carrier.

After the determination of the parameters in equation (2), the values of the constants in equation (3) were determined using the method of least squares. From experimental data and the adoption of stationary regime of the system the kinetic constants were μ_m = 1,49 h^{-1}; K_S=9,113 g/dm^3; $Y^m_{X/S}$=0,184 g/g (R^2 = 91,68%). Data showed a relatively high specific growth rate, due to the high volume of viable biomass in the reactor and the fact that the biomass grew without leaving the apparatus. $K_S << S$, which indicated high affinity of the biomass to the substrate. The value of the constant K_S was much smaller than the values listed in Rovinski and Yarovenko, 1978, where the value was commensurable with the concentration of substrate in the medium.

The other parameters of the system of differential equations (1) could easily be determined from the stationary regime of the system:

$$\widetilde{P} = q_P \frac{\widetilde{X}}{D} \qquad (5)$$

$$\widetilde{S} = S_0 - \frac{q_P \widetilde{X}}{Y^m_{P/S} D} \qquad (6)$$

Equations (5) and (6) give the relationship between the kinetics of the fermentation process and the steady state of the system. The parameters in them can be identified by experimental data of the fermentation system, but only if it is in stationary mode.

By experimental data from Fig. 3 and Fig. 4 using the method of least squares it was found out that q_P = 2,047 g/(g.h), $Y^m_{P/S}$ = 0,298 g/g. The correlation coefficient between experimental data and the models ranged from 74.5% to 84.4%. The model for the description of product accumulation underestimated experimental data at low dilution rates, but fairly accurately described the experimental values at medium and high values of D, i.e. within the operating range of the system. The specific rate of product accumulation was close in value to that of the batch process, which confirmed the accepted assumption that the constants could be calculated by the batch process.

A system of explicit equations for the stationary regime was obtained as a result of the analytical identification conducted:

$$\left|\begin{aligned}
&\widetilde{X} = 0{,}382.D.\left(\frac{118{,}4}{D+3{,}48} - \frac{9{,}113}{1{,}49-D}\right)\\
&\widetilde{P} = 0{,}782.\left(\frac{118{,}4}{D+3{,}48} - \frac{9{,}113}{1{,}49-D}\right)\\
&\widetilde{S} = 118{,}4 - 2{,}322.\left(\frac{118{,}4}{D+3{,}48} - \frac{9{,}113}{1{,}49-D}\right)\\
&\widetilde{Q} = 0{,}782.D\left(\frac{118{,}4}{D+3{,}48} - \frac{9{,}113}{1{,}49-D}\right)
\end{aligned}\right. \qquad (7)$$

The fourth equation in the system (7) represented the functional dependence of system productivity in stationary regime. The results for the fermentation process dynamics showed that the system productivity had its maximum in a relatively narrow area of dilution rates - about 0,7-0,8 h^{-1}. The product concentration, therefore its yield, decreased with increasing the dilution rate, which was related to the lack of sufficient time for stay of the fluid in the fermentation area and therefore insufficient time to conduct the biological transformation. Through numerical optimization of the fourth equation in system (7), the optimal value of the dilution rate could be determined (Table 1).

Table 1: Parameters of the System at the Optimal Dilution Rate

Dm	$\widetilde{X}_m$	$\widetilde{P}$	$\widetilde{S}$	$\widetilde{Q}_m$,
h^{-1}	g/dm^3			g/(dm^3.h)
0,725	4,52	12,76	80,72	9,265

Kinetics of continuous fermentation process in a cascade of reactors with ideal mixing

Data in Table 1 showed very low product yield (about 21%) and poor sugars comsuption. This required that the fermentation was carried out in several consecutive steps (a cascade of reactors with ideal mixing). A simulation study was conducted and the parameters of the fermentation process in successive fermentation steps were defined based on the kinetics in one step under the following assumptions (Table 2): the dilution rate was determined by the dilution rate in the first step; the output parameters of a given apparatus formed the input parameters of the next apparatus; the kinetics in successively connected apparatuses was described by the same equations; the biomass concentration was a constant value for each one of the fermentation steps.

Data in Table 2 showed that the system productivity in the last two apparatuses of the cascade was from 1.5 to 3 times lower than that in the first 3 apparatuses. This can be easily explained, after the determination of the kinetic parameters of the cascade.

The following system of differential equations was used for description of the kinetics of the processes in the cascade of apparatuses (Rovinski and Yarovenko, 1978):

$$\left|\begin{aligned}&\frac{dX_i}{d\tau}=D_i(X_{i-1}-X_i)+\mu_i X_i\\&\frac{dP}{d\tau}=D_i(P_{i-1}-P_i)+q_{pi}\mu_i X_i\\&\frac{dS}{d\tau}=D_i\left(S_{i-1}-S_i\right)-\frac{\mu_i X_i}{Y}-\frac{k_i P_i S_i}{K_{Mi}+S_i}\end{aligned}\right. \qquad (8)$$

wherein: X$_i$, S$_i$, P$_i$ were the concentration at the outlet of the system of biomass, substrate and product in the i-th apparatus, g/dm^3; μ$_i$ -specific growth rate of the biomass in the i-th apparatus, h^{-1}; q$_{pi}$ - specific ethanol production rate in the i-th apparatus, g/(dm^3.h); K$_{Mi}$ - saturation constants of the metabolic products in the i-th apparatus; k$_i$ - kinetic constant/coefficient of the reaction rate in the i-th apparatus;

In the third equation of the differential equation system compared with the system (1), a correction that shows growth inhibition in the cascade of apparatuses after the first one due to ethanol accumulation should be made. This adjustment is made by the member:

$$\frac{k_i P_i S_i}{K_{Mi}+S_i}.$$

The following dependencies were valid in stationary regime of the system (Rovinski and Yarovenko, 1978):

$$\begin{aligned}&X_i=\frac{D_i X_i-1}{D_i-\mu i};\\&P_i=P_{i-1}+q_{pi}(X_i-X_{i-1})\end{aligned} \qquad (9)$$

The Monod equation was used for the description of the specific growth rate:

$$\mu_i=\mu_{mi}\frac{S_i}{K_S+S_i} \qquad (10)$$

Table 2: Fermentation Process Parameters in a System of Consequtively Connected Apparatuses with Ideal Mixing

№ of apparatus	Dm	$\widetilde{X}_m$	$\widetilde{P}$	$\widetilde{S}$	$\widetilde{Q}_m$ *
	h^{-1}	g/dm^3			g/(dm^3.h)
1	0,725	4,52	12.76	80.72	9.27
2			25.61	55.6	9.31
3			36.55	34.2	7.93
4			44.84	18	6.01
5			49.95	8	3.71

* calculated based on the ethanol produced in the step

The system of equations (8) can be solved analytically if taking into account the already described assumptions as well as equations (9) and (10). The methodology for solving the system of equations was presented in Rovinski and Yarovenko, 1978 and the following relationships were obtained as a result of its implementation:

$$\begin{aligned}&y_i=b_0+b_1x_i\\&y_i=\frac{Y(\alpha_T-\alpha'_{i-1})}{\alpha'_{i-1}-\alpha_T(1-Y)};\ \mathrm{x_i}=\frac{1}{i};\\&b_0=\frac{\mu_m K_M}{kq_pK_S};\ b_1=\frac{D(K_S-K_M)}{kq_pK_S}\end{aligned} \qquad (11)$$

wherein: b$_0$ and b$_1$ - coefficients; Y - economic coefficient (biomass yield); α', α$_T$ - real and theoretical ethanol yield (α' - the yield was calculated for each step based on the sugars utilized in the step; α$_T$ = 0,5114 - calculated based on the stoichiometry of the alcohol fermentation process at 20 °C (Yarovenko and Rovinski, 1978)); μ$_m$ - maximum specific growth rate of the biomass, h^{-1}; q$_p$ - specific ethanol production rate in one fermentation step or in a cascade of apparatuses, g/(dm^3.h); K$_S$, K$_M$ - constants of saturation of the substrate and the metabolic products; k - kinetic constant/coefficient of the reaction rate;

The coefficients b$_0$ and b$_1$ could be determined by the fermentation process dynamics. The kinetic parameters in the cascade of reactors with ideal mixing could be defined using the coefficients b$_0$ and b$_1$. To carry out the calculations all kinetic parameters were assumed to be equal to the kinetic parameters calculated in the first step.

To determine the kinetic parameters in the equations (11) it was necessary to determine the value of y_i in each of the steps. Thus, the actual ethanol production, which is the ratio between the amounts of product accumulated in the step and the theoretical amount of ethanol, which would be the result of full utilization of the substrate entering the given apparatus, was calculated. After the determination of the parameter y_i, $y_i = f(x_i)$ was plotted, which was compared with the equation of a straight line (Fig. 6). The calculations after the first apparatus were made at a constant dilution rate according to the methodology of calculating (Rovinski and Yarovenko, 1978). The parameters of the equation of the straight line were the coefficients b_0 and b_1.

The kinetic constants for the cascade of reactors with ideal stirring were calculated based on the graphical determination of the parameters in equation (11): kq_P = 11.67 g/kg and K_M=12.03 g/dm^3. These two constants were the averages for the entire cascade of apparatuses with ideal mixing.

A major disadvantage of this method was that the whole cascade of apparatuses was formalized to one apparatus with ideal mixing.

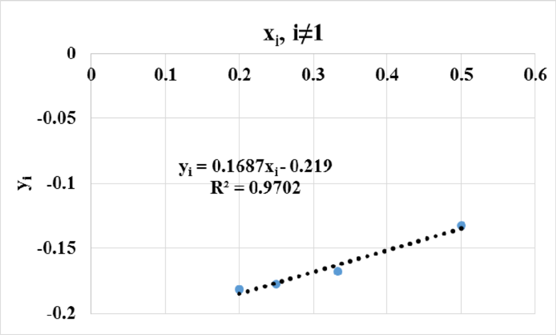


Figure 6: Graphical Definition of the Parameters in Equation (11)

However, the use of the method allowed to determine some important system parameters. Firstly, the minimum number of apparatuses, which was necessary to effect complete transformation of the substrate was 3 (Table 2). Secondly, the estimation of K_M indicated the presence of product inhibition, which was particularly strong in the last two apparatuses of the cascade (apparatus 4 and 5).

Data from Table 2 and the identified kinetic parameters indicated that fermentation could be completed in 3 apparatuses. This could be done by two approaches – by optimization of the dilution rate of each apparatus and by optimization of the fluidisation conditions, that is, by reduction of the impact of the so-called external diffusion resistances in the system. These two optimization problems themselves are interesting and will be the subject of subsequent publications. However, the kinetic parameters as shown in the present publication need to be determined in order these new problems to be solved.

Another important feature of the system in question is the fact that there was a significant amount of ethanol in the capsules, which was supposed to leave them by diffusion. For this purpose it is also necessary to optimize the hydrodynamic environment in the apparatus, but it must not distort the structure of flows in it.

The comparison of the results in Table 2 with data of real experiment is the subject of study in the present work. Initial results showed that in the second stage productivity of about 9 g/dm3 was achieved upon reaching ethanol yield of about 75% of the theoretical yield (total yield in two fermentation steps). This is encouraging, and is due to the fact that the fermentation is carried out in optimized hydrodynamic conditions that are also object of the present study.

The results obtained are comparable with the data presented in Rovinski and Yarovenko, 1978, differing only in the the dilution rate that was several times higher in value. The high value of the dilution rate established was due to the increased concentration of cells in the volume of the apparatus, which was one of the advantages of immobilized cell systems.

CONCLUSION

The present publication presents an analytical approach for determination of part of the kinetic parameters of alcohol fermentation process performed in a single apparatus or a cascade of apparatuses with ideal mixing. The kinetic parameters were defined graphically and analytically based on experimental data on the dynamics of the fermentation process with immobilized cells in a fluidised bed reactor. They were used to determine by simulation the dynamics in the cascade of fermentation apparatuses, as well as for the determination of the minimum number of steps in the cascade, and its kinetic characteristics.

A great advantage of the proposed method is that it eliminates to a great extent the specific hydrodynamic dependencies in the apparatus (especially fluidized bed reactors) and formalizes relationships to systems with ideal mixing. Thus, if it is possible to prove that an apparatus is with ideal mixing, the proposed methodology can be applied to the apparatus and to a cascade of similar apparatuses. An important condition for the application of the methodology is the identification of kinetic parameters in the apparatus to be done in a stationary mode, otherwise kinetics will be linked to the specific type of bioreactor used. It is important to note that the fermentation kinetics will be affected by the type of the selected fermentation - with free or immobilized cells, which would change the specific representation of the differential equation systems (1) and (8).

REFERENCES

Berg C., 2004. “World fuel ethanol - analysis and outlook”, http://www.distill.com/World-Fuel-Ethanol-A&O-2004.html

Demirbas A., 2008. “Biofuels sources, biofuel policy, biofuel economy and global biofuel projections.” *Energy Conversion and Management*, 49, 2106–2116.

Directive 2003/30/EC of the European parliament and of the council of 8 may 2003 on the promotion of the use of biofuels or other renewable fuels for transport.

Kosaric N., Vardar-Sukan F., 2001. “The Biotechnology of Ethanol - Classical and Future Applications”. WILEY-VCH Verlag GmbH, 2001.

Kostov, G. 2007. “Investigation of systems for ethanol fermentation” PhD thesis, Plovdiv, 2007 (in Bulgarian)

Kostov, G. 2015. “Intensification of fermentation processes with immobilized bio catalysis” DSC thesis, Plovdiv, 2015 (in Bulgarian)

Levenspiel, O. 1999. “Chemical reactor engineering.” John Wiley & Sons.

Malek I., Fencl Z. 1968. “Continuous cultifation of microorganisms – theory and methodology.” Pishevaya promishenosti, Moscow. (in Russian).

Pirt, J. 1975. “Priciples of microbe and cell cultivation.” Blackwell Sicentific Publication, Londov.

Rovinskii L., Yarovenko V. 1978. “Modeling and optimization of microbiological processes in ethanol production.” Pishevaya promishenosti, Moscow. (in Russian).

Thomsen A.B., Medina C., Ahring B.K., 2003. “Biotechnology in ethanol production” *Risø Energy Report*, 2.

Willaert, R. G., Baron, G. V., De Backer L. 1996. “Immobilised living cell systems.” John Wiley & Sons, Chichester.

Yarovenko V.L. 2002. “Technology of spirit”, Colos-Press, Moscow (in Russian).

Zaldivar J.,· Nielsen J., Olsson L. 2001. “Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration”, *Appl Microbiol Biotechnol.*, 56, 17–34.

Zhou, Y., Martins E., Groboillot A., Champagne C.P., Neufeld R.J. 1998. “Spectrophotometric quantification of lactic bacteria in alginate and control of cell release with chitosan coating.” *Journal of Applied Microbiology*, 84, 342-348.

AUTHOR BIOGRAPHIES

GEORGI KOSTOV is associated professor at the department “Technology of wine and brewing” at University of Food Technologies, Plovdiv. He received his MSc in “Mechanical engineering” in 2007, PhD on “Mechanical engineering in food and flavor industry (Technological equipment in biotechnology industry)” in 2007 from University of Food Technologies, Plovdiv and DSc on “Intensification of fermentation processes with immobilized biocatalysts”. His research interests are in the area of bioreactors construction, biotechnology, microbial population’s investigation and modeling, hydrodynamics and mass transfer problems, fermentation kinetics, beer production.

ZAPRYANA DENKOVA is professor at the department “Microbiology” at University of Food Technologies, Plovdiv. She received her MSc in “Technology of microbial products” in 1982, PhD in „Technology of biologically active substances“ in 1994 and DSc on “Production and application of probiotics” in 2006. Her research interests are in the area of selection of probiotic strains and development of starters for food production, genetics of microorganisms, and development of functional foods.

VESELA SHOPSKA is assistant professor at the department “Technology of wine and brewing” at University of Food Technologies, Plovdiv. She received her MSc in “Technology of wine and brewing” in 2006 at University of Food Technologies, Plovdiv. She received her PhD in “Technology of alcoholic and non-alcoholic beverages (Brewing technology)” in 2014. Her research interests are in the area of beer fermentation with free and immobilized cells; yeast and bacteria metabolism and fermentation activity.

ROSITSA DENKOVA is assistant professor at the department “Biochemistry and molecular biology” at University of Food Technologies, Plovdiv. She received her MSc in “Industrial biotechnologies” in 2011 and PhD in “Biotechnology (Technology of biologically active substances)” in 2014. Her research interests are in the area of isolation, identification and selection of probiotic strains and development of starters for functional foods.

BOGDAN GORANOV is researcher in company “LBLact”, Plovdiv. He received his PhD in 2015 from University of Food Technologies, Plovdiv. The theme of his thesis was “Production of lactic acid with free and immobilized lactic acid bacteria and its application in food industry”. His research interests are in the area of bioreactors construction, biotechnology, microbial population’s investigation and modeling, hydrodynamics and mass transfer problems, fermentation kinetics.

VASIL ILIEV is a service manager in “Weissbiotech”, Ascheberg, Germany. He received his PhD in 2016 from University of Food Technologies, Plovdiv. His research interests are in the area of bioreactors construction, biotechnology, microbial population’s investigation and modeling, hydrodynamics and mass transfer problems, fermentation kinetics, beer and ethanol production.

IVAN PETELKOV is a PhD student at the Department of Wine and Beer Technology at the University of Food Technologies, Plovdiv. He received his MSc in Technology of wine and brewing at the University of Food Technologies, Plovdiv.

PREDICTIVE CONTROL OF TWO-INPUT TWO-OUTPUT SYSTEM WITH NON-MINIMUM PHASE

Marek Kubalcik, Vladimír Bobál
Tomas Bata University in Zlín
Faculty of Applied Informatics
Nad Stráněmi 4511, 760 05, Zlín, Czech Republic
E-mail: kubalcik@fai.utb.cz, bobal@fai.utb.cz

Tomáš Barot
Department of Mathematics with Didactics
University of Ostrava, Pedagogical Faculty
Mlýnská 5, 701 03 Ostrava, Czech Republic
E-mail: Tomas.Barot@osu.cz

KEYWORDS
Simulation, Model Predictive Control, Linear Systems, Non-Minimum Phase Systems, TITO Systems.

ABSTRACT

In this paper, a simulation of predictive control of a two-input two-output (TITO) system with non-minimum phase is presented. The proposed controller is based on extended setting of constraints. This setting represents a modification for purposes of control of non-minimum phase multivariable systems. The main problem of control of this particular type of system is undesirable undershoot in the initial phase of the control. Known methods can properly reduce the undershoot in case of predictive control of single-input single-output (SISO) systems. The paper proposes a modification of a predictive controller for two-input two-output systems with non-minimum phase behaviour. The non-minimum phase TITO models are simulated using its mathematical representation in the form of transfer function matrix.

INTRODUCTION

Control of systems which are characterized by non-minimum-phase behavior (Hoagg and Bernstein, 2007) requires a quite sophisticated approach for a controller design. The main problem is an undershoot which is present during the control using classical controllers without any modifications taking into account the non-minimum phase behaviour of the controlled system. A suitable method for control of the non-minimum-phase systems is model predictive control (MPC) (Huang, 2002; Rawlings and Mayne, 2009; Corriou, 2004). A successful implementation of MPC for control of non-minimum phase SISO systems is presented for example in (Barot and Kubalcik, 2014). Generally, many technological processes require a simultaneous control of several variables related to one system. In this case, a design of a controller is more sophisticated. One of the most effective approaches to control of multivariable systems is model predictive control. An advantage of model predictive control is that multivariable systems can be handled in a straightforward manner. Therefore, model predictive control appears as a suitable approach for control of non-minimum phase multivariable systems.

Model Predictive Control is one of the control methods which have developed considerably over a few past years. Predictive control is essentially based on discrete or sampled models of processes. Computation of appropriate control algorithms is then realized especially in the discrete domain. The basic idea of the generalized predictive control (Camacho and Bordons, 2007; Kwon, 2005) is to use a model of a controlled process to predict a number of future outputs of the process. A trajectory of future manipulated variables is given by solving an optimization problem incorporating a suitable cost function and constraints. Only the first element of the obtained control sequence is applied. The whole procedure is repeated in the following sampling period. This principle is known as the receding horizon strategy.

The known approach for control of non-minimum-phase SISO systems (Camacho and Bordons, 2007) consists of increasing of a minimum control horizon. However, this classical modification may not be generally successful. The further possibility is setting of equality constraints applied in several initial sampling periods of control (Barot and Kubalcik, 2014). This modification was also tested for control of non-minimum phase TITO systems. The obtained results were not satisfactory. The undershoots were not eliminated using this approach.

In this paper, an appropriate setting of constraints in MPC which eliminates undershoots in the initial phase of model predictive control of non-minimum phase multivariable systems is presented. This approach was implemented for control of a TITO non-minimum phase system.

MODEL OF THE CONTROLLED TITO SYSTEM

A continuous TITO system can be expressed using the transfer function matrix:

$$\boldsymbol{G}(s)=\begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \tag{1}$$

A non-minimum phase behaviour is characterized by positive values of roots $\vartheta_i; i \in \langle 1;4 \rangle$ in numerators in partial transfer functions $G_{ij}; i, j \in \langle 1;2 \rangle$ of the transfer function matrix (2). $\pi_l; l \in \langle 1;8 \rangle$ are poles.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

$$\boldsymbol{G}(s)=\begin{bmatrix}\dfrac{s-\vartheta_1}{(s-\pi_1)(s-\pi_2)} & \dfrac{s-\vartheta_2}{(s-\pi_3)(s-\pi_4)}\\ \dfrac{s-\vartheta_3}{(s-\pi_5)(s-\pi_6)} & \dfrac{s-\vartheta_4}{(s-\pi_7)(s-\pi_8)}\end{bmatrix} \qquad (2)$$

In the discrete simulation of MPC, the model (2) is expressed in Z-transform (Kučera, 1991) for a given sampling period T as (3).

$$\boldsymbol{G}(z^{-1})=\begin{bmatrix}G_{11}(z^{-1}) & G_{12}(z^{-1})\\ G_{21}(z^{-1}) & G_{22}(z^{-1})\end{bmatrix} \qquad (3)$$

For the simulation purposes, the mathematical model of TITO system (3) can have a form of the matrix fraction (4).

$$\boldsymbol{G}(z^{-1})=\boldsymbol{A}^{-1}(z^{-1})\boldsymbol{B}(z^{-1}) \qquad (4)$$

The structure of the particular matrices $\boldsymbol{A}(z^{-1})$ and $\boldsymbol{B}(z^{-1})$ can be seen in (5)-(6) with description of polynomials in (7)-(8).

$$\boldsymbol{A}(z^{-1})=\begin{bmatrix}\alpha_{11}(z^{-1}) & \alpha_{12}(z^{-1})\\ \alpha_{21}(z^{-1}) & \alpha_{22}(z^{-1})\end{bmatrix} \qquad (5)$$

$$\boldsymbol{B}(z^{-1})=\begin{bmatrix}\beta_{11}(z^{-1}) & \beta_{12}(z^{-1})\\ \beta_{21}(z^{-1}) & \beta_{22}(z^{-1})\end{bmatrix} \qquad (6)$$

$$\left.\begin{aligned}\alpha_{11}(z^{-1})&=1+\alpha_{111}z^{-1}+\alpha_{112}z^{-2};\\ \alpha_{12}(z^{-1})&=\alpha_{121}z^{-1}+\alpha_{122}z^{-2};\\ \alpha_{21}(z^{-1})&=\alpha_{211}z^{-1}+\alpha_{212}z^{-2};\\ \alpha_{22}(z^{-1})&=1+\alpha_{221}z^{-1}+\alpha_{222}z^{-2}\end{aligned}\right\} \qquad (7)$$

$$\left.\begin{aligned}\beta_{11}(z^{-1})&=\beta_{111}z^{-1}+\beta_{112}z^{-2};\\ \beta_{12}(z^{-1})&=\beta_{121}z^{-1}+\beta_{122}z^{-2};\\ \beta_{21}(z^{-1})&=\beta_{211}z^{-1}+\beta_{212}z^{-2};\\ \beta_{22}(z^{-1})&=\beta_{221}z^{-1}+\beta_{222}z^{-2}\end{aligned}\right\} \qquad (8)$$

The transformation from the continuous model (1) to its discrete version (4) is possible using the least squares method (Nelles, 2001). The parameters of the polynomials (7)-(8) were included in an ARX model (Nelles, 2001) of the TITO system.

MODEL PREDICTIVE CONTROL OF TITO SYSTEMS

The model predictive control is a control strategy which incorporates a model of the controlled system for predictions of output variables. The calculations are performed on the receding horizon window which corresponds to a maximum prediction horizon N_2. The control action signal is denoted as $\boldsymbol{u}(k)$. The output signal is $\boldsymbol{y}(k)$, $\boldsymbol{e}(k)$ is a control error and $\boldsymbol{w}(k)$ is a reference signal. Each variable in TITO MPC is two-dimensional.

The vector of future increments of manipulated variable $\Delta\boldsymbol{u}$ with N_u elements is determined by solving an optimization task which comprises a suitable cost function and constraints of variables. N_u is a control horizon. In the optimization task, the unknown variable $\boldsymbol{y}$ is determined by prediction equations (11) where the future outputs of the controlled model are determined by CARIMA model (Controlled AutoRegressive Integrated Moving Average) (Rossiter, 2003) (9). Equation (9) can be rewritten to equations (10) and (11) without consideration of the noise signal $\boldsymbol{e}_\mathrm{s}(k)$. N_1 and N_2 are minimum and maximum prediction horizons. Matrices $\boldsymbol{P}$ and $\mathcal{G}$ are defined in (12)-(14) where $\boldsymbol{Z}$ is a zero matrix of a given dimension.

$$\left.\begin{aligned}\boldsymbol{A}(z^{-1})\boldsymbol{y}(k)&=\boldsymbol{B}(z^{-1})\boldsymbol{u}(k)+\Delta^{-1}(z^{-1})\boldsymbol{C}(z^{-1})\boldsymbol{e}_s(k)\\ \Delta(z^{-1})&=\begin{bmatrix}1-z^{-1} & 0\\ 0 & 1-z^{-1}\end{bmatrix}\end{aligned}\right\} \qquad (9)$$

$$\left.\begin{aligned}\boldsymbol{y}(k)&=\boldsymbol{A}_1\boldsymbol{y}(k-1)+\boldsymbol{A}_2\boldsymbol{y}(k-2)+\boldsymbol{A}_3\boldsymbol{y}(k-3)+\\ &+\boldsymbol{B}_1\Delta\boldsymbol{u}(k-1)+\boldsymbol{B}_2\Delta\boldsymbol{u}(k-2);\\ \boldsymbol{A}_1&=\begin{bmatrix}(1-\alpha_{111}) & -\alpha_{121}\\ -\alpha_{211} & (1-\alpha_{221})\end{bmatrix};\\ \boldsymbol{A}_2&=\begin{bmatrix}(\alpha_{111}-\alpha_{112}) & (\alpha_{121}-\alpha_{122})\\ (\alpha_{211}-a_{212}) & (\alpha_{221}-\alpha_{222})\end{bmatrix};\\ \boldsymbol{A}_3&=\begin{bmatrix}\alpha_{112} & \alpha_{122}\\ \alpha_{212} & \alpha_{222}\end{bmatrix};\ \boldsymbol{B}_1=\begin{bmatrix}\beta_{111} & \beta_{121}\\ \beta_{211} & \beta_{221}\end{bmatrix};\\ \boldsymbol{B}_2&=\begin{bmatrix}\beta_{112} & \beta_{122}\\ \beta_{212} & \beta_{222}\end{bmatrix}\end{aligned}\right\} \qquad (10)$$

$$\underbrace{\begin{bmatrix}\boldsymbol{y}(k+N_1)\\ \vdots\\ \boldsymbol{y}(k+N_2)\end{bmatrix}}_{\boldsymbol{y}}=\boldsymbol{P}\begin{bmatrix}\boldsymbol{y}(k)\\ \boldsymbol{y}(k-1)\\ \boldsymbol{y}(k-2)\\ \Delta\boldsymbol{u}(k-1)\end{bmatrix}+\mathcal{G}\underbrace{\begin{bmatrix}\Delta\boldsymbol{u}(k)\\ \Delta\boldsymbol{u}(k+1)\\ \vdots\\ \Delta\boldsymbol{u}(k+N_u-1)\end{bmatrix}}_{\Delta\boldsymbol{u}} \qquad (11)$$

$$\left.\begin{aligned}\boldsymbol{P}&=\begin{bmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \cdots & \boldsymbol{P}_{14}\\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \cdots & \boldsymbol{P}_{24}\\ \vdots & & \ddots & \vdots\\ \boldsymbol{P}_{i1} & \boldsymbol{P}_{i2} & \cdots & \boldsymbol{P}_{i4}\end{bmatrix};\\ \mathcal{G}&=\begin{bmatrix}\mathcal{G}_{11} & \mathcal{G}_{12} & \cdots & \mathcal{G}_{1j}\\ \mathcal{G}_{21} & \mathcal{G}_{22} & \cdots & \mathcal{G}_{2j}\\ \vdots & & \ddots & \vdots\\ \mathcal{G}_{i1} & \mathcal{G}_{i2} & \cdots & \mathcal{G}_{ij}\end{bmatrix}\end{aligned}\right\} \qquad (12)$$

$$\left.\begin{array}{c} \mathcal{G} \in \mathcal{R}^{2N_2\text{-}2N_1+2,2N_2}; \\ \left.\begin{array}{c} \mathcal{G} = \boldsymbol{Z}; \boldsymbol{Z} \in \mathcal{R}^{2N_2\text{-}2N_1+2,2N_2}; \\ \mathcal{G}_{11} = \mathcal{G}_{22} = \mathcal{G}_{33} = \boldsymbol{B}_1; \\ \mathcal{G}_{21} = \mathcal{G}_{32} = (\boldsymbol{A}_1\boldsymbol{B}_1 + \boldsymbol{B}_2); \\ \mathcal{G}_{31} = (\boldsymbol{A}_1^2\boldsymbol{B}_1 + \boldsymbol{A}_1\boldsymbol{B}_2 + \boldsymbol{A}_2\boldsymbol{B}_1); \\ \left(\begin{array}{c} \mathcal{G}_{i1} = \boldsymbol{A}_{\boldsymbol{1}}\mathcal{G}_{(i-1)1} + \boldsymbol{A}_2\mathcal{G}_{(i-2)1} + \boldsymbol{A}_3\mathcal{G}_{(i-3)1} \\ \mathcal{G}_{i(j-1)} = \boldsymbol{A}_{\boldsymbol{1}}\mathcal{G}_{(i-1)(j-1)} + \boldsymbol{A}_2\mathcal{G}_{(i-2)(j-1)} + \\ + \boldsymbol{A}_3\mathcal{G}_{(i-3)(j-1)} + \boldsymbol{B}_{\boldsymbol{2}} \\ \mathcal{G}_{ij} = \boldsymbol{B}_{\boldsymbol{1}} \end{array}\right), \\ i = 4,\ldots,N_2;\ j = 1,\ldots,i \end{array}\right\} \end{array}\right\} \tag{13}$$

$$\left.\begin{array}{c} \boldsymbol{P} \in \mathcal{R}^{2N_2,8}; \\ \boldsymbol{P}_{11} = \boldsymbol{A}_1; \boldsymbol{P}_{12} = \boldsymbol{A}_2; \boldsymbol{P}_{13} = \boldsymbol{A}_3; \boldsymbol{P}_{14} = \boldsymbol{B}_2; \\ \boldsymbol{P}_{21} = \boldsymbol{A}_1^2 + \boldsymbol{A}_2; \boldsymbol{P}_{22} = \boldsymbol{A}_1\boldsymbol{A}_2 + \boldsymbol{A}_3; \\ \boldsymbol{P}_{23} = \boldsymbol{A}_1\boldsymbol{A}_3; \boldsymbol{P}_{24} = \boldsymbol{A}_1\boldsymbol{B}_2; \\ \boldsymbol{P}_{31} = \boldsymbol{A}_1^3 + \boldsymbol{A}_1\boldsymbol{A}_2 + \boldsymbol{A}_3 + \boldsymbol{A}_1\boldsymbol{A}_2; \\ \boldsymbol{P}_{32} = \boldsymbol{A}_1^2\boldsymbol{A}_2 + \boldsymbol{A}_1\boldsymbol{A}_3 + \boldsymbol{A}_2^2; \\ \boldsymbol{P}_{33} = \boldsymbol{A}_1^2\boldsymbol{A}_3 + \boldsymbol{A}_3\boldsymbol{A}_2; \boldsymbol{P}_{34} = \boldsymbol{A}_1^2\boldsymbol{B}_2 + \boldsymbol{A}_2\boldsymbol{B}_2; \\ \left(\boldsymbol{P}_{ij} = \boldsymbol{A}_{\boldsymbol{1}}\boldsymbol{P}_{(i-1)j} + \boldsymbol{A}_2\boldsymbol{P}_{(i-2)j} + \boldsymbol{A}_3\boldsymbol{P}_{(i-3)j}\right), \\ i = 4,\ldots,N_2;\ j = 1,\ldots,i \end{array}\right\} \tag{14}$$

The optimization problem is then solved as a minimization by the quadratic programming. A cost function J is defined by (15)-(16) where the vector $\Delta\boldsymbol{u}$ is solved with regard to the constraints defined by matrix inequality (17). Matrix $\boldsymbol{I}$ is an identity matrix.

$$J = (\boldsymbol{y} - \boldsymbol{w})^T(\boldsymbol{y} - \boldsymbol{w}) + \Delta\boldsymbol{u}^T\Delta\boldsymbol{u} \tag{15}$$

$$\left.\begin{array}{c} J = \frac{1}{2}\Delta\boldsymbol{u}^T\boldsymbol{H}\Delta\boldsymbol{u} + \boldsymbol{b}^T\Delta\boldsymbol{u}; \\ \boldsymbol{H} \in \mathcal{R}^{2N_u,2N_u}; \\ \boldsymbol{b} \in \mathcal{R}^{2N_u,1}; \\ \boldsymbol{H} = \mathcal{G}^T\mathcal{G} + \boldsymbol{I}; \\ \boldsymbol{b} = \mathcal{G}^{\boldsymbol{T}}\left(\boldsymbol{P}\begin{bmatrix} \boldsymbol{y}(k) \\ \boldsymbol{y}(k-1) \\ \boldsymbol{y}(k-2) \\ \Delta\boldsymbol{u}(k-1) \end{bmatrix} - \boldsymbol{w}\right) \end{array}\right\} \tag{16}$$

$$\boldsymbol{M}\Delta\boldsymbol{u} \leq \boldsymbol{\gamma} \tag{17}$$

CLASSICAL MODIFICATIONS FOR MPC OF NON-MINIMUM PHASE SISO SYSTEMS

The method of predictive control enables a particular setting of the horizons for the purposes of control of non-minimum phase systems. The recommended approach increases the minimum horizon N_1. It means reducing of N_1 upper rows in matrices $\boldsymbol{P}$ and $\mathcal{G}$ for SISO control or $2N_1$ rows for TITO control. Simulations performed for SISO systems proved that this method can not be generally successfully applied with appropriate results. The undesired undershoots were not eliminated in all cases.

Therefore, further possible modification was proposed in (Barot and Kubalcik, 2014) which successfully eliminated undershoots for SISO systems. The principal of this method consists of particular equality constraints settings (18) of $\Delta u_{\max}$ in the initial part of control. E.g., the constraint value was obtained experimentally as a relatively small constant ζ. For TITO systems, the experiments were not successful using the approach which was successfully used for SISO systems.

$$\boldsymbol{M} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}; \boldsymbol{M} \in \mathcal{R}^{N_u,N_u} \tag{18}$$

$$\boldsymbol{\gamma} = \begin{bmatrix} \Delta u_{\max} \equiv \zeta \\ \vdots \\ \Delta u_{\max} \equiv \zeta \end{bmatrix}; \boldsymbol{\gamma} \in \mathcal{R}^{N_u,1} \tag{19}$$

MODIFICATION FOR MPC OF NON-MINIMUM PHASE TITO SYSTEMS

TITO MPC of non-minimum phase systems is more complicated than SISO MPC. The efforts to eliminate the undershoot using the methods described in the previous section were not successful. In this section it is introduced a modification which provided satisfactory control results without undesired undershoot also for TITO non-minimum phase systems.

The main principle consists of restriction of the controlled variable $\boldsymbol{y}$. In this case, restricted is the lower limit of $\boldsymbol{y}$ in the initial part of control. This restriction is applied only in the initial part of control. The matrix inequality (17) is then modified to matrices (20) and (21).

$$\boldsymbol{M} = -\mathcal{G}; \boldsymbol{M} \in \mathcal{R}^{2(N_2-N_1+1),2N_u} \tag{20}$$

$$\boldsymbol{\gamma} = \left[-\boldsymbol{P}\begin{bmatrix} \boldsymbol{y}(k) \\ \boldsymbol{y}(k-1) \\ \boldsymbol{y}(k-2) \\ \Delta\boldsymbol{u}(k-1) \end{bmatrix}\right]; \boldsymbol{\gamma} \in \mathcal{R}^{2(N_2-N_1+1),1} \tag{21}$$

SIMULATION RESULTS

As a simulation example it was chosen a TITO non-minimum phase system given by equations (22)-(24). The continuous-time model of this system has only positive roots in the numerators of the transfer functions (22). The sampling period was chosen as 0.5 [s]. The step responses of the system are in Fig. 1 – 4.

$$\boldsymbol{G}(s)=\begin{bmatrix}\dfrac{-1.5s+1.5}{s^2+2s+1} & \dfrac{-s+1}{s^2+2s+2}\\ \dfrac{-0.5s+0.5}{s^2+2s+2} & \dfrac{-1.2s+1}{s^2+2s+1}\end{bmatrix} \tag{22}$$

$$\left.\begin{aligned}\alpha_{11}(z^{-1})&=1-1.2722z^{-1}+0.4151z^{-2};\\ \alpha_{12}(z^{-1})&=0.1879z^{-1}-0.095z^{-2};\\ \alpha_{21}(z^{-1})&=0.0662z^{-1}-0.0331z^{-2};\\ \alpha_{22}(z^{-1})&=1-1.2708z^{-1}+0.4119z^{-2}\end{aligned}\right\} \tag{23}$$

$$\left.\begin{aligned}\beta_{11}(z^{-1})&=-0.321z^{-1}+0.551z^{-2};\\ \beta_{12}(z^{-1})&=-0.2019z^{-1}+0.3455z^{-2};\\ \beta_{21}(z^{-1})&=-0.101z^{-1}+0.1764z^{-2};\\ \beta_{22}(z^{-1})&=-0.2741z^{-1}+0.4304z^{-2}\end{aligned}\right\} \tag{24}$$

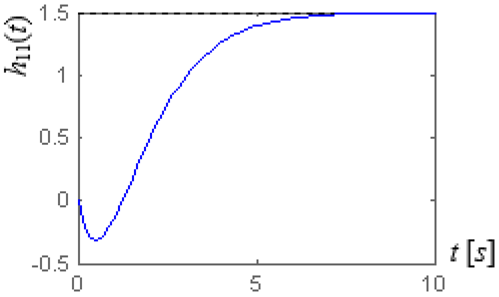


Figure 1: Step Functions of $\boldsymbol{G}_{11}$(s) of TITO Model (22)

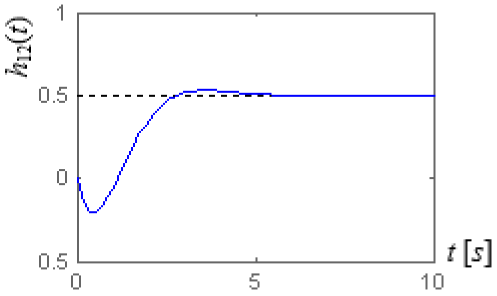


Figure 2: Step Functions of $\boldsymbol{G}_{12}$(s) of TITO Model (22)

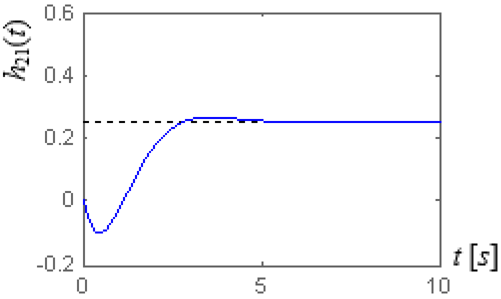


Figure 3: Step Functions of $\boldsymbol{G}_{21}$(s) of TITO Model (22)

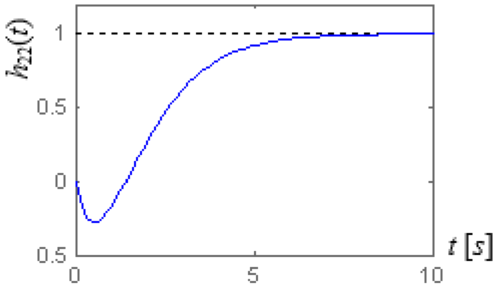


Figure 4: Step Functions of $\boldsymbol{G}_{22}$(s) of TITO Model (22)

MPC of the non-minimum phase TITO system was implemented using MATLAB scripts. The optimization part was programmed using the Hildreth's dual method (Wang, 2009).

The previously introduced methods applied for control of SISO systems were not successful. In this case of the multivariable system undershoots were not eliminated, as can be seen in Fig. 5. The approach proposed in this paper provided satisfactory results, as can be seen in Fig. 6. The minimum, control and prediction horizons were chosen as N_1=1, N_u=30 and N_2=40. The constraints of variables were set according to (21) in several initial steps which cover the undershoot. In this particular case it was 10 steps.

The proposed method was applied for control of the introduced TITO systems with an elimination of significant undershoots in the initial part of the predictive control. The modification was active only during start-up of the control process. It is not appropriate for other step changes which also causes some minor undershoots.

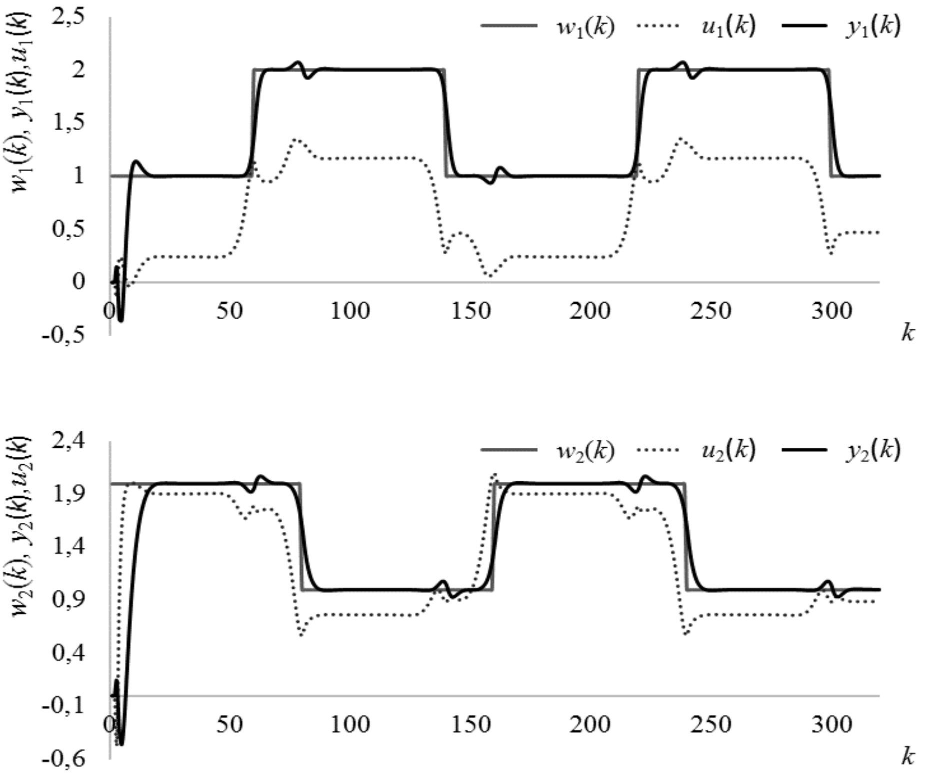


Figure 5: Simulation of Control without Proposed Modification

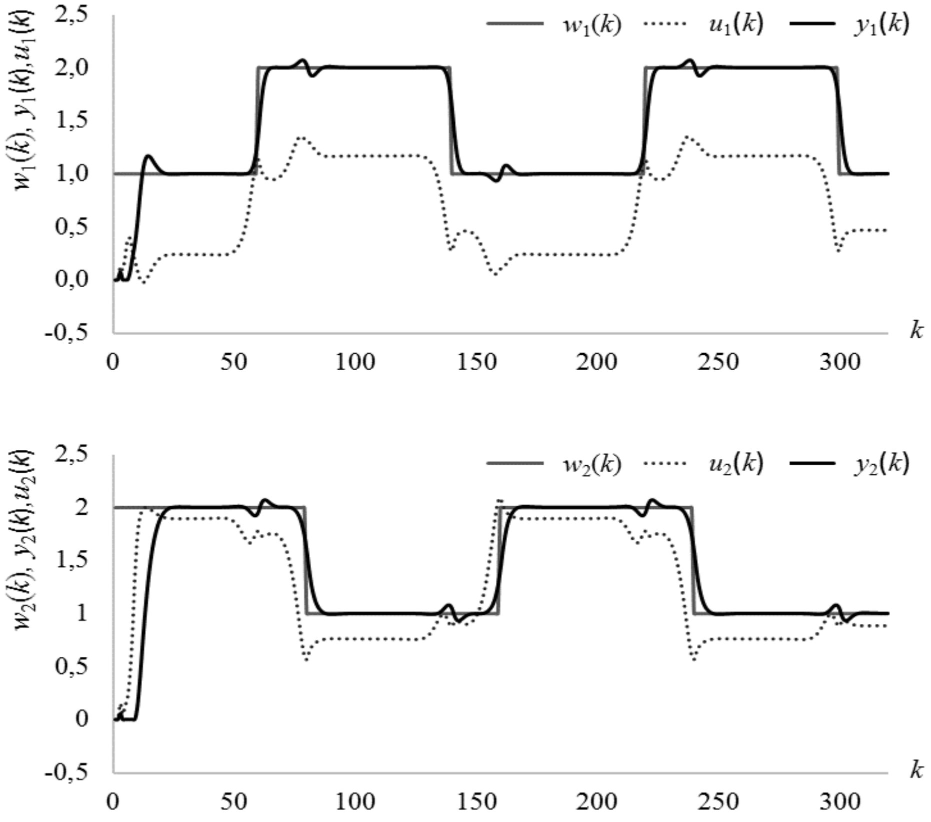


Figure 6: Simulation of Control with Proposed Modification

CONCLUSIONS

The undershoot during the control of the non-minimum phase TITO system was eliminated using the proposed particular setting of constraints in the predictive control. This modification enables successful control of systems with this specific behaviour. The presented restriction is applied in the initial part of control. The approaches suitable for SISO systems were not satisfactory in case of the TITO system. The principal of the proposed modification is based on inequality constraints settings in the optimization task. The modification is used in several initial sampling periods in MPC and the undershoots are successfully eliminated.

REFERENCES

Barot, T., and Kubalčík, M. 2014. Predictive Control of Non-Minimum Phase Systems. In *15th International Carpathian Control Conference* (ICCC). Velke Karlovice, Czech Republic.

Camacho, E.F., and Bordons, C. 2007. *Model predictive control.* London: Springer.

Corriou, J.P. 2004. *Process control: theory and applications*. London: Springer.

Hoagg, J.B., and Bernstein, D.S. 2007. Nonminimum-phase zeros - much to do about nothing - classical control - revisited part II. *Control Systems*, IEEE, 27 (3), 45-57.

Huang, S. 2002. *Applied predictive control*. London: Springer.

Kučera, V. 1991. *Analysis and Design of Discrete Linear Control Systems*. Prague: Nakladatelství Československé akademie věd.

Kwon, W.H. 2005. *Receding horizon control: model predictive control for state models*. London: Springer.

Nelles, O. 2001. *Nonlinear System Identification*, Springer-Verlag, Berlin.

Rawlings, J.B., and Mayne, D.Q. 2009. *Model Predictive Control Theory and Design.* Nob Hill Pub.

Rossiter J. A. 2003. *Model Based Predictive Control: a Practical Approach,* CRC Press.

Wang, L. 2009. *Model Predictive Control System Design and Implementation Using MATLAB*. London: Springer-Verlag Limited.

AUTHOR BIOGRAPHIES

MAREK KUBALČÍK graduated in 1993 from the Brno University of Technology in Automation and Process Control. He received his Ph.D. degree in Technical Cybernetics at Brno University of Technology in 2000. From 1993 to 2007 he worked as senior lecturer at the Faculty of Technology, Brno University of Technology. From 2007 he has been working as an associate professor at the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín, Czech Republic. Current work cover following areas: control of multivariable systems, self-tuning controllers, predictive control. His e-mail address is: kubalcik@fai.utb.cz.

VLADIMÍR BOBÁL graduated in 1966 from the Brno University of Technology, Czech Republic. He received his Ph.D. degree in Technical Cybernetics at Institute of Technical Cybernetics, Slovak Academy of Sciences, Bratislava, Slovak Republic. He is now Professor at the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín, Czech Republic. His research interests are adaptive control and predictive control, system identification and CAD for automatic control systems. You can contact him on email address bobal@fai.utb.cz.

TOMÁŠ BAROT graduated in Information Technology of study program Engineering Informatics at the Faculty of Applied Informatics of the Tomas Bata University in Zlín, Czech Republic in 2010. He received his Ph.D. degree in Automatic Control and Informatics at the same faculty in 2016. From 2016, he has been working at the Department of Mathematics with Didactics of the Pedagogical Faculty at the University of Ostrava in Czech Republic. In his research, he interests in applied mathematics (optimization and numerical methods in control theory) and pedagogy (pedagogical cybernetics and quantitative methods in statistics). His e-mail address is: Tomas.Barot@osu.cz.

VERIFICATION OF ROBUST PROPERTIES OF DIGITAL CONTROL CLOSED-LOOP SYSTEMS

Vladimír Bobál, Ľuboš Spaček and Peter Hornák
Tomas Bata University in Zlín
Faculty of Applied Informatics
Nad Stráněmi 4511
760 05 Zlín
Czech Republic
E-mail: bobal@fai.utb.cz

KEYWORDS

Digital Control, Polynomial Methods, Robustness, Robustness Margins, LQ Method, Simulation of Control Loop Systems.

ABSTRACT

Robustness is specific property of closed-loop systems when the designed controller guarantees control not only for one nominal controlled system but also for all predefined class of systems (perturbed models). The robust theory is mainly exploited for design of the continuous-time systems. This paper deals with an experimental simulation investigation of robust properties of digital control closed-loop systems. Minimization of the Linear Quadratic (LQ) criterion was used for the design of control algorithm. Polynomial approach is based on the structure of the controller with two degrees of freedom (2DOF). Four types of process models (stable, non-minimum phase, unstable and integrating) were used for controller design. The Nyquist plot based characteristics of the open-loop transfer function (gain margin, phase margin and modulus margin) served as robustness indicators. The influence of change of process gain was chosen as a parametric uncertainty. The experimental results demonstrated that a robustness of examined digital control closed-loop systems could be improved by addition of user-defined poles (UDP).

INTRODUCTION

One of possible approaches to digital control systems is the polynomial theory. Polynomial methods are design techniques for complex systems (including multivariable), signals and processes encountered in control, communications and computing that are based on manipulations and equations with polynomials, polynomial matrices and similar objects. Systems are described by input-output relations in fractional form and processed using algebraic methodology and tools (Šebek and Hromčík 2007). The design procedure is thus reduced to algebraic polynomial equations. Controller design consists of solving polynomial (Diophantine) equations. The Diophantine equations can be solved using the uncertain coefficient method – which is based on comparing coefficients of the same power. This is transformed into a system of linear algebraic equations (Kučera 1997).

It is obvious that the majority processes met in industrial practice are influenced by uncertainties. The uncertainties suppression can be solved by implementation of either adaptive control or robust control. The robust control and the adaptive control are viewed as two control techniques, which are used for controller design in the presence of process model uncertainty - process model variations (Landau 1999; Landau et al. 2011).

The design of a robust controller deals in general with designing the controller in the presence of process uncertainties. This can be simultaneously: parameter variations (affecting low- and medium-frequency ranges) and unstructured model uncertainties (often located in high-frequency range).

The aim of this paper is the experimental examination of the robustness of digital controllers based on LQ method. Robustness is the property when the dynamic response of control closed loop (including stability of course) is satisfactory not only for the nominal process transfer function used for the design but also for the entire (perturbed) class of transfer functions that expresses uncertainty of the designer in dynamic environment in which a real controller is expected to operate. The design of a robust digital pole assignment controller is investigated in (Landau and Zito 2006), the robust stability of discrete-time systems with parametric uncertainty is analysed in (Matušů 2014). A more comprehensive discussion of robustness is taken when the design based on frequency methods is considered. One can readily compare the system gain at the desired operating point and the point(s) of onset of instability to determine how much gain change is acceptable. Only this method will be used for investigation of the robustness of digital control stable, unstable, non-minimum phase and integrating processes.

The paper is organized in the following way. The fundamental principle of the robustness of digital

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

control-loop systems with basic concepts are illustrated in Section 2. Design of 2DOF digital LQ controller is presented in Section 3. The simulation verification of robust properties of digital closed-loop control with their results are presented in Section 4. Section 5 concludes this paper.

ROBUSTNESS PROBLEM AND BASIC CONCEPTS

Digital Control Loops

The “degrees of freedom” concept is frequently used in the development of control strategies and control system design. The degree of freedom of a control system is defined as the number of closed-loop transfer functions that can be adjusted independently. The design of control systems is a multi-objective problem, so a **T**wo-**D**egree-of-**F**reedom (2DOF) control system naturally has advantages over a **O**ne-**D**egree-of-**F**reedom (abbreviated as 1DOF) control system. This fact was already stated by (Horowitz 1963). Typical structure of the digital control-loop of 2DOF modification is depicted schematically in Fig. 1.

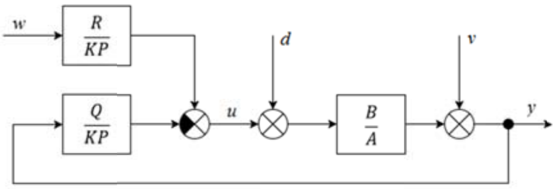


Figure 1: Block diagram of a closed loop 2DOF control system

The specification of individual discrete transfer functions and signals in Fig. 1:

$$G_p(z^{-1}) = \frac{Y(z)}{U(z)} = \frac{B(z^{-1})}{A(z^{-1})} \tag{1}$$

is the process model and

$$G_r\left(z^{-1}\right) = \frac{R\left(z^{-1}\right)}{P\left(z^{-1}\right)K\left(z^{-1}\right)} \tag{2}$$

$$G_q\left(z^{-1}\right) = \frac{Q\left(z^{-1}\right)}{P\left(z^{-1}\right)K\left(z^{-1}\right)} \tag{3}$$

are the feedback part G_q and the feedforward part G_r of individual controllers ; y, u, w are the process output, the controller output and the reference signal; d and v are disturbances and

$$K\left(z^{-1}\right) = 1 - z^{-1} \tag{4}$$

Robustness

The mathematical model of the process is just inaccurate interpretation of the real system, therefore the discrepancies between the model and the real system do not enable the optimal controller design. A number of factors may be responsible for modelling errors. A generic term for this modelling error is *model uncertainty* which can itself be represented mathematically in different ways. Two major classes of uncertainty are: structured uncertainty (parametric uncertainty) and unstructured uncertainty (specified in the frequency domain). The model used for controller design will be termed the *nominal model.* Robustness is a property of the controller that was designed for control of the nominal model and is suitable for control of the family (similar class) models.

Robust control methods are analyzed e.g. in (Morari and Zafirou 1989; Doyle et al.1990; Sánches-Peňa and Sznainer 1998; Skogestad and Postlethwaite 2005; Matušů and Prokop 2011).

Robustness Margins

The open-loop transfer function of 2DOF is

$$G_{OL}\left(z^{-1}\right) = \frac{B\left(z^{-1}\right)Q\left(z^{-1}\right)}{A\left(z^{-1}\right)K\left(z^{-1}\right)P\left(z^{-1}\right)} \tag{5}$$

The frequency characteristic can be obtained by substitution $z = e^{j\omega}$ where ω is the normalized frequency. The Nyquist plot is a parametric plot of a frequency response of the open-loop transfer function $G_{OL}\left(e^{-j\omega}\right)$.

Robustness of closed loops is closely related to the distance between Nyquist plot of open loop transfer function and critical point (-1, *j*0). There are some elements, which help to evaluate this distance: gain margin, phase margin, delay margin and modulus margin.

The *gain margin* G_m is defined as the change in open-loop gain required to make the system unstable. Systems with greater gain margins can withstand greater changes in system parameters before becoming unstable in closed loop. It can be considered as an inverse value of the gain corresponding to a phase shift that occurs at the critical frequency ω_{pc}

$$G_m = \frac{1}{\left|G_{OL}\left(e^{-j\omega_{pc}}\right)\right|} \tag{6}$$

Minimal recommended value of the gain margin is 0.5 (Landau and Zito 2006). The *phase margin* P_m is defined as the change in open-loop phase shift required to make a closed-loop system unstable. The phase margin also measures the system’s tolerance to time-delay. If there is a time-delay greater than $180/\omega_{pc}$ in the open-loop, the system will become unstable in closed-loop (Messner and Tilbury 2011). The phase margin is the additional phase that it must be add at the crossover frequency ω_{cr}, for which the gain of the open-loop system equals 1. It is defined (in degrees) as

$$P_m = 180^0 - \varphi(\omega_{cr}) \quad (7)$$

where $\varphi(\omega_{cr})$ is phase shift at crossover frequency. Typical values for a good phase margin are $30^o \le P_m \le 60^o$ (Landau and Zito 2006).

The *delay margin* τ_m is amount of added time delay that the system can tolerate before it becomes unstable.

$$\tau_m = \frac{P_m}{\omega_{cr}} \quad (8)$$

The *modulus margin* M_m is minimum distance in the Gauss plane between Nyquist plot of open-loop transfer function and critical point (-1, *j*0)

$$M_m = \left|1 + G_{OL}\left(z^{-1}\right)\right|_{min} \quad (9)$$

Notice that the modulus margin is a much more reliable robustness indicator than gain and phase margin. It is possible to consider this indicator as modern approach so-called infinity norm H_∞ (Kwakernaak 1993). It can be shown that good gain margin does not guarantee good phase margin, and vice versa. On the other side, values of modulus margin bigger than 0.5 guarantee $G_m \ge 2$ and $P_m \rangle 29^o$ (Landau 1998).

The gain, phase and modulus margins in the Nyquist plot are depicted in Fig. 2.

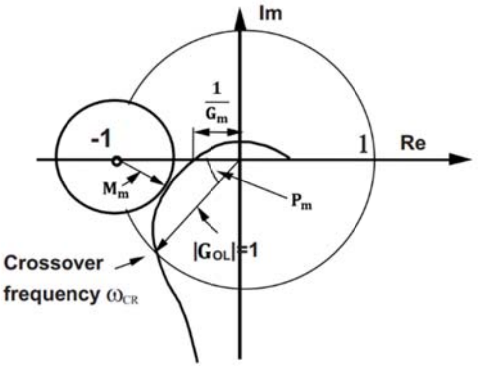


Figure 2: Gain, phase and modulus margins

DESIGN OF POLYNOMIAL 2DOF LQ CONTROLLER

The design of the control algorithm is based on a general block scheme of closed-loop with two degrees of freedom (2DOF) as seen in Fig. 1.

The second-order discrete process model is considered

$$G_P\left(z^{-1}\right) = \frac{B\left(z^{-1}\right)}{A\left(z^{-1}\right)} = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \quad (10)$$

Then the individual parts of digital controller can be expressed in the form of discrete transfer functions

$$G_r\left(z^{-1}\right) = \frac{R\left(z^{-1}\right)}{P\left(z^{-1}\right)K\left(z^{-1}\right)} = \frac{r_0}{\left(1 + p_1 z^{-1}\right)\left(1 - z^{-1}\right)}$$

(11)

$$G_q\left(z^{-1}\right) = \frac{Q\left(z^{-1}\right)}{P\left(z^{-1}\right)K\left(z^{-1}\right)} = \frac{q_0 + q_1 z^{-1} + q_2 z^{-2}}{\left(1 + p_1 z^{-1}\right)\left(1 - z^{-1}\right)} \quad (12)$$

The controller synthesis consists of solving linear polynomial (Diophantine) equations. From first polynomial equation

$$A(z^{-1})K(z^{-1})P(z^{-1}) + B(z^{-1})Q(z^{-1}) = D(z^{-1}) \quad (13)$$

it is possible to calculate 4 feedback controller parameters – coefficients of the polynomials *Q*, *P*. Polynomial $D\left(z^{-1}\right)$ is the characteristic polynomial.

Asymptotic tracking of the reference signal *w* is provided by the feedforward part of the controller which is given by solution of the following polynomial Diophantine equation

$$S(z^{-1})D_w\left(z^{-1}\right) + B(z^{-1})R(z^{-1}) = D(z^{-1}) \quad (14)$$

For a step-changing reference signal value, polynomial $D_w(z^{-1}) = 1 - z^{-1}$ and *S* is an auxiliary polynomial which does not enter into the controller design. Then it is possible to derive the polynomial *R* from equation (14) by substituting $z = 1$

$$R = r_0 = \frac{D(1)}{B(1)} \quad (15)$$

The 2DOF controller output is given by

$$u(k) = \frac{r_0}{K\left(z^{-1}\right)P\left(z^{-1}\right)} w(k) - \frac{Q\left(z^{-1}\right)}{K\left(z^{-1}\right)P\left(z^{-1}\right)} y(k) \quad (16)$$

Polynomial LQ controllers are derived in this paper using minimization of LQ criterion (Kučera 1991).

Minimization of LQ Criterion

The linear quadratic control methods try to minimize the quadratic criterion which uses penalization of the value of the controller output

$$J = \sum_{k=0}^{\infty} \left\{ \left[w(k) - y(k)\right]^2 + q_u \left[u(k)\right]^2 \right\} \quad (17)$$

where q_u is so-called penalization constant, which relates the controller output to the value of the criterion. In this paper, criterion minimization (17) will be realized through the spectral factorization for an input-output description of the system

$$A(z) q_u A\left(z^{-1}\right) + B(z) B\left(z^{-1}\right) = D(z) \delta D\left(z^{-1}\right) \quad (18)$$

where δ is a constant chosen so that $d_0 = 1$.

Because the experimental process model (10) is second-order with second-degree polynomials $A\left(z^{-1}\right), B\left(z^{-1}\right)$, the second-degree polynomial is also obtained from the spectral factorization:

$$D\left(z^{-1}\right)=1+d_1z^{-1}+d_2z^{-2} \tag{19}$$

Spectral factorization of polynomials of the first- and the second- degree can be computed by analytical way; the procedure for higher degrees must be performed iteratively (Bobál et al. 2005). The MATLAB Polynomial Toolbox (Šebek 2014) can be used for a computation of spectral factorization of the higher degree polynomials using file *spf.m* by command

$$\text{d = spf(a*qu*a' + b*b')} \tag{20}$$

It is known that by using the spectral factorization (18), it is possible to compute only two suitable poles (α, β). It is obvious from equation (13) that in this case a choice of the fourth-degree polynomial $D(z)$ is suitable

$$D_4\left(z^{-1}\right)=1+d_1z^{-1}+d_2z^{-2}+d_3z^{-3}+d_4z^{-4} \tag{21}$$

Therefore, the other poles (γ_1, γ_2) are needed to be user-defined. The application of user-defined poles is significant mainly in nonstandard process models (e.g. non-minimum phase, unstable, integrating or with time-delay). The usage of user-defined poles makes it possible to improve the robustness of these processes.
Then the digital 2DOF controller (16) can be expressed in the form

$$\begin{aligned}u(k)=r_0w(k)-q_0y(k)-q_1y(k-1)-q_2y(k-2)\\+(1-p_1)u(k-1)+p_1u(k-2)\end{aligned} \tag{22}$$

where

$$r_0=\frac{1+d_1+d_2+d_3+d_4}{b_1+b_2} \tag{23}$$

and parameters q_0, q_1, q_2, p_1 are computed from (13).

SIMULATION VERIFICATION AND RESULTS

A simulation verification of designed algorithms was performed in MATLAB/SIMULINK environment. The robustness of individual control loops was experimentally investigated by changing characteristic polynomial degree by adding user-defined poles. In addition, the influence of a change of the gain of the nominal process model was examined. From the point of view of robust theory, it is possible to consider these experiments as determination of individual robust stability margins that are caused by parametric uncertainty influence.
This paper presents robustness properties of the digital second-order control systems that can be described by the following continuous-time transfer function:

1) Stable system:

$$G_1(s)=\frac{1}{10s^2+7s+1} \tag{24}$$

2) Non-minimum phase stable system:

$$G_2(s)=\frac{1-2s}{10s^2+7s+1} \tag{25}$$

3) Unstable system:

$$G_3(s)=\frac{1}{8s^2-2s-1} \tag{26}$$

4) Integrating system:

$$G_4(s)=\frac{1}{s(2s+1)} \tag{27}$$

Let us now discretize (24) - (27) using a sampling period $T_0=2$ s. Discrete forms of these transfer functions are (see (10))

$$G_{N1}\left(z^{-1}\right)=\frac{0.1281z^{-1}+0.0803z^{-2}}{1-1.0382z^{-1}+0.2466z^{-2}} \tag{28}$$

$$G_{N2}\left(z^{-1}\right)=\frac{-0.0736z^{-1}+0.2820z^{-2}}{1-1.0382z^{-1}+0.2466z^{-2}} \tag{29}$$

$$G_{N3}\left(z^{-1}\right)=\frac{0.3104z^{-1}+0.3656z^{-2}}{1-3.3248z^{-1}+1.6487z^{-2}} \tag{30}$$

$$G_{N4}\left(z^{-1}\right)=\frac{0.7385z^{-1}+0.5285z^{-2}}{1-1.3679z^{-1}+0.3679z^{-2}} \tag{31}$$

Transfer functions (28) – (31) represent nominal models. From the second-order model (10) and two parts (11), (12) of digital 2DOF controller, the open-loop transfer function can be expressed using individual process and controller parameters

$$G_{OL}(z)=\frac{e_1z^3+e_2z^2+e_3z+e_4}{z^4+f_1z^3+f_2z^2+f_3z+f_4} \tag{32}$$

where

$$\begin{aligned}&e_1=b_1q_0;\quad e_2=b_1q_1+b_2q_2;\quad e_3=b_1q_2+b_2q_1;\\&e_4=b_1q_0;\quad f_1=p_1+a_1-1;\quad f_2=p_1\left(a_2-a_1\right)-a_2;\\&f_3=p_1\left(a_1-1\right)-a_1+a_2;\quad f_4=-a_2p_1\end{aligned} \tag{33}$$

Experimental process models (28) – (31) were used for simulation experiments. Individual simulation experiments are realized subsequently:

- The individual nominal models (28) – (31) G_{Ni} for $(i=1,2,3,4)$ were multiplied by the parameter K_{Pi}, then the perturbed models are given as

$$G_{Pi}=K_{Pi}G_{Ni} \tag{34}$$

The parameter K_{Pi} was increased (decreased) as far as control closed-loops were in the stability boundary - the critical gain K_{ci} was determined.

- These experiments were realized for the case when the poles α, β were computed by spectral factorization and user-defined poles (UDF) were $\gamma_1 = \gamma_2 = 0$. The penalization factor $q_u = 10$ was used for all experiments.
- Obtained critical gains were used for simulation in control-loops when the poles α, β were computed by spectral factorization and $\gamma_1 = 0.2$, $\gamma_2 = 0$.
- The same experimental conditions were used as in the second case, only $\gamma_2 = 0.4$.
- Individual control behaviours of models (28) – (31) are shown in Figs. 3, 5, 7a, 7b and 11.
- Frequency plots of open-loops with the nominal models G_{Ni} were depicted, see (Figs. 4, 6, 8 and 12).
- Values of individual robustness margins - G_m, P_m, M_m were computed (see Tabs. 1 – 4).

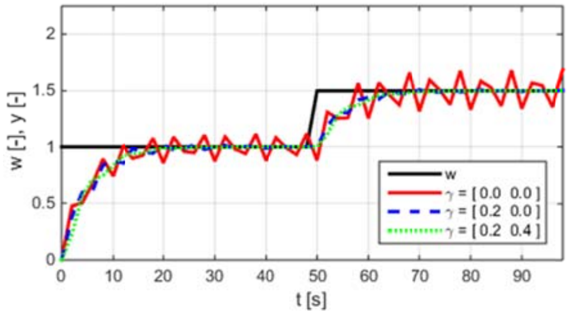


Figure 3: Control of stable model (28) and using UDP $K_{c1} = 3.55$

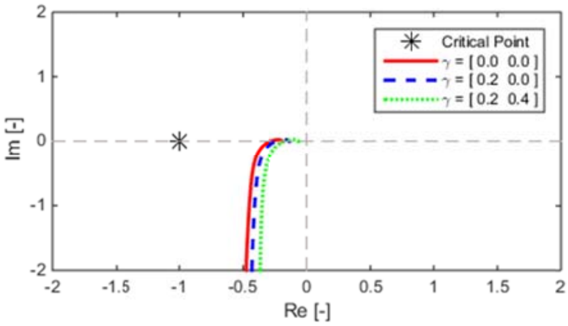


Figure 4: Nyquist plots of open loop with model (28)

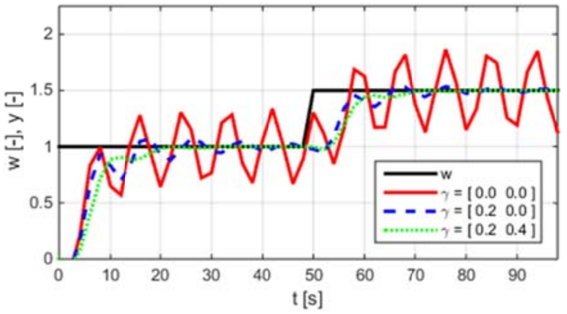


Figure 5: Control of non-minimum phase model (29) and using UDP, $K_{c2} = 1.72$

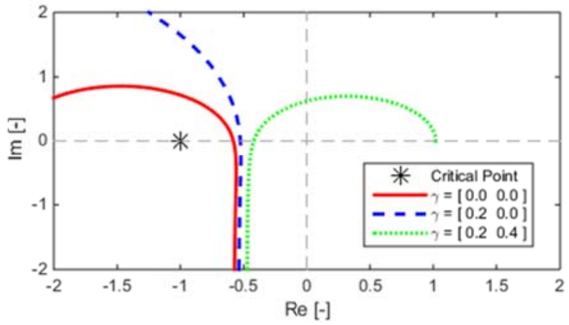


Figure 6: Nyquist plots of open loop with model (29)

In the case of the unstable model (30), the Nyquist plot crosses the real axis in two points (see Fig. 8). Therefore, two critical gains exist and two cases of control closed-loops are in the stability boundary (see Figs. 7a, 7b). It is obvious from Fig. 7a that the addition of UDP does not stabilize the control process, on the contrary the control process is unstable by addition γ_1, γ_2.

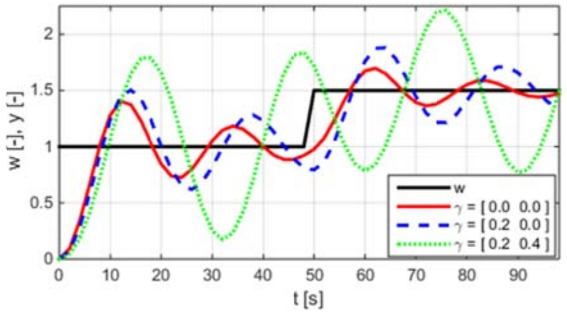


Figure 7a: Control of unstable model (30) and using UDP, $K_{c3a} = 0.78$

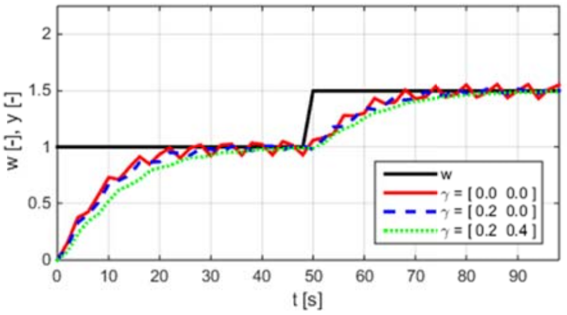


Figure 7b: Control of unstable model (30) and using UDP, $K_{c3b} = 1.29$

Figs. 9 ($K_{c3a} = 0.78$) and 10 ($K_{c3b} = 1.29$) show where Nyquist plots cross the real axis relatively to critical points.

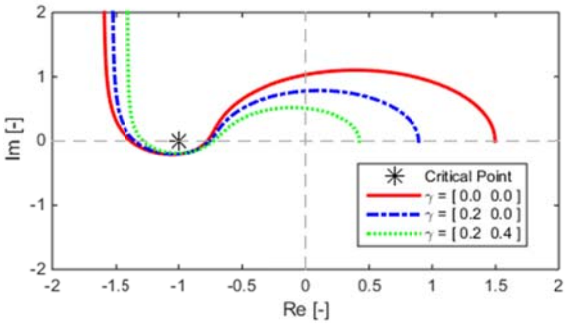


Figure 8: Nyquist plots of open loop with model (30)

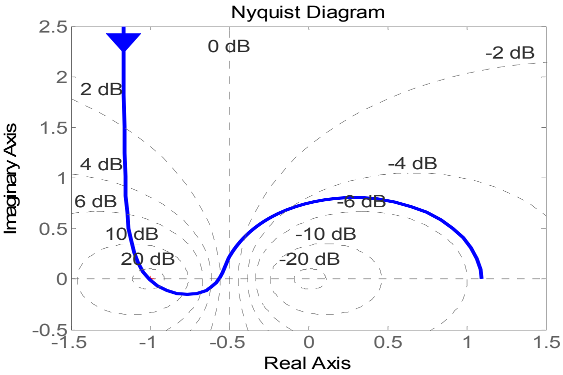


Figure 9: Cross point of Nyquist plot with real axis is located on the right side from critical point (-1, *j*0)

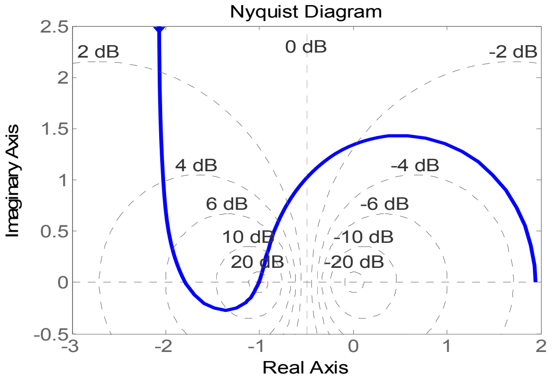


Figure 10: Cross point of Nyquist plot with real axis is located on the left side from critical point (-1, *j*0)

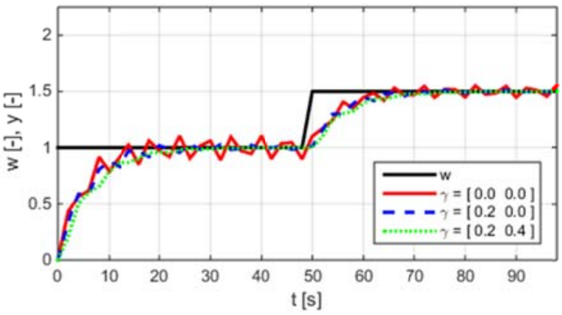


Figure 11: Control of integrating model (31) and using UDP, $K_{c2} = 2.37$

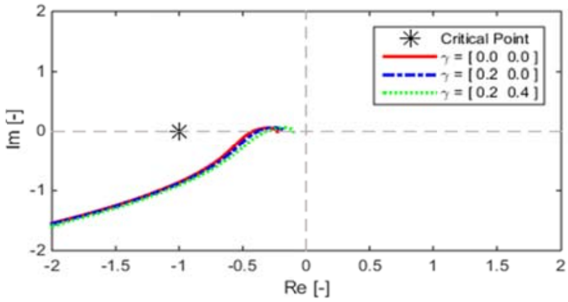


Figure 12: Nyquist plots of open loop with model (31)

Table 1: Robustness Margins, Model (28)

UDP	G_m	P_m	M_m
$\gamma_1 = \gamma_2 = 0$	3.55	63.40	0.64
$\gamma_1 = 0.2; \gamma_2 = 0$	4.18	65.96	0.68
$\gamma_1 = 0.2; \gamma_2 = 0.4$	5.84	69.84	0.74

Table 2: Robustness Margins, Model (29)

UDP	G_m	P_m	M_m
$\gamma_1 = \gamma_2 = 0$	1.72	-34.56	0.41
$\gamma_1 = 0.2; \gamma_2 = 0$	1.90	-50.87	0.47
$\gamma_1 = 0.2; \gamma_2 = 0.4$	2.37	62.18	0.58

Table 3: Robustness Margins, Model (30)

UDP	G_m	P_m	M_m
$\gamma_1 = \gamma_2 = 0$	*	11.42	0.19
$\gamma_1 = 0.2; \gamma_2 = 0$	**	11.37	0.20
$\gamma_1 = 0.2; \gamma_2 = 0.4$	***	10.82	0.19

[*] Interval (0.73 – 1.29)
[**] Interval (0.74 – 1.32)
[***] Interval (0.78 – 1.37)

Table 4: Robustness Margins, Model (31)

UDP	G_m	P_m	M_m
$\gamma_1 = \gamma_2 = 0$	2.37	38.26	0.50
$\gamma_1 = 0.2; \gamma_2 = 0$	2.61	39.31	0.52
$\gamma_1 = 0.2; \gamma_2 = 0.4$	3.15	40.88	0.57

Typical values for stability margins in a robust design are recommended in (Landau and Zito 2006):

1. Gain margin: $G_m \geq 2$, [min: 1.6]
2. Phase margin: $30^{\circ} \leq P_m \leq 60^{\circ}$
3. Modulus margin: $M_m \geq 0.5$, [min: 0.4]

It is obvious from Tabs. 1 – 4 that these recommendations are fulfilled subsequently:

Recommendation 1 is fulfilled for stable system (28) and integrating system (31) – their Nyquist plots are located on the left side of the complex plane (see Figs.

4 and 12) and they cross the real axis relatively near the point (0, 0).

Recommendation 2 is almost fulfilled for system (28) and fully for system (31).

Recommendation 3 can be fulfilled for all systems except of unstable system (30).

Remarks: The negative $-P_m$ in the system (29) is caused by an unstable zero (see the first and second Nyquist plot – Fig. 6). The unstable system (30) does not fulfil any recommended value. The open-loop of this system has two cross points of Nyquist plot with real axis (-0.73, *j*0) and (-1.29, *j*0) – see Figs. 8 -10. The perturbed system is at the stability boundary for $K_{p3} = K_{c3a}$ and $K_{p3} = K_{c3b}$ and is stable between these points $\left(K_{c3a} < K_{p3} < K_{c3b}\right)$ - see Tab.3.

CONCLUSION

It is well known that the robust theory is applied mainly for design of continuous-time controllers for control of stable systems. The paper presents an experimental simulation investigation of robustness of algorithms for digital LQ control. Individual control algorithms were verified using 2DOF modification in the MATLAB/Simulink environment. Four types of process models (stable, non-minimum phase, unstable and integrating) were used for controller design. The influence of change of process gain (parametric uncertainty) was used for determination of the stability boundary. The gain, phase and modulus margins of the Nyquist plot of open-loop sampled-data system were used as robustness indicators. Values of these indicators showed that only stable and integrating processes are in recommended intervals. Simulation experiments demonstrated that robustness margins of examined digital control closed-loop systems can be improved by addition of user-defined poles.

REFERENCES

Bobál, V., Böhm, J., Fessl, J. and J. Macháček. 2005. *Digital Self-tuning Controllers: Algorithms, Implementation and Applications.* Springer-Verlag, London**.**

Doyle, J., Francis, B. and A. Tannenbaum. 1990. *Feedback Control Theory.* 1990. Macmillan Publishing.

Horowitz, I., M. 1963. *Synthesis of Fedback Systems.* Academia Press.

Kučera, V. 1991. *Analysis and Design of Discrete Linear Control Systems.* Prentice-Hall, Englewood Cliffs, NJ.

Kučera, V. 1993. "Diophantine equations in control – a survey". *Automatica* 29, 1361-1375.

Kučera, V. 1991. *Analysis and Design of Discrete Linear Control Systems.* Prentice-Hall, Englewood Cliffs, NJ.

Kwakernaak, H. 1993. „Robust control and H_∞ optimization – Tutorial paper". *Automatica* 29, 255-273.

Landau, I. D. 1998. "The R-S-T digital controller design and applications". *Control Engineering Practice* 7,155-165.

Landau, I. D. 1999. "From robust control to adaptive control," *Control Engineering Practice* 7, 1113-1124.

Landau, I. D. and G. Zito. 2006. *Digital Control Systems.* Springer-Verlag, London.

Landau, I. D., Lozano, R., M'Saad, M. and A. Karimi. 2011. *Adaptive Control, Algorithms. Analysis and Applications.* Springer-Verlag, London.

Matušů, R. 2014. "Robust stability analysis of discrete-time systems with parametric uncertainty," *International Journal of Mathematical Models and Methods* 8, 95-102.

Matušů, R. and R. Prokop. 2011. "Graphical analysis of robust stability for systems with parametric uncertainty: an overview," *Transactions of the Institute of Measurement and Control* 33, 274-290.

Messner, B. and D. Tilbury. 2011. *Control Tutorials for MATLAB and Simulink.* MathWorks, Natick, MA, USA.

Morari, M. and E. Zafiriou. 1989. *Robust Process Control.* Prentice-Hall, Englewood Cliffs, NJ.

Sánches-Peňa, R. S. and M. Szainer. 1998. *Robust Systems. Theory and Application.* J. Willey & Sons, New York.

Šebek, M. Polynomial Toolbox for MATLAB, Version 3.0. 2014. PolyX, Prague, Czech Republic.

Šebek, M. and M. Hromčík. 2007. "Polynomial design methods*,"* *International Journal of Robust and Nonlinear Control* 17, 679-681.

Skogestad, S. and I. Postletwaithe. 2005. *Multivariable Feedback Control – Analyses and Design.* J. Wiley & Sons, Chichester.

AUTHOR BIOGRAPHIES

VLADIMÍR BOBÁL graduated in 1966 from the Brno University of Technology, Czech Republic. He received his Ph.D. degree in Technical Cybernetics at Institute of Technical Cybernetics, Slovak Academy of Sciences, Bratislava, Slovak Republic. He is now Professor at the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interests are adaptive and predictive control, system identification and CAD for automatic control systems. You can contact him on e-mail address bobal@fai.utb.cz.

ĽUBOŠ SPAČEK studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master's degree in Automatic Control and Informatics in 2016. He currently attends PhD study at the Department of Process Control. His e-mail address is lspacek@fai.utb.cz.

PETER HORNÁK studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master's degree in Automatic Control and Informatics in 2016. He is currently working in private company focused on industrial robotics. His e-mail address is pet.hornak@gmail.com.

MODELING OF CORN EARS BY DISCRETE ELEMENT METHOD (DEM)

Ádám Kovács and György Kerényi
Department of Machine and Product Design
Budapest University of Technology and Economics
H-1111, Budapest, Hungary
E-mail: kovacs.adam@gt3.bme.hu

KEYWORDS
Ear of corn, maize, discrete element method, agricultural material.

ABSTRACT

One of the main aims of today's agriculture is to reduce the losses of different agricultural processes. With more than 1000 billion tons of production annually, corn is one of the most essential agricultural crops in the world so our study focuses on the modeling of corn ears by discrete element method (DEM) from the point of view of losses during corn harvesting. To describe the physical and mechanical properties of corn ears and its connections to the corn stalk, laboratorial and in-situ tests and observations were conducted. During an iteration process the mechanical model of the corn ear and the connection between the corn ear and the corn stalk were validated. There has been a good match between the numerical results and the experimental data. The simulation results clearly demonstrate that the discrete element model of corn ear is capable of modeling losses during corn harvesting in the future.

INTRODUCTION

The increasing demand for more and more and better quality agricultural products presents a big challenge for farmers, breeders and developers of agricultural machineries. Thus, one of the perpetual goals of precision agriculture is to reduce the losses during the agricultural processes. There are several different methods to prevent agricultural losses but in our study the reduction of losses in agricultural machineries is discussed.

Due to the seasonal characteristics of agricultural products in situ tests of constructions are limited in time and often prove to be very expensive. In the field of agricultural machine design numerical methods are not available which could properly replace field tests.

The utilization of corn plants and crops is remarkable worldwide; the corn production in the world is almost 1000 million tons annually. In 2015 almost 7 million tons of corns were harvested by farmers in Hungary (Hungarian Central Statistical Office 2017), which demonstrates the significance of the plant in the agriculture of the country.

During mature corn processing the first step is harvesting, when combines with special corn headers gather the crop. The first time maize gets in contact with the machine is, when the corn header picks up the corn ears from the stalk. In this process two types of losses can occur: losses of corn kernels and losses of corn ears. Losses of corn kernels are caused by collisions among the ears and different parts of the corn header. Losses of corn ears may be caused by high ear picking forces or collisions in the corn header. Because the average number of corn kernels per ears is 800 pieces, the loss of a corn ear is more significant than the loss of kernels. Therefore, our study focuses on the modeling of corn ears by discrete element method (DEM) from point of view of corn ear loss in corn header during harvesting.

DEM is widely used to investigate bulk agricultural materials. The micromechanical parameters of a sunflower DEM model were calibrated based on odometer tests so that the model can sufficiently approach the macro mechanical behavior of the real bulk material (Keppler et al. 2011). In another study the effect of particle shape of corn kernels on flow was investigated by discrete element method simulation of a rotary batch seed coater (Pasha et al. 2016).

In connection with complete plant and corn ear modeling fewer literatures can be found. The iteration among grass stalk and rotation mower was investigated by DEM (Kemper et al. 2014). A special solid geometrical structure of DEM was analysed for corn stalks in quantitative and qualitative ways (Kovács et al. 2015). Several possible DEM geometrical structures for modelling of fibrous agricultural materials were compared (Kovács and Kerényi 2016). The whole corn ear, the corn cob, corn kernels and the connection among the kernels and the cob, were modelled with DEM in order to simulate the corn threshing process (Yu et al. 2015). An interactive digital design system was developed by C++ programming language based on OpenGL graphic library for corn modelling (Xiao et al. 2010).

Consequently, based on the literature review it is clear that there are no DEM model that can model the connection between the corn ear and the corn stalk and the collision among different materials and corn ears. Therefore, experimental apparatus has been developed for measuring the ear picking force and the coefficient of restitution among corn ears and different surfaces in order to develop a DEM model for modeling losses during corn harvesting.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

MATERIALS AND METHODS

Discrete element method (DEM) was developed to investigate bulk materials which contain separate parts. The DEM model is defined as follows: it contains separated, discrete particles which have independent degrees of freedom and the model can simulate the finite rotations and translations, connections can break and new connections can come about in the model (Cundall and Hart 1993).

Based on the harvest and product processes of harvest-ready maize the main loads of the corn ear were determined. Leaves and husk of the corn ear were neglected in our study. First of all, the physical and physiological properties of the corn ear (mass, length, diameter, shape, position in the stalk, center of mass) were measured and observed. Laboratorial ear picking force and collision tests were conducted to define the main mechanical parameters and the behavior of corn stalks. The quantitative results of the measures weren't directly usable for the modeling method so necessary data and graphs were calculated with mathematical and statistical and image processing methods for the numerical modeling.

The examination of the available contact models was the first step of the modeling. After that the models were compared and the Timoshenko-Beam-Bonded model (Brown et al. 2014), which is based on the Timoshenko-beam theory, was selected for the study.

In the next step, the DEM physical geometry of the corn ear was created to calibrate its mechanical properties based on the experimental results. After that the geometry model of the stalk, the shank were created so as to calibrate the mechanical properties of the stalk-ear connection.

With modifications of the micro-mechanical parameters of the contact models, during an iteration process, the right assembly was found.

MEASUREMENTS AND OBSERVATIONS

Two types of measurements and observations were conducted in connection with the corn ears: in-situ and laboratorial. In both cases the same type of maize from the middle region of Hungary (GPS coordinates: N 47.743692; E 19.613025) were analyzed in October 2016.

In-situ measurements and observations

During the in-situ measurements and observations the position of the ears, the center of the mass of the ears from the stalk and from the ground were analyzed.

There are two types of positions of the corn ear on the corn stalk: hanging and standing, as shown on Figures 1. In the first case the shank of the corn ear is broken while on the other hand it is unbroken. Based on 100 observed plants, 59% of the plants had hanging corn ears.

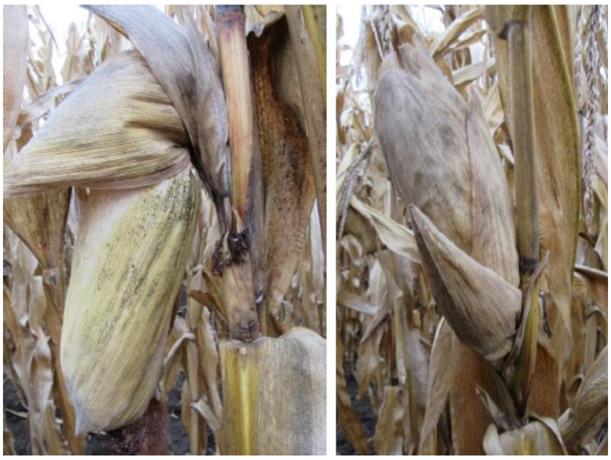


Figure 1: Positions of corn ears: hanging position on the left side and standing position on the right side

Needless to say, these two different positions mean different parameters for the center of the mass (CM) and from this reason the different positions can cause different dynamic behavior for the corn stalk during harvesting. Based on 20 samples, the distances of CM from the ground and from the stalk of the standing corn ears are higher than the case of hanging corn ears, as shown in Table 1.

Table 1: Center of mass of corn ears

Property	Value (P=0.05) [cm]
CM from ground, hanging	92.9 ± 5.1
CM from stalk, hanging	10.1 ± 1.3
CM from ground, standing	108.2 ± 4.5
CM from stalk, standing	5.6 ± 0.5

Laboratorial measurements and observations

First, the main diameter, the length and the mass of corn ears were measured, as shown in Table 2.

Table 2: Physical properties of corn ears based on 20 samples

Property	Value (P=0.05)
Average main diameter	47.6 ± 0.9 mm
Average length	169.8 ± 8.7 mm
Average mass	211.2 ± 16.0 g

To model the shape of the ears an image based analysis was conducted with 20 corn ears being observed. In this process an axis symmetric property of the corn ears was assumed. First, a picture was taken from each corn ear in front of the same background. After that a background was completely removed by image processing software. In the next step the pictures were transformed to binary images with luminance factor 0.95, so the boundaries of the images were easily found, the process is shown on Figure 2.

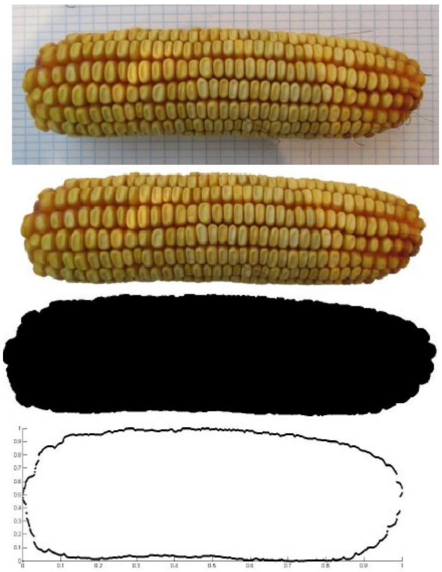


Figure 2: Image processing steps of corn ears shape

Based on the boundaries, the centers of mass of the corn ears could be calculated by the equations of first moment of area and an average ear can be formulated, as shown on Figure 3.

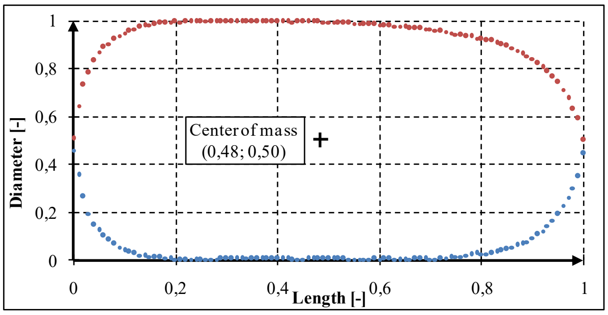


Figure 3: Ideal shape and average center of mass of corn ears

The size of the shank of the corn ear impacts on the ear picking force essentially. The shank has two main dimensions: diameter and length, so these parameters were measured on 20 samples. The results are shown in Table 3.

Table 3: Physical properties of the shank based on 20 samples

Property	Value (P=0.05) [mm]
Average diameter	10.6 ± 0.4
Average length	115.8 ± 15.7

One of the most significant mechanical parameters of the corn ears is the coefficient of restitution among the corn ears and different surfaces. Thus, the coefficient of restitution among corn ears and steel (S235) and plastic (Polyethylene, PE) sheets were measured in axial and radial directions of the corn ears. The samples were dropped into the sheets (size 50x50 cm, thickness of the plastic: 5 mm; thickness of the steel: 1 mm) from 1 meter height and the collision was detected with 1000 frames per second by a high-speed camera. The coefficient of restitution (C_R) was calculated from the bounce height (h) and the drop height (H) with Equation (1).

$$C_R = \sqrt{\frac{h}{H}} \tag{1}$$

Based on 10 samples, respectively, significant differences between the axial and radial coefficients of restitution were not found and the coefficient of restitution between the ear and the sheet of plastic was higher, as it was expected (Figure 4).

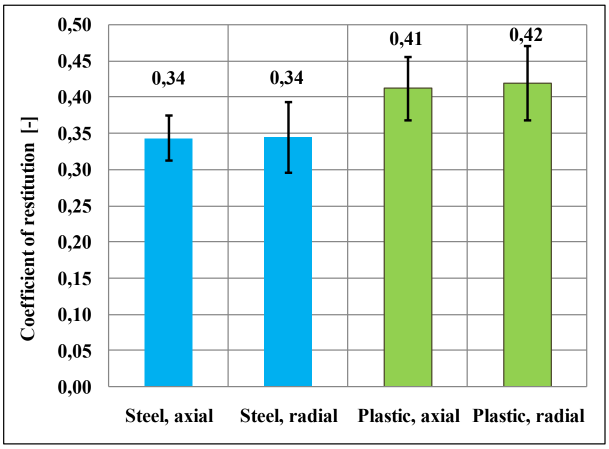


Figure 4: Coefficients of restitution of corn ears

To model the connection between the corn ears and stalks, the ear picking force was measured on a self-developed apparatus that can copy the ear gathering process in the corn header, as shown on Figure 5.

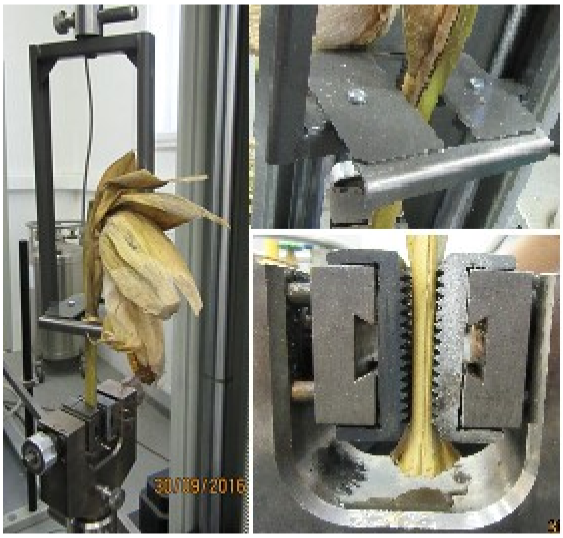


Figure 5: Apparatus for measurement of ear picking force

Based on 20 samples, the average ear picking force was 457.3 ± 58.2 N (P=0.05) and the dimensionless characteristics with the confidence interval (P=0.05) of the ear picking process is shown on Figure 6.

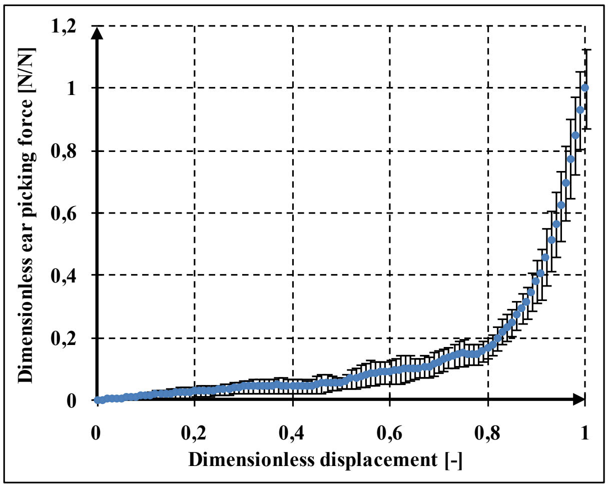


Figure 6: Characteristics of ear picking force

MODEL FORMATION

The discrete element model formation is constituted of four main parts: geometrical and contact structure formation of the stalk, the shank and the ear of the plant and the modeling of the test environment. During the model formation the measured and observed parameters of the plants and EDEM 2.7 Academic (DEM Solution Ltd.) were used.

In the stalk a hollow geometrical structure with 18 particles in one cross section was used based on our previous studies. In this geometrical structure different bonds among the particles were used in tangential and axial directions to model different mechanical behaviors in these directions of the stalk (Kovács and Kerényi 2016), as seen on Figure 7.

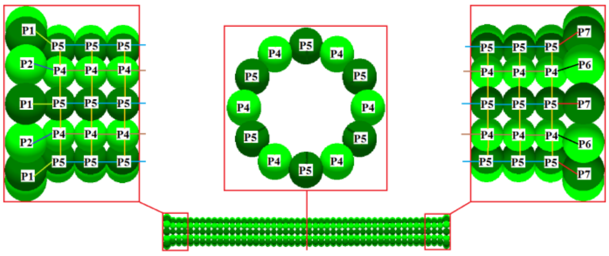


Figure 7: Geometrical and contact structure of the stalk

The shank provides the contact between the stalk and the ear. It was modeled by a simpler geometrical structure of chain of spheres. In this model there is only one type of bonds among the particles in its axial direction but there is another bond among the shank and the stalk. The mechanical properties in tangential properties are provided by the stiffness of the particles. The shape of the shank was formed in such a way that it could carry a hanging ear because this type of ears were more common during the in-situ observations. It is necessary to note that it is impossible to model this broken condition of the shank so a curved shank with unbroken bonds among the particles was modeled, as shown on Figure 8.

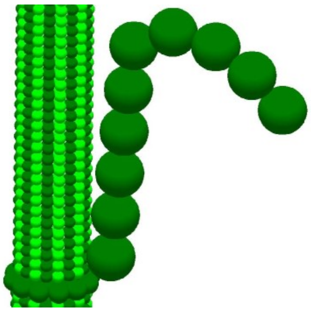


Figure 8: The shank on the stalk

The geometrical model of the corn ear is one particle that is formed by several sphere surfaces and it doesn't contain bonds. The previously described ideal shape of the corn ears (Figure 3) was approached by 25 sphere surfaces, supposing that the corn ears are axis symmetric, as shown on Figure 9.

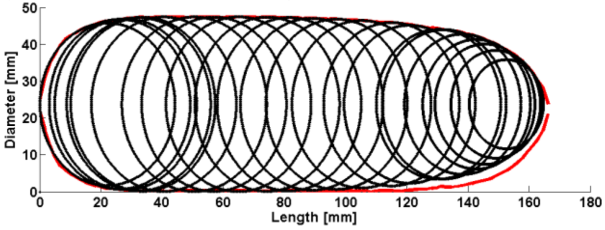


Figure 9: Ideal (red line) and modeled (black spheres) shape of the corn ear

The corn ear was situated in such a way that its center of the mass was near the same as the results from the measures, as shown on Figure 10. Naturally, a bond is needed between the end of the shank and the corn ear.

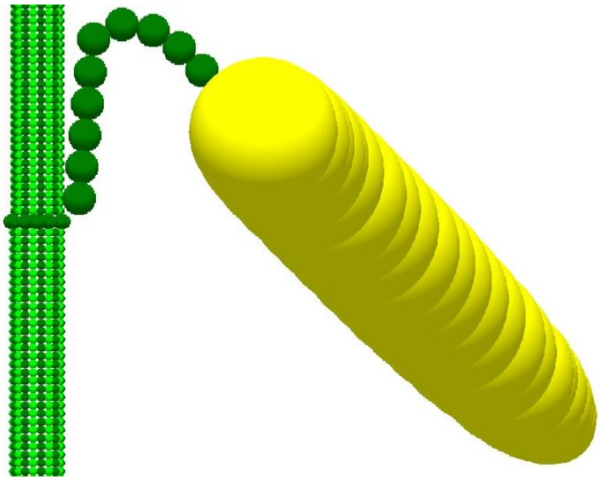


Figure 10: The entire model: stalk, shank and corn ear

Lastly, the testing machines were integrated into the model. In the case of the collision simulation, only a steel and a plastic sheet were modeled, based on the material properties of the real ones. In the other case the clamps that hold the stalk and the half of the ear picking apparatus were modeled, as shown on Figure 11.

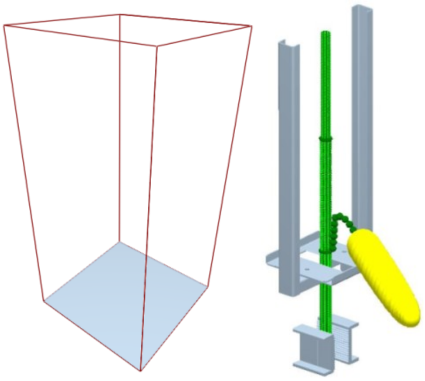


Figure 11: Model of the collision (left) and the ear picking (right) simulation

RESULTS

After the model formation the mechanical parameters of the models were modified during an iteration process in order to find the right set of parameters that can accurately mimic the real behavior of the different plant parts.

First, the coefficients of restitution between the corn ears and the different surfaces were analyzed. In the simulations the collision between the ear and the sheet was taken place in radial direction. Based on the collision test the coefficients of restitution among the corn ear and sheets of steel and plastic were chosen at 0.34 and 0.41, respectively. From these values the bounce heights could be calculated with Equation (1): 115.6 mm; 176.4 mm on steel and plastic respectively.

The collisions are calculated in the software from the following numerical material properties: coefficient of restitution between the materials and shear moduli of the materials. The shear moduli of the steel and plastic are well known: 8e10 Pa and 1.17e8 Pa, respectively. These values were set to the material parameters of the model.

After that the coefficients of restitution were set to the measured values and the shear modulus of the corn ear was modified in the range of 1e8 – 1e14 Pa until it reached the maximum value (1e14 Pa) that the software allowed. Unfortunately, the expected bounce height was not reached with these parameters so another calibration method was chosen.

In the next step the shear modulus of the corn ear was chosen for a fixed value 1e10 Pa, based on our experiences and the literature (Yu et al. 2015). With the fixed shear moduli the numerical coefficient of restitution was modified from 0.5 to 1.0.

In both cases the relationship between the bounce height and the coefficient of restitution shows a second order polynomial characteristics with R^2=0.99. The exact numerical coefficient of restitution values for the observed bounce heights are 0.77 and 0.88 for steel and plastic, respectively, as shown on Figure 12. The difference among the measured and the numerical coefficient of restitution values comes from the calculation method of the bounce in the discrete element software where it uses an algorithmic damping.

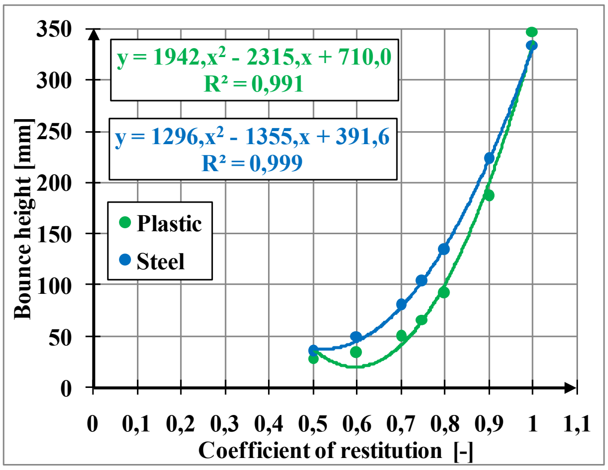


Figure 12: Characteristics of bounce height as a function of coefficient of restitution among corn ear and steel and plastic

After the right set of parameters for the collision of the ear had been selected, the ear picking force of the plant was calibrated. In this step the main objective was the calibration of the mechanical parameters of the bonds among the stalk, the shank and the corn ear in order to simulate the ear gathering process.

With the right set of parameters the maximum ear picking force from the simulation was 439 N that is in accordance with the measured value. The simulated characteristics coincide very well with the measured characteristics of the real process, as shown on Figure 13.

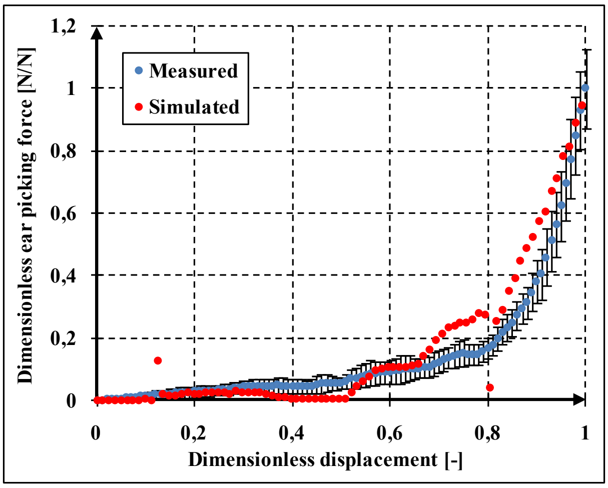


Figure 13: Measured and simulated characteristics of ear picking process

Based on our practical experiences the bonds were calibrated in such a way that the shank is torn from the stalk but it is in contact with the ear after the gathering, as shown on Figure 14.

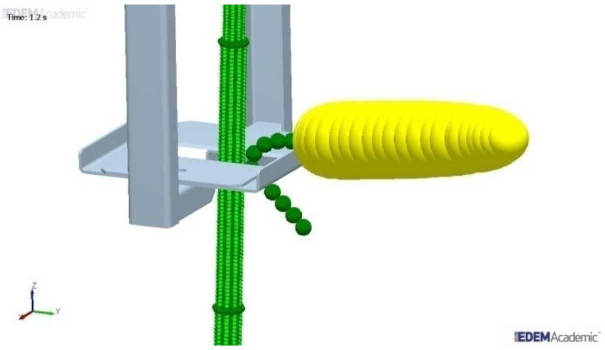


Figure 14: The final state of the ear gathering process in the simulation

The obtained simulation results are referred to a quasi-static measurement method but they are adaptable to analyze the basic phenomena during ear gathering process.

CONCLUSIONS

A numerical and experimental study of the corn ear collision and picking force was undertaken using discrete element method in order to simulate corn ear losses during the harvesting process of harvest-ready maize. The effect of numerical shear moduli, coefficient of restitution on collision among corn ear and steel and plastic sheets was analyzed and a right set of bond parameters among the stalk, the shank and the corn ear was calibrated for the simulations in the future. The following conclusions could be drawn:

(1) The applied measurement method is usable to provide experimental results for discrete element models of shank and ear of maize and for the contacts among these parts.
(2) The applied geometrical structures of the stalk, shank and corn ear are suitable for further analysis.
(3) The bounce height as a function of coefficient of restitution shows a second order polynomial characteristics for the collisions between the corn ear and sheets of plastic and steel.
(4) The calibrated physical and mechanical properties of the interactions and bonds among the stalk, shank, ear of corn and geometrical elements are suitable for further analysis.
(5) The characteristics of simulation and the measurement are coincided very well; thus, the discrete element model of ear picking is capable to simulate the harvesting process.
(6) The discrete element numerical models are capable of simulating the interactions among the testing apparatus and different parts of maize and compare the simulation and experimental results.

As to the measurement, the applied method should be extended for more samples, more maize species and different fertilizing conditions.

In the future, the current results and models can be adapted to more detailed and realistic simulations on losses of corn ears in a corn header of a combine harvester during harvesting process.

ACKNOWLEDGEMENTS

The authors thank Dr. Gábor Szebényi from the Department of Polymer Engineering, Dr. Imre Obrulov from the Department of Material Science and Engineering, Szabolcs Berezvai from the Department of Applied Mechanics and Dr. Tibor Poós from the Department of Building Services and Process Engineering at BUTE for the help they provided in the laboratorial tests.

REFERENCES

Brown N.J.; J-F. Chen; J. Y. Ooi. 2014. “A bond model for DEM simulation of cementitious materials and deformable structures.” *Granular Matter*, No. (2014) 16, 299–311.

Cundall P.A.; R.D. Hart. 1993. “Numerical Modeling of Discontinua.” *Analysis and Design Methods*, No. 1993, 231-243.

Kemper S.; T. Lang; L. Frerichs. 2014. “The overlaid cut in a disc mower - results from field tests and simulation.” *Landtechnik*, No. 69(4), 171-175.

Keppler, I.; L. Kocsis; I. Oldal; A. Csatár. 2011. „Determination of the discrete element model parameters of granular materials.” *Hungarian Agricultural Engineering*, No. 23/2011, 30-32.

Kovács Á.; K. Kotrocz; Gy. Kerényi. 2015. “The adaptability of discrete element method (DEM) in agricultural machine design.” *Hungarian Agricultural Engineering*, No. 27, 14-19.

Kovács Á.; GY. Kerényi. 2016. “Comparative analysis of different geometrical structures of discrete element method (DEM) for fibrous agricultural materials.” In *4th CIGR International Conference of Agricultural Engineering* (Aarhus, Denmark, June 26-29), 1-8.

Pasha M.; C. Hare; M. Ghadiri; A. Gunadi; P. M. Piccione. 2016. “Effect of particle shape on flow in discrete element method simulation of a rotary batch seed coater.” *Powder Technology*, Volume 296, 29–36.

Xiao B.; X. Guo; X. Du; W. Wen; X. Wang; S. Lu. 2010. “An interactive digital design system for corn modeling.” *Mathematical and Computer Modeling*, No. 51 (11-12), 1383-1389.

Yu Y.; H. Fu; J. Yu. 2015. “DEM-based simulation of the corn threshing process.” *Advanced Powder Technology*, No. 26 (5), 1400-1409.

AUTHOR BIOGRAPHIES

ÁDÁM KOVÁCS was born in Debrecen, Hungary and went to Budapest University of Technology and Economics, where he studied agricultural machine design and obtained his MSc. degree in 2016. Currently he is a PhD student in the same institution and his topic is discrete element modeling of maize. He worked for the WIGNER Research Center for Physics at the Department of Plasma Physics for two years, where he designed diagnostic devices for fusion reactors. His e-mail address is: kovacs.adam@gt3.bme.hu and his Web-page can be found at http://gt3.bme.hu/en.

GYÖRGY KERÉNYI studied agricultural machine design at Szent István University, Gödöllő and after that he went to Budapest University of Technology and Economics, where he obtained his PhD degree in 1997. Currently he is an associate professor and deputy head of Department of Product and Machine Design in the same institution and his research topic is numerical methods in agricultural machine design. His e-mail address is: kerenyi.gyorgy@gt3.bme.hu and his Web-page can be found at http://gt3.bme.hu/en.

OPTIMAL CONTROL WITH DISTURBANCE ESTIMATION

František Dušek,
Daniel Honc,
Rahul Sharma K.
Department of Process control
Faculty of Electrical Engineering and Informatics, University of Pardubice, Czech Republic
E-mail: {frantisek.dusek,daniel.honc}@upce.cz, rahul.sharma@student.upce.cz

KEYWORDS
Control, optimal control, LQ controller, model predictive control, disturbance, estimation, thermal process.

ABSTRACT

The paper deals with a very common situation in many control systems and this is the fact that, for zero control action, the controlled variable is nonzero. This is often caused by the existence of another process input which is uncontrolled. Classic controllers do not take into account the second input, so deviation variables are considered or some feedforward controller is used to compensate the variable. The authors propose a solution, that the process is considered as a system with two inputs and single output (TISO). Here, the uncontrolled input is estimated with the state observer and the controller is designed as the multivariable controller. A Linear-quadratic (LQ) state-feedback control and model predictive control (MPC) of simple thermal process simulations are provided to demonstrate the proposed control strategy.

INTRODUCTION

Control theory is frequently using models in the form of transfer functions, which from the definition, consider zero initial conditions (Åström and Murray 2010; Nise 2010; Ogata 1995; Skogestad and Postlethwaite 2005). This means practically that for zero control action the controlled variable will be zero as well. Unfortunately, this is not true for many practical applications. Even a P controller will not work very well and the situation is even worse for advanced controllers based on state space process models. These models are similar to P or PD controller formulations – without integral control action. One solution is subtracting working point variables and introducing deviation variables – "zero initial condition" will be met. Integral control action, to ensure offset-free reference tracking, is another interesting problem to solve (Maeder and Morari 2010; Dušek et al. 2015). But why not use the natural process model with disturbance variables and their dynamical effects, directly for the controller design? Then the control task can be solved as a multivariable control problem when only some of the process inputs are used as control variables, while the others are considered as disturbances. Disturbance modelling and state estimation for offset free reference tracking control problems, was published in (Muske and Badgwell 2002; Pannocchia and Rawlings 2003; Tatjewski 2014).

Authors propose to estimate the disturbance variable by the augmented state observer. Extended formulation of a standard LQ state-feedback controller and predictive controller, so that the disturbance information is an integral part of the controller, is presented in the paper. A simple thermal process with electrical heating, ambient temperature effect and temperature sensor is modeled analytically by the first principle approach. The model has two inputs and one output. One of the inputs is heating power, while the other is ambient temperature. The output is the temperature sensor measured temperature. A discrete time linear time invariant process model is used for LQ controller design with infinity horizon and asymptotic set point tracking and predictive controller with finite horizon and special formulation of the cost function. Deviations of future states from desired states, calculated from the future set point knowledge, are considered instead of the future control errors which are commonly used in the literature (Camacho and Bordons 2007; Kouvaritakis and Cannon 2015; Maciejowski 2002; Rawlings and Mayne 2009; Rossiter 2003).

PROCESS MODEL WITH OFFSETS

Let us consider the controlled process with variable u_m as the control variable (control action) and y_m as controlled variable. Disturbance (offset) variables u_0 and y_0 are considered as process input and additive disturbance on the process output – see block diagram in Fig. 1.

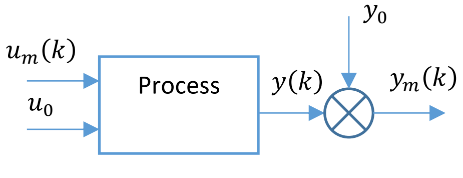


Figure 1: Process model

Discrete time state space process model can be written as,

$$\begin{aligned} \boldsymbol{x}(k+1) &= \mathbf{A}\boldsymbol{x}(k) + \mathbf{b}u_m(k) + \mathbf{b}_0 u_0 \\ y_m(k) &= \mathbf{c}\boldsymbol{x}(k) + y_0 \end{aligned} \tag{1}$$

If we know the steady state input and both the offsets, the steady state output can be calculated as

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

$$y_m = \underbrace{\mathbf{c}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}}_{p}\, u_m + \underbrace{\mathbf{c}(\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}_0}_{p_0}\, u_0 + y_0 \quad (2)$$

and the steady state input (we will use this in controller design)

$$u_m(k) = \frac{1}{p}(y_m - y_0) - \frac{p_0}{p}u_0 \qquad (3)$$

DISTURBANCE STATE ESTIMATION

The disturbances can be measured or estimated. In our case, we are using augmented state estimation for estimating the state vector and disturbance variable u_0, while y_0 must be known. It is not possible to estimate both offsets simultaneously. If u_0 is known, y_0 can be calculated from the steady state.

We introduce the augmented state space model as

$$\underbrace{\begin{bmatrix}\mathbf{x}(k+1)\\ u_0\end{bmatrix}}_{\mathbf{x}_r(k+1)} = \underbrace{\begin{bmatrix}\mathbf{A} & \mathbf{b}_0\\ 0 & 1\end{bmatrix}}_{\mathbf{A}_r}\underbrace{\begin{bmatrix}\mathbf{x}(k)\\ u_0\end{bmatrix}}_{\mathbf{x}_r(k)} + \underbrace{\begin{bmatrix}\mathbf{b}\\ 0\end{bmatrix}}_{\mathbf{b}_r} u_m(k) \qquad (4)$$

$$\underbrace{y_m(k) - y_0}_{y_r(k)} = \underbrace{[\mathbf{c} \quad 0]}_{\mathbf{c}_r}\underbrace{\begin{bmatrix}\mathbf{x}(k)\\ u_0\end{bmatrix}}_{\mathbf{x}_r(k)}$$

State estimator with gain **K** has the form

$$\hat{\mathbf{x}}_r(k+1) = \mathbf{A}_r\hat{\mathbf{x}}_r(k) + \mathbf{b}_r u_m(k) + \\ + \mathbf{K}\big(y_m(k) - y_0 - \mathbf{c}_r\hat{\mathbf{x}}_r(k)\big) \quad (5)$$

We are estimating in vector $\hat{\mathbf{x}}_r(k)$, all the state variables and disturbance variable u_0, from variables $u_m(k)$, $y_m(k)$ and from the known output offset y_0.

CONTROLLER DESIGN

Two types of controllers based on the state space process model are modified so that the estimation of the disturbance variable u_0 can be used as an integral part of the controller design.

LQ controller

Linear-quadratic state-feedback controller with infinite horizon cost function is

$$J = \sum_{i=1}^{\infty}\begin{bmatrix}\mathbf{x}^T(k+i)\mathbf{Q}\mathbf{x}(k+i) + \\ u^T(k+i-1)\mathrm{R}u(k+i-1)\end{bmatrix} \qquad (6)$$

Negative state feedback controller part is

$$u(k) = -\mathbf{L}\mathbf{x}(k) \qquad (7)$$

To be able to follow the set point asymptotically we are introducing a feedforward path with control variable $u_f(k)$ – see Fig. 2.

Control action is

$$u_m(k) = u(k) + u_f(k) \qquad (8)$$

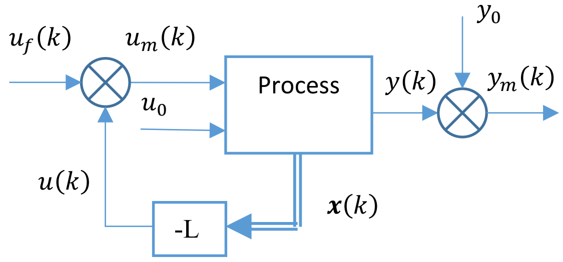


Figure 2: LQ Controller

Steady state output can be calculated as

$$y_m = \underbrace{\mathbf{c}(\mathbf{I}-\mathbf{A}+\mathbf{b}\mathbf{L})^{-1}\mathbf{b}}_{p_L}\, u_f + \\ + \underbrace{\mathbf{c}(\mathbf{I}-\mathbf{A}+\mathbf{b}\mathbf{L})^{-1}\mathbf{b}_0}_{p_{oL}}\, u_0 + y_0 \qquad (9)$$

If $y_m = w(k)$ then the feedforward control variable is

$$u_f(k) = \frac{1}{p_L}(w(k) - y_0) - \frac{p_{oL}}{p_L}u_0 \qquad (10)$$

and the control action is

$$u_m(k) = -\mathbf{L}\mathbf{x}(k) + u_f(k) \qquad (11)$$

Model predictive controller

We consider the special matrix form cost function formulation for model predictive controller as

$$J(N) = (\mathbf{x}_N - \mathbf{x}_{Nw})^T\mathbf{Q}_N(\mathbf{x}_N - \mathbf{x}_{Nw}) + \mathbf{u}_N^T\mathbf{R}_N\mathbf{u}_N \ (12)$$

where $\mathbf{u}_N$ is the vector of future control actions deviations from previous control action for a prediction horizon of N, which is given by,

$$\mathbf{u}_N = \underbrace{\begin{bmatrix}u_m(k)\\ u_m(k+1)\\ \vdots\\ u_m(k+N-1)\end{bmatrix}}_{\mathbf{u}_{Nm}} - \underbrace{\begin{bmatrix}u_m(k-1)\\ u_m(k-1)\\ \vdots\\ u_m(k-1)\end{bmatrix}}_{\mathbf{u}_{Nm0}}$$

and $\mathbf{x}_N$ is the vector of future predicted states deviations from future desired states $\mathbf{x}_{Nw}$

$$\mathbf{x}_N - \mathbf{x}_{Nw} = \underbrace{\mathbf{S}_{xx}\mathbf{x}(k) + \mathbf{S}_{xu}\mathbf{u}_{Nm} + \mathbf{S}_{xu0}\mathbf{u}_{N0}}_{\mathbf{x}_N} - \mathbf{x}_{Nw} = \\ = \mathbf{S}_{xx}\mathbf{x}(k) + \mathbf{S}_{xu}\mathbf{u}_N + \underbrace{\mathbf{S}_{xu}\mathbf{u}_{Nm0} + \mathbf{S}_{xu0}\mathbf{u}_{N0} - \mathbf{x}_{Nw}}_{\mathbf{o}}$$

$$\mathbf{x}_N = \begin{bmatrix}\mathbf{x}(k+1)\\ \mathbf{x}(k+2)\\ \vdots\\ \mathbf{x}(k+N)\end{bmatrix}, \mathbf{S}_{xx} = \begin{bmatrix}\mathbf{A}\\ \mathbf{A}^2\\ \vdots\\ \mathbf{A}^N\end{bmatrix}, \mathbf{u}_{N0} = \begin{bmatrix}u_0\\ u_0\\ \vdots\\ u_0\end{bmatrix},$$

$$
\mathbf{S}_{xu}=\begin{bmatrix}\mathbf{b} & 0 & \cdots & & 0\\ \mathbf{Ab} & \mathbf{b} & & & \\ \vdots & & \ddots & & \vdots\\ \mathbf{A}^{N-2}\mathbf{b} & \mathbf{A}^{N-3}\mathbf{b} & \cdots & \mathbf{b} & 0\\ \mathbf{A}^{N-1}\mathbf{b} & \mathbf{A}^{N-2}\mathbf{b} & & \mathbf{Ab} & \mathbf{b}\end{bmatrix}
$$

$$
\mathbf{S}_{xu0}=\begin{bmatrix}\mathbf{b}_0 & 0 & \cdots & & 0\\ \mathbf{Ab}_0 & \mathbf{b}_0 & & & \\ \vdots & & \ddots & & \vdots\\ \mathbf{A}^{N-2}\mathbf{b}_0 & \mathbf{A}^{N-3}\mathbf{b}_0 & & \mathbf{b}_0 & 0\\ \mathbf{A}^{N-1}\mathbf{b}_0 & \mathbf{A}^{N-2}\mathbf{b}_0 & \cdots & \mathbf{Ab}_0 & \mathbf{b}_0\end{bmatrix}.
$$

Cost function (12) can be transformed to a form,

$$
\begin{gathered}
J(N)=\mathbf{u}_N^T\underbrace{(\mathbf{R}_N+\mathbf{S}_{xu}^T\mathbf{Q}_N\mathbf{S}_{xu})}_{\mathbf{M}}\mathbf{u}_N+ \qquad (13)\\
\mathbf{u}_N^T\underbrace{\mathbf{S}_{xu}^T\mathbf{Q}_N[\mathbf{S}_{xx}\mathbf{x}(k)+\mathbf{o}]}_{\mathbf{m}}+\underbrace{[\mathbf{S}_{xx}\mathbf{x}(k)+\mathbf{o}]^T\mathbf{Q}_N\mathbf{S}_{xu}}_{\mathbf{m}^T}\mathbf{u}_N+\\
\underbrace{\mathbf{x}^T(k)\mathbf{S}_{xx}^T\mathbf{Q}_N\mathbf{S}_{xx}\mathbf{x}(k)+\mathbf{x}^T(k)\mathbf{S}_{xx}^T\mathbf{Q}_N\mathbf{o}+\mathbf{o}^T\mathbf{Q}_N\mathbf{S}_{xx}\mathbf{x}(k)+\mathbf{o}^T\mathbf{Q}_N\mathbf{o}}_{c}
\end{gathered}
$$

Solution for the unconstrained case to this quadratic form can be calculated analytically as

$$
\mathbf{u}_N=-\mathbf{M}^{-1}\mathbf{m} \qquad (14)
$$

and the actual control action is

$$
u_m(k)=u_m(k-1)+\mathbf{u}_N(1) \qquad (15)
$$

where $\mathbf{u}_N(1)$ is first element of vector of optimal future control actions deviation from previous control action.

Vector of future desired states $\mathbf{x}_{Nw}$ is calculated from the future set points as

$$
\mathbf{x}_{Nw}=\begin{bmatrix}\mathbf{x}_w(k+1)\\ \mathbf{x}_w(k+2)\\ \vdots\\ \mathbf{x}_w(k+N)\end{bmatrix} \qquad (16)
$$

where

$$
\mathbf{x}_w(k+i)=(\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}u_w(k+i)+(\mathbf{I}-\mathbf{A})^{-1}\mathbf{b}_0u_0
$$

and

$$
u_w(k+i)=\frac{1}{p}[w(k+i)-y_0]-\frac{p_0}{p}u_0
$$

THERMAL PROCESS

We consider the simple thermal process, where E is a heating power, T_o is ambient temperature, T_E, T and T_C are temperatures of the heating element, body of the system and the temperature sensor respectively. The system has two inputs and one output – see Fig. 3.

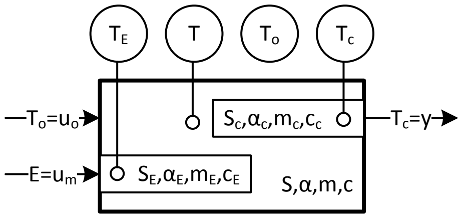


Figure 3: Thermal process

We are modeling the process analytically with first principle and we consider individual subsystems as systems with lumped parameters for the sake of simplicity.

Energy balance of the heating element is

$$
E=\underbrace{\alpha_E S_E}_{s_1}(T_E-T)+\underbrace{m_E c_E}_{m_1}\frac{dT_E}{dt} \qquad (17)
$$

Energy balance of body of the system is

$$
\begin{gathered}
\alpha_E S_E(T_E-T)=\alpha_c S_c(T-T_c)+\underbrace{\alpha S}_{s_2}(T-T_o)+\\
+\underbrace{mc}_{m_2}\frac{dT}{dt} \qquad (18)
\end{gathered}
$$

Energy balance of the temperature sensor is

$$
\underbrace{\alpha_c S_c}_{s_3}(T-T_c)=\underbrace{m_c c_c}_{m_3}\frac{dT_c}{dt} \qquad (19)
$$

State space model of the whole process is

$$
\begin{bmatrix}\frac{dT_E}{dt}\\ \frac{dT}{dt}\\ \frac{dT_c}{dt}\end{bmatrix}=\begin{bmatrix}-\frac{s_1}{m_1} & \frac{s_1}{m_1} & 0\\ \frac{s_1}{m_2} & -\frac{s_1+s_2+s_3}{m_2} & \frac{s_3}{m_2}\\ 0 & \frac{s_3}{m_3} & -\frac{s_3}{m_3}\end{bmatrix}\begin{bmatrix}T_E\\ T\\ T_c\end{bmatrix}+
$$

$$
+\begin{bmatrix}\frac{1}{m_1}\\ 0\\ 0\end{bmatrix}E+\begin{bmatrix}0\\ \frac{s_2}{m_2}\\ 0\end{bmatrix}T_o \qquad (20)
$$

For the following simulations, we consider the parameters of the process as given in Table 1.

Table 1: Process parameters

		J.s^{-1}.K^{-1}		J. K^{-1}
Heating	s_1	0.5	m_1	1
Body	s_2	2.5	m_2	25
Sensor	s_3	0.1	m_3	0.5

SIMULATION RESULTS

The gain of the state observer is calculated as a solution of dual problem to a linear-quadratic state-feedback controller for discrete-time state-space system calculated in MATLAB as with command

$$[\mathbf{K}^T,\sim,\sim] = \mathrm{dlqr}(\mathbf{A}_r^T,\mathbf{c}_r^T,\mathbf{Q}_e,R_e)$$

The penalization matrices are selected as $\mathbf{Q}_e$ = eye(4) and R_e = 0.1, and the sample time Ts = 2.5 s. State and disturbance estimation are demonstrated in Fig. 4. After a few seconds the state estimation errors drop to zero and the disturbance variable T_o is correctly estimated.

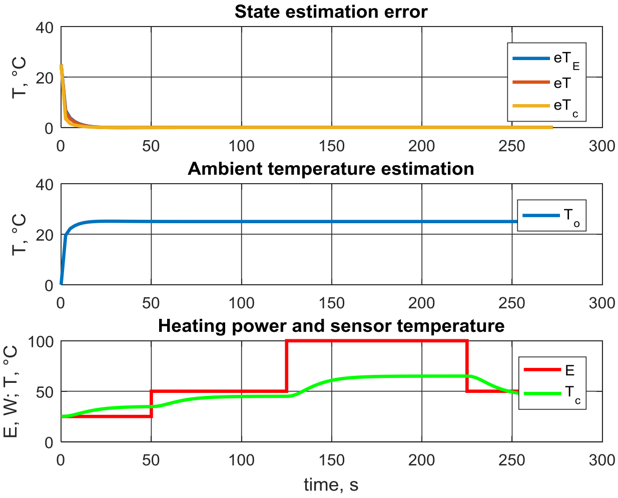


Figure 4: State and disturbance estimation

The gain of the LQ controller is calculated in the same way as the observer gain with MATLAB command, but only with modified penalization matrix **Q**

$$[\mathbf{L},\sim,\sim] = \mathrm{dlqr}(\mathbf{A},\mathbf{b},\mathbf{Q},\mathrm{R})$$

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 100 \end{bmatrix}$$

The control experiment can be seen in Fig. 5. The set point w is followed by the output T_c by controlling the heating power E (we do not consider constrains).

The predictive controller has identical parameters. The prediction horizon N = 15. Fig. 6 shows the control response with the predictive controller. It can be seen that, the predictive controller starts in advance before the set point change and the quality of the control is slightly higher – standard deviation (SD) of control error is 8.9 °C compared to 12.7 °C for LQ controller.

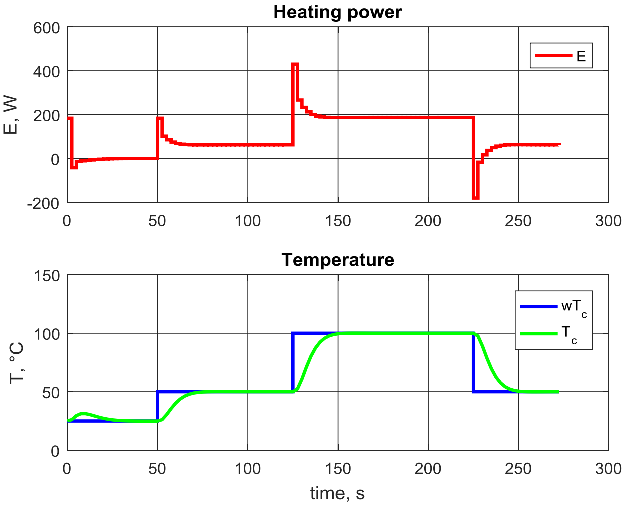


Figure 5: Control with LQ controller

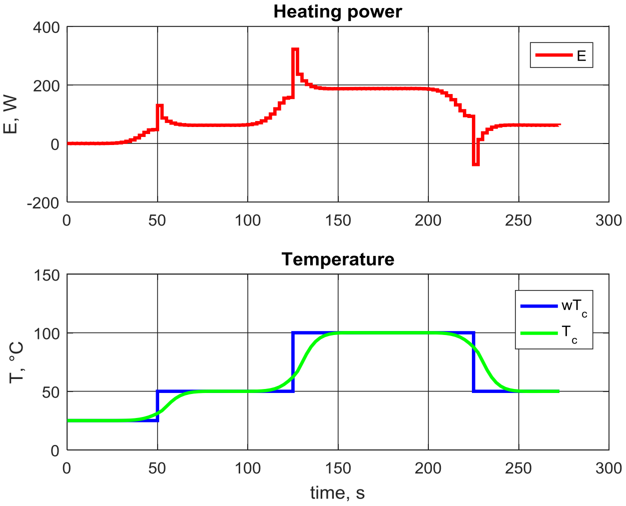


Figure 6: Control with predictive controller

CONCLUSION

The paper deals with a practical control issue, where for the steady state zero control input, the controlled variable is nonzero because of the offset or disturbance. Classic control methods are dealing with this problem by introducing deviations from a working point, integral actions, or feedforward parts of the standard single input single output (SISO) controllers. Authors propose, to work with the processes as with multivariable systems, and to design the controllers as a multivariable system. The disturbance (uncontrolled input variable) estimation method is presented in the paper. Subsequently, LQ and predictive controller design methods are modified that the estimated disturbance can be used as an integral part of the controllers. The set point is followed asymptotically with the LQ controller – feedforward controller path uses the offset information. Similarly, offset is used in the model predictive controller in free response calculation and for future desired states calculation as well. We are controlling sensor temperature but, it is also possible (without any problems) to control temperature of the body of the system; only by changing vector **c** of the process model for the controller design.

The paper is a nice example of the strength and elegance of state space methods for modelling, estimation and control. According to authors' opinion these methods will become acutely relevant in the future.

This research was supported by project SGS, Modern Methods for Simulation, Control and Optimization at FEI, University of Pardubice. This support is very gratefully acknowledged.

REFERENCES

Åström, K.J., Murray, R.M. 2010. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press.

Camacho, E.F., Bordons, C. 2007. *Model Predictive control, Advanced Textbooks in Control and Signal Processing*, Springer.

Dušek, F., Honc, D., Sharma, K. 2015. A comparative study of state-space controllers with offset-free reference tracking. In *20th International Conference on Process Control*. IEEE, 2015, 176-180.

Kouvaritakis, B., Cannon, M. 2015. *Model Predictive Control, Classical, Robust and Stochastic, Advanced Textbooks in Control and Signal Processing*, Springer.

Maciejowski, J. M. 2002. *Predictive Control with Constraints*. Pearson.

Maeder, U., Morari, M. 2010. Offset-free reference tracking with model predictive control. *Automatica*. 2010, vol. 46, issue 9, 1469-1476.

Muske, K.R., Badgwell, T.A. 2002. Disturbance modeling for offset-free linear model predictive control. *Journal of Process Control*. 2002, vol. 12, issue 5, 617-632.

Nise, N.S. 2010. *Control Systems Engineering*. Wiley.

Ogata, K. 1995. *Discrete-time control systems*. Prentice Hall.

Pannocchia, G., Rawlings, J.B. 2003. Disturbance models for offset-free model-predictive control. *AIChE Journal*. 2003, vol. 49, issue 2, 426-437.

Rawlings, J.B. and Mayne D.Q. 2009. *Model Predictive Control: Theory and Design*, Nob Hill Publishing, LLC.

Rossiter, J.A. 2003. *Model-Based Predictive Control: A Practical Approach*, CRC Press.

Skogestad, S., Postlethwaite, I. 2005. *Multivariable Feedback Control: Analysis and Design*. Wiley.

Tatjewski, P. 2014. Disturbance modeling and state estimation for offset-free predictive control with state-space process models. *International Journal of Applied Mathematics and Computer Science*. 2014, vol. 24, issue 2.

AUTHOR BIOGRAPHIES

FRANTIŠEK DUŠEK was born in Dačice, Czech Republic and studied Faculty of Chemical Technology Pardubice field of Automation and obtained his MSc. degree in 1980. He worked for pulp and paper research institute IRAPA. Now he is a vice dean of Faculty of Electrical Engineering and Informatics. In 2001 he became Associate Professor. His e-mail address is: `frantisek.dusek@upce.cz`

DANIEL HONC was born in Pardubice, Czech Republic and studied University of Pardubice field of Process Control and obtained his Ph.D. degree in 2002. He is head of the Department of Process Control at Faculty of Electrical Engineering and Informatics. His e-mail address is: `daniel.honc@upce.cz`

RAHUL SHARMA K., was born in Kochi, India and went to the Amrita University, where he studied electrical engineering and obtained his M.Tech degree in 2013. He is now doing his Ph.D. studies at Department of process control, Faculty of Electrical and Informatics, University of Pardubice, Czech Republic. His e-mail address is: `rahul.sharma@student.upce.cz`

MODELLING AND MODEL PREDICTIVE CONTROL OF MAGNETIC LEVITATION LABORATORY PLANT

Petr Chalupa
Jakub Novák
Martin Malý
Faculty of Applied Informatics
Tomas Bata University in Zlin
nam. T. G. Masaryka 5555, 760 01, Czech Republic
E-mail: chalupa@fai.utb.cz

KEYWORDS
State-space model, model predictive control, MATLAB, Simulink, magnetic levitation, CE152 model

ABSTRACT

The paper is focused on creating a mathematical model of a magnetic levitation plant and usage of the model for a design of a predictive controller. The magnetic levitation laboratory plant CE 152 by Humusoft Company is used to determine values of model parameters and for real time control experiments. From the control point of view, the CE152 represents a nonlinear and very fast system. Both the mathematical model and the model predictive controller are created using MATLAB / Simulink environment. This environment extended by Real time toolbox is used for real time experiments with the laboratory plant.

INTRODUCTION

Design techniques of current controllers are in most cases based on some kind of model of the real-time plant to be controlled (Bobál et al. 2005). There are many types of mathematical models of the real-time plats but linear models is the most popular category for the control design. More sophisticated nonlinear models can represent the plant in more details than linear models but the controller design is much more complicated in most cases. Moreover, controllers based on nonlinear models are in general more computationally demanding and thus are not suitable for fast systems.

This paper is based on previous work (Chalupa et al. 2016) where detailed nonlinear Simulink model of the CE152 Magnetic Levitation plant (Humusoft 1996) is presented. The model was derived using first principle modelling (Himmelblau and Riggs 2004) and values of the model parameters were specified by real-time measurements. This approach is often referred to as “grey-box modelling” (Tan and Li 2002). The final Simulink model is used to linearize the model around a suitable operating point in the proposed paper (Ljung 1999).

The linear model is used to design a model predictive controller (MPC) (Camacho and Bordons, 2004). The state space linear model is used and the receding horizon principle is applied to the process of minimization of the MPC criterion (Kwon and Han, 2005).

The main goal of the work presented in this paper was to design and verify a model predictive controller for the CE152 Magnetic levitation system. There are several challenges concerning the design process: instability of the controlled plant, nonlinearity of the controlled plant and very fast response of the plant.

CE152 MAGNETIC LEVITATION PLANT

A photo of the CE 152 magnetic levitation plant is presented in Figure 1.

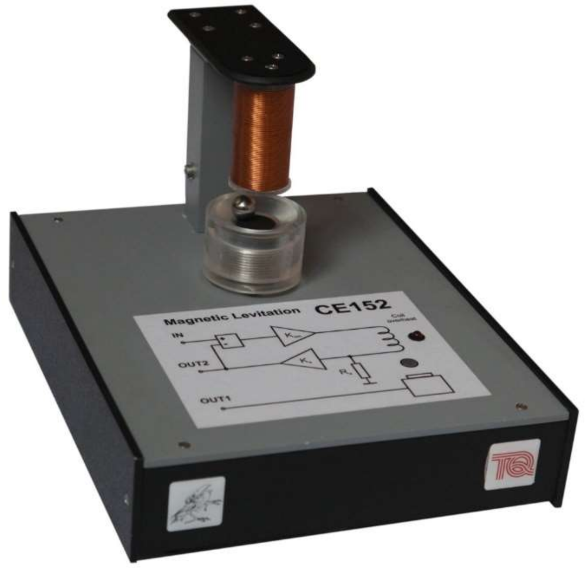


Figure 1: Photo of the CE152 plant

The CE152 Magnetic levitation system is a nonlinear unstable dynamic system with one input and one output. The input signal affects the ball position while the output signal gives information about position of a steel ball. Both signal values are converted and scaled to the specific range of the machine unit [MU].

Structure of the CE152 Magnetic levitation system

The system consists of a model of the magnetic levitation system, power supply and a universal data acquisition card MF624. MF624 is a standard PCI card with A/D, D/A converters, analogue/digital inputs and outputs,

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

counters, timers and appropriate drivers. The model is connected to the PC via this card.

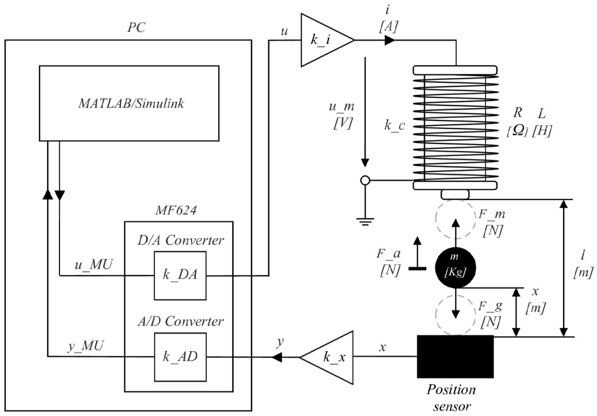


Figure 2: Scheme of the magnetic levitation plant

Simplified inner structure is shown in Figure 2. A steel ball levitates in magnetic field of the coil driven by power amplifier connected to D/A converter. Position of the steel ball is measured by inductive linear position sensor connected to A/D converter. Both control and measured parameters are sent and received by Simulink.

System behaviour

The CE152 Magnetic levitation system is a nonlinear unstable single input single output dynamic system. When an input control signal of certain value is sent to the system, the ball is lifted upwards to the magnetic core and it stops when it hits the core. This behaviour is caused by electromagnetic force of the magnetic core, which overcame the force of gravity. As the ball getting closer to the coil core, accelerating force grows. Because of an obstacle in the form of magnetic core, both the ball and accelerating force stops. Higher input signal means higher electromagnetic force and much more rapidly increasing acceleration. If control value decreases under a certain value, electromagnetic force is not big enough to overcome gravitational force and the ball falls down.

MODELLING OF THE CE 152 PLANT

The process of creating a mathematical model of the CE152 Magnetic Levitation laboratory plant is described in detail in (Chalupa et al. 2016). This chapter presents just results important for a design of a control system.
The magnetic levitation plant can be divided into several parts and each of these parts can be modelled separately:

- D/A converter,
- power amplifier,
- ball and coil subsystem,
- position sensor,
- A/D converter.

D/A converter

The D/A converter converts digital signal u_{MU} from PC into an analogue voltage signal u. The D/A converter can be described by a linear function (1):

$$u = k_{DA}u_{MU} \tag{1}$$

where u is D/A converter output signal/coil input voltage [V], u_{MU} is D/A converter input signal [MU] and k_{DA} is D/A converter gain [V/MU].

Power amplifier

The power amplifier represents a source of constant current, which is proportional to its control voltage:

$$i = k_i u \tag{2}$$

where k_i is a gain of the power amplifier.

Ball and coil subsystem

Lagrange's method can be used for modelling ball and coil subsystem. The motion equation is based on the balance of all acting forces.

$$m_k\ddot{x} + k_{fv}\dot{x} = \frac{i^2 k_c}{(x-x_0)^2} - m_k g \tag{3}$$

where: g - gravitational acceleration [m.s^{-2}]
x - ball position [m]
m_k - ball mass [kg]
x_0 - coil offset [m]
kc - coil constant [-]
i - coil current [A]
k_{fv} - dumping constant [N.s.m^{-1}]

Position sensor

An inductive position sensor is used to measure the ball position x. The relation between ball position and voltage is approximately linear.

$$y = k_x x + y_0 \tag{4}$$

where x is ball position [m], y_0 position sensor offset [V], y position sensor voltage [V], k_x - position sensor gain [V/m]

A/D converter

The A/D converter converts analogue voltage signal y into a digital signal y_{MU}. The A/D converter can be described by a linear function (5):

$$y_{MU} = k_{AD}y \tag{5}$$

where y represents A/D converter output signal/position sensor voltage [V], y_{MU} is A/D converter output signal [MU] and k_{AD} is A/D converter gain [MU/V].

Model of the whole system

The model of the whole system consists of joined models of individual parts – equations (1) – (5). A Simulink scheme of the whole model is presented in Figure 3.

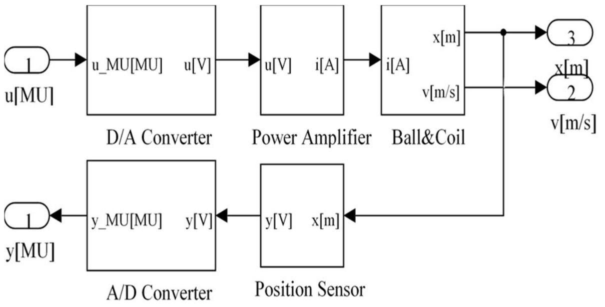


Figure 3: Simulink model of the whole system

LINEARIZATION OF THE NONLINEAR MODEL

Linearized state-space model will be created in this part. The CE 152 plant can be represented by a Single Input Single Output (SISO) model with a state vector consisting of ball velocity - *v* and ball position - *x*. Input voltage is represented by *u* and *y* represent the output of the system – i.e. A/D converter output. System can be described as presented in Figure 4.

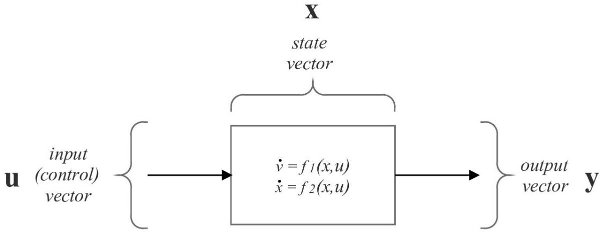


Figure 4: State space system

In general, a nonlinear continuous time SISO system can be described by the following equations:

$$\dot{\boldsymbol{X}} = \boldsymbol{f}(\boldsymbol{X}, u_{MU}) \tag{6}$$

$$y = g(\boldsymbol{X}, u_{MU}) \tag{7}$$

Where $\boldsymbol{X}$ is a state vector, *u* is an input signal and *y* represents the output signal. In case of CE 152 the sate vector consists of two scalars:

$$\boldsymbol{X} = [x, v]^T \tag{8}$$

The system described by equations (6) – (8) can be linearized around some operating point and system matrix (A), input matrix (B), output matrix (C) and feedthrough matrix (D) are calculated.

Parameters of these matrices are obtained through linearization of following differential equations at selected operating (nominal) point.

$$\frac{\boldsymbol{dv}}{\boldsymbol{dt}} = \frac{(k_{DA}k_i u_{MU})^2 k_c}{m_k(x-x_0)^2} - \frac{k_{fv}}{m_k} v - \mathrm{g} \tag{9}$$

$$\frac{dx}{dt} = v \tag{10}$$

$$y_{MU} = k_{AD}k_x x + k_{AD}y_0 \tag{11}$$

Determination of matrices parameters is done based on partial derivatives of differential equations with respect to state variables and input variables at the selected operating point.

$$A = \frac{\partial f_i}{\partial x_i}(x, u_{MU})\,|_{v_{ss}, x_{ss}, u_{MUss}} \tag{12}$$

$$B = \frac{\partial f_i}{\partial u_i}(x, u_{MU})\,|_{v_{ss}, x_{ss}, u_{MUss}} \tag{13}$$

$$C = \frac{\partial g_i}{\partial x_i}(x, u_{MU})\,|_{v_{ss}, x_{ss}, u_{MUss}} \tag{14}$$

$$D = \frac{\partial g_i}{\partial u_i}(x, u_{MU})\,|_{v_{ss}, x_{ss}, u_{MUss}} \tag{15}$$

Steady state is assumed and then the derivatives in equations (9), (10) are set to be zero, which leads to

$$v_{ss} = 0 \tag{16}$$

$$u_{MUss} = \pm\sqrt{\frac{m_k g(x-x_0)^2}{k_C k_{DA}^2 k_i^2}} \tag{17}$$

Now final shape of the matrices is:

$$A = \begin{bmatrix} -\frac{k_{fv}}{m_k} & -\frac{2g}{x_{ss}-x_0} \\ 1 & 0 \end{bmatrix} \tag{18}$$

$$B = \begin{bmatrix} -\frac{2k_{DA}k_i\sqrt{k_C g}}{\sqrt{m_k}(x_{ss}-x_0)} \\ 0 \end{bmatrix} \tag{19}$$

$$\mathrm{C} = [0 \quad k_{AD}k_x] \tag{20}$$

$$D = [0] \tag{21}$$

State-space representation can be converted to transfer function form. For this case of a SISO model, the resulting transfer function is only one in the form:

$$G(s) = \frac{b_0}{a_2 s^2 + a_1 s + a_0} \tag{22}$$

where partial coefficients are:

$$b_0 = -\frac{2k_{DA}k_{AD}k_x k_i\sqrt{k_C g}}{\sqrt{m_k}(x_{ss}-x_0)} \tag{23}$$

$$a_0 = -\frac{2g}{x_{ss}-x_0} \quad a_1 = \frac{k_{fv}}{m_k} \quad a_2 = 1 \tag{24}$$

The transfer function gain is a static variable and it doesn't depend on the chosen operating point.

Selection of an operating point

In general, not all operating points are necessarily suitable for linearization. It is a question, how adequate is the linearized approximation in comparison with the real system dynamics. First step is to specify value of the operating point parameters, in this case only one parameter is needed and this is to specify a position of the ball x_{ss}. The chosen operating point corresponds to the position of a ball approximately in the middle of the space between the magnetic core and the head of the inductive sensor. Then the result value is:

$$x_{ss} = -\frac{l}{2} = 2.85 \cdot 10^{-3}[m] \tag{25}$$

Verification and validation of the linearized model

It is assumed that nonlinear model is sufficiently accurate and so given linearized model should be too, but it must be considered, that linear model will be accurate only, when the system state is near the operating point.

DISCRETIZATION OF THE LINEAR MODEL

The continuous-time linear model presented in previous chapter is not suitable for controller design because many MPC algorithms are based on discrete-time model of the controlled system. Therefore, continuous-time model was transformed to its discrete-time version. Sampling period T_0=0.001 s was used for all experiments.

It is assumed, that discrete-time state space system looks as follows:

$$x_{k+1} = \Phi x_k + \Gamma u_k \tag{26}$$

$$y_k = \mathrm{C} x_k + \mathrm{D} u_k \tag{27}$$

where

$$\Phi = e^{AT_0} \quad \Gamma = \int_0^{T_0} e^s ds\, B \tag{28}$$

There are more solutions, how get the matrices Φ and Γ. Some of them are presented in (Kwon and Han, 2005) and MATLAB offers the *c2d* function to process this problem.

STATE ESTIMATION

State variables are mostly unmeasurable and they need to be estimated if the state space form is used. For observation (prediction) of state variables are used so-called filters because a part of this estimator works in the same way as classic filter. Other part estimates state variables from measurable variables.

Moving Horizon Estimation (MHE) is used in this paper to obtain estimation of the system states. This optimization approach uses observed measurements over a horizon, containing noise and disturbances, and produces estimates of unknown desired variables. Unlike Kalman filter, which can be solved by deterministic (non-iterative) approach, MHE requires an iterative approach that relies on linear or nonlinear programming solvers. There are many MHE filters and the one used in this work is known as the L_2E filter (also LEF filter). It gives very satisfying results in terms of computation speed and precision. Detailed procedure of derivation the LEF filter can be find in in (Kwon and Han, 2005).

The state estimations for the current step can be computed using the following equation:

$$\hat{x}(k) = HY_{k-1} + LU_{k-1} \tag{29}$$

Output values Y_{k-1} and input values U_{k-1} on defined horizon are given as:

$$Y_{k-1} = [y_{k-N_f}, y_{k-N_f+1}, \dots\, y_{k-1}] \tag{30}$$

$$U_{k-1} = [u_{k-N_f}, u_{k-N_f+1}, \dots\, u_{k-1}] \tag{31}$$

Matrices H and L are derived from the state space description as given by equations (26) and (27) and length of the filter horizon N_f.

MODEL PREDICTIVE CONTROL

Main strategy of the MPC can be described as follows (Orukpe 2012), (Camacho and Bordons, 2004), (Kwon and Han, 2005):

- Step 1: Model is an explicit part of controller and it is used to predict future outputs $y(t)$. Predictions are calculated at each time step based on available information and the unknown future control signal course $u(t)$.
- Step 2: The sequence of future control signals is computed as a result of optimization of the performance criterion. Performance criterion consists of cost function and constraints. Cost functions include future output predictions, reference trajectory and future control signals. Constraints can be applied on both the output and the input of the process.
- Step 3: Only the first control action u_k is transmitted to the process. At the next sampling time, a new output value y_{k+1} is measured and whole sequences is repeated. This strategy is known as the receding horizon concept.

Cost function

The cost function (objective function) is necessary part of MPC. Cost function of linear models is mostly quadratic. Basic task is to minimize this function, which results in a sequence of input control samples. For better results, penalization constant can be added. The cost function is often expressed in matrices and is computed by simulation in the prediction horizon for all sequences of the manipulated variable and the manipulated variable sequence is calculated by a numerical algorithm such as gradient method or some even more sophisticated algorithms. Duration of the calculation is extended after addition of constraints (Haber et al., 2011), (Camacho and Bordons, 2004). Typical cost function look as follows:

$$J_k(N_d, N_y, N_u) = \sum_{j=N_d}^{N_y} \left(\hat{y}(k+j|k) - w(k+j)\right)^2 + \lambda \sum_{j=1}^{N_u-1} \Delta u(k+j-1) \tag{32}$$

where:

$\hat{y}(k+1|k)$ –sequence of future output values
$w(k+j)$ – sequence of desired values
$\Delta u(k+j-1)$ – sequence of differences of future control efforts

Cost function contains some optional parameters. Parameters N_d and N_y represents minimal, maximal prediction horizon and they determine the future interval, when a reference signal trajectory should have been followed. It is assumed that reference trajectory is known. That helps to make a timely reaction, before any changes have been effectively made. Parameter N_u represents control horizon, which does not necessarily have to coincide with the maximum prediction horizon. Lower value of N_u brings fewer computations. The coefficient λ represent a weight of the future control signal differences. All these parameters are assumed as tuning parameters (Haber et al., 2011).

Constraints

In practice, we often encounter with constraints. It can be constraints of sensor, actuator or some technological limitations. Usually input variables are constrained because they operate only in a certain range of values. In addition, there are also some constraints for the process output variables (with respect to the environment or the safety of workplaces). Ability to work with constraints is one of the main advantages of predictive control, which had an impact on MPC expansion in the industry, where large number of industrial processes is controlled to values close to restrictive conditions (Orukpe 2012).
Constraints can be categorized to hard constraints or soft constraints.
Hard constraints can never be exceeded (they must be satisfied). Typical examples are:

- constraint of control input variable: $u_{min} \leq u(k) \leq u_{max}$
- constraint of output variable: $y_{min} \leq y(k) \leq y_{max}$

Soft constraints are allowed to exceed the established limits on certain limit tolerance ε. Typical examples are:

- constraint of control input: $u_{min} + \varepsilon \leq u(k) \leq u_{max} + \varepsilon$
- constraint of output: $y_{min} + \varepsilon \leq y(k) \leq y_{max} + \varepsilon$

Control law

For the control of the CE152 magnetic levitation system a GPC algorithm was used. Prediction of output values is given by equation (33).

$$Y_k = Px_k + HU_k + Ld_k \qquad (33)$$

where parameter d_k represents compensation parameter, which gives value of prediction error and can be computed as:

$$d_k = y_k - CX_k \qquad (34)$$

The cost function (32) can be rewritten as follows:

$$J_k = E_k^T E_k + \Delta U_k^T \lambda \Delta U_k \qquad (35)$$

where E_k represents a vector of predicted future control errors. Sequence of optimal control actions is computed by solving equation (35). This can be done using MATLAB function *quadprog*.
The MATLAB Embedded function block in Simulink environment was used for implementation of the MPC controller. Control law is then computed in this function. Input values are current state estimates x_k, previous control action $u_{k\text{-}1}$, compensation constant d_k and reference trajectory represented by vector of values of desired future positions W_k. Output value is new action value u_k.

REAL TIME EXPERIMENTS

This section presents several real-time control experiments that were performed using CE 152 magnetic levitation plant. The following tuning parameters of the MPC were used for all presented experiments:

$$N_d = 1,\ N_y = 15,\ N_u = 15,\ N_f = 15,\ \lambda = 1 \quad (36)$$

Figure 5 presents control loop courses of the model predictive control of the CE 152 Magnetic Levitation plant:

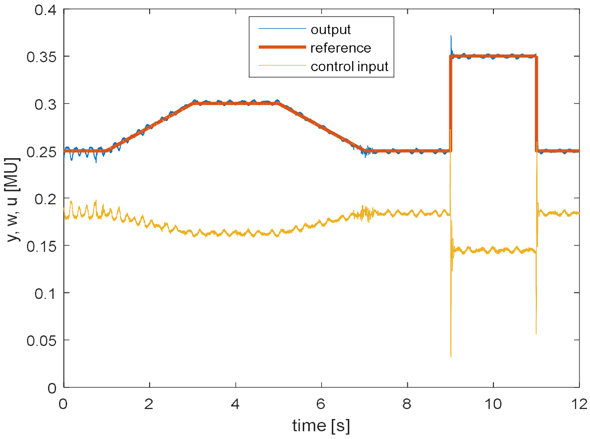


Figure 5: MPC control of the real plant

Comparison with simple controllers

Performance of the MPC controller was tested by comparing its control circuit courses with two PID-based controllers:

- PID-DEMO – pure PID controller with parameters defined as optimal by Humusoft – the manufacturer of the CE 152 plant.
- PID-Humusoft – advanced controller based on PID algorithm designed by the Humusoft company

All experiments were performed under the same conditions (same starting point). Figures 6 and 7 presents controlled outputs and control actions respectively. The course of reference trajectory contains only step changes.

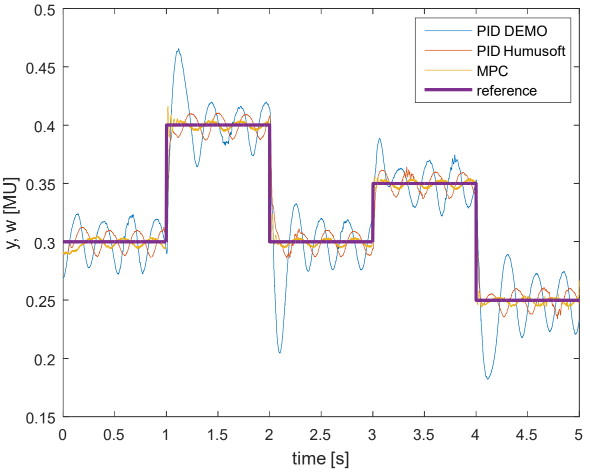


Figure 6: Comparison of outputs – step changes

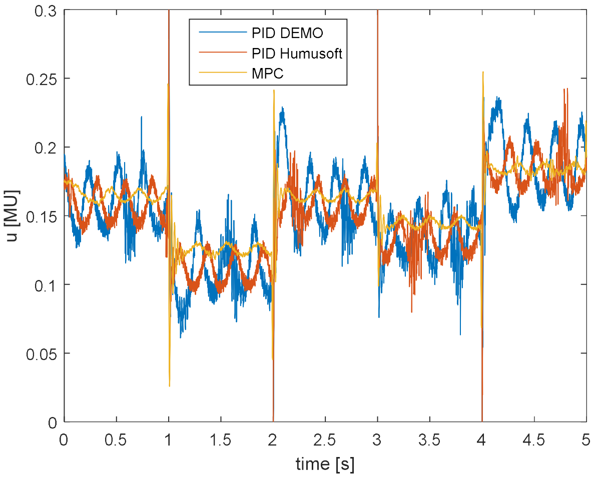


Figure 7: Comparison of control signals – step changes

Figures 8 and 9 presents comparison of the three controllers in case of ramp and step changes of reference signal.

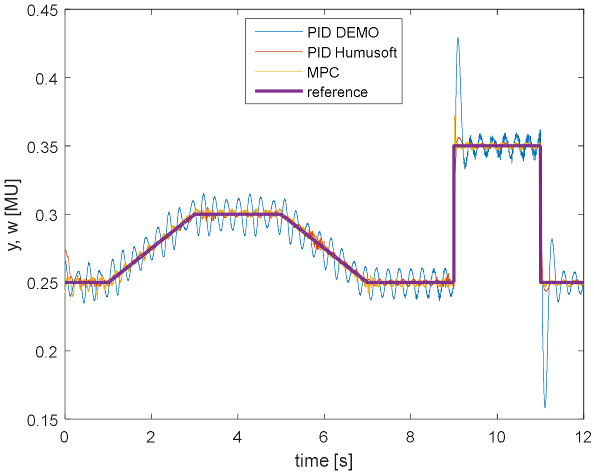


Figure 8: Comparison of outputs – steps and ramps

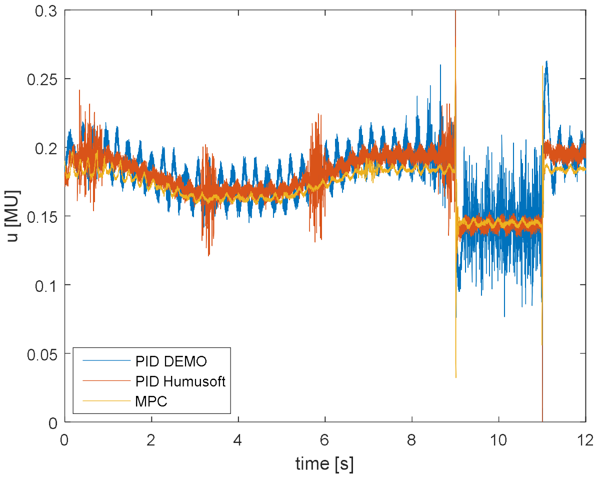


Figure 9: Comparison of control signals – steps and ramps

The control outputs presented in Fig. 7 and Fig. 9 are values produced directly by the controllers. These values are saturated to range <0 MU; 1 MU> by DA converter before the signal enters magnetic levitation laboratory plant. The limits of control signal were incorporated directly into MPC design and therefore it is ensured that MPC output is always within given limits. On the other hand, both PID produced control values out of the saturation range.

Criterions of control quality for both steps and ramp courses of reference signal are summarized in Table 1 where e_k represents control error in step k and $\Delta u_k = u_k - u_{k-1}$ represents control signal difference in step k. The sums presented in Table 1 are calculated over the whole time range presented in Fig 6 -9. As sample time T_0=1 ms was used for all experiments, all compared sums have the same number of summands.

Table 1: Control quality criterions

		PID DEMO	PID Humusoft	GPC
Steps (Fig. 6, 7)	$\sum_{k=1}^{N} e_k$	3.3705	0.7387	0.0992
	$\sum_{k=2}^{N} \Delta u_k$	1.7822	1.8641	0.0225
Ramps (Fig. 8, 9)	$\sum_{k=1}^{N} e_k$	2.2693	0.3836	0.0734
	$\sum_{k=2}^{N} \Delta u_k$	1. 3445	1.1175	0.0219

It can be seen that performance of MPC is significantly better then performance of classical PID based controllers.

CONCLUSION

The CE152 Magnetic levitation system was investigated and its first principle model was derived. Consequently, this model was linearized to obtain a model suitable for control design. A model predictive controller was derived using the state space model and verified by real-time control experiments.

The designed MPC was compared with two more simple PID-based controllers. Comparison of control courses led to the conclusion that the performance of the MPC is significantly better than the performance of the PID-based controllers. On the other hand, computational demands of the MPC are much higher comparing to PID controllers.

Further work will be focused on more detailed examination of the proposed MPC. Special attention will be paid to the role of tuning parameters.

ACKNOWLEDGEMENTS

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme project No. LO1303 (MSMT-7778/2014).

REFERENCES

Bobál, V.; J. Böhm; J. Fessl and J. Macháček. 2005. *Digital Self-tuning Controllers: Algorithms, Implementation and Applications.* Springer - Verlag London Ltd., London.

Camacho, E. and C. Bordons. 2004. *Model predictive control.* 2nd. ed., New York, Springer.

Chalupa, P.; M. Maly and J. Novak. 2016. "Nonlinear Simulink Model Of Magnetic Levitation Laboratory Plant", In: *ECMS 2016 Proceedings,* T. Claus, F. Herrmann, M. Manitz, O. Rose (Eds.), European Council for Modeling and Simulation. doi:10.7148/2016-0293.

Haber, R., R. Bars and U. Schmitz. 2011. *Predictive control in process engineering: From the basics to the applications*. Weinheim: Wiley-VCH

Kwon, W. and S. Han. 2005. *Receding horizon control: model predictive control for state models*. Springer, London.

Ljung, L. 1999. *System identification: theory for the user*. Upper Saddle River, N.J.: Prentice Hall PTR.

Himmelblau, D. M. and J. B. Riggs. 2004. *Basic principles and calculations in chemical engineering*, Upper Saddle River, N.J.: Prentice Hall.

Humusoft. 1996. *CE 152 Magnetic levitation model educational manual*

Orukpe, P. E. 2012. "Model Predictive Control Fundamentals". *Nigerian Journal of Technology (NIJOTECH)*. Vol. 41, No. 2, 139-148.

Tan, K. C. and Y. Li. 2002. "Grey-box model identification via evolutionary computing." *Control Engineering Practice*, 10, 673–684.

AUTHOR BIOGRAPHIES

Petr Chalupa was born in Zlin in 1976 and graduated from Brno University of Technology in 1999 and received the Ph.D. degree in Technical cybernetics from Tomas Bata University in Zlin in 2003.

He worked as a programmer and designer of an attendance system and as a developer of a wireless alarm system. He was a researcher at the Centre of Applied Cybernetics. Currently he works as a researcher at the Faculty of Applied Informatics, Tomas Bata University in Zlin as a member of CEBIA-Tech team. His research interests are adaptive and predictive control and modelling of real-time systems.

Jakub Novak was born in 1976 and received the Ph.D. degree in chemical and process engineering from Tomas Bata University in Zlin in 2007.

He is a researcher at the Faculty of Applied Informatics, Tomas Bata University in Zlin under a CEBIA-Tech project. His research interests are modeling and predictive control of the nonlinear systems.

Martin Malý graduated from Tomas Bata University in Zlin, Faculty of Applied Informatics in 2015. Nowadays he works as an engineer in TES Vsetin, Czech Republic

PREDICTIVE CONTROL OF A SERIES OF MULTIPLE LIQUID TANKS SUBSTITUTED BY A SINGLE DYNAMICS WITH TIME-DELAY

Stanislav Talaš, Vladimír Bobál, Adam Krhovják and Lukáš Rušar
Tomas Bata University in Zlin
Department of Process Control, Faculty of Applied Informatics
T. G. Masaryka 5555
760 01 Zlin
Czech Republic
E-mail: talas@fai.utb.cz

KEYWORDS
Time-delay, Predictive control, flow liquid level control.

ABSTRACT

The article focuses on control of a system consisting of a series of liquid tanks. Accumulation of individual dynamics causes, that the overall system exhibits high order behaviour. Another effect is a summation of slow responses of individual systems on an input signal leading to a significant time gap in reaction time of the whole system. In order to make control operations more straightforward and increase calculation speed, the mathematical description of gathered dynamics was approximated into a simplified form containing time-delay. The resulting form of the system is regulated by a predictive controller with time-delay compensation. The whole process is simulated in the Matlab environment.

INTRODUCTION

Systems created by a serial connection of individual elements are often used in chemical and petrochemical industry. Connection of multiple subsystems often creates a complicated dynamic behaviour described by high-order differential equations. In order to control systems of this type, it is appropriate to apply procedures that are able to handle a complex dynamic with a sufficient precision. Nevertheless, these techniques very often require a noteworthy amount of computing power to function with sufficient speed. In order to minimize this negative phenomenon, the mathematical description of the controlled systems tends to be simplified, in exchange for its precision. In such case, a time-delay effect may represent a slow initial system response and it can be used for a simplified description of the slow dynamic.

The computation complexity of model predictive control (MPC) was a subject of several studies as (Morari and Lee, 1998), (Angeli et al., 2012). Even systems with seemingly simple dynamics may cause a significant decrease in performance during on-line optimization. Basic countermeasures like lowering the length of the horizon or the number of variables often lead to a drop in quality of control. Moreover, they limit advanced functions that make use of MPC viable. Therefore, more elaborate techniques tend to be applied, such as an explicit predictive control (Kvasnica, 2010), faster optimization procedures (Wang et al., 2009) or just the simplification of mathematical description.

The utilization of time-delay effect in order to approximate the original complex behaviour with a simpler expression was a part of several studies (Richard, 2003), (Kubalčík and Bobál, 2012).

The article focusses attention on system formed by subsystems connected in series, which together create a complex dynamic. In this case, the representing subject of regulation is a set of eight liquid tanks. In order to achieve a sufficiently precise control, the mathematical description of the overall system is approximated into a simplified expression containing time-delay which replaces slow reactions accumulated from individual sections. The resulting form is then applied as a reference model for the predictive controller.

The article is organized as follows. The control technique of the predictive control is described in the first section, followed by an analysis of the series of liquid tanks. Next part describes the simplification process of the original system description. The Results section presents simulation outcomes of regulation processes.

METHODS

The predictive control principle is based on using an optimization to determine the most suitable system input. The core element is an internal model of the controlled process, from which estimates of future output values are predicted. The control algorithm searches for such a vector of input values that will cause the system output value to reach the reference state, while the change of the control input is minimal. In order to achieve this function, several parameters are established to define optimization properties. The whole search for the optimal outcome happens on a time interval from the present to a defined point in the future. This interval is called prediction horizon. The distance of the output signal from the reference value is considered in the area limited by the minimal horizon N_1 and the maximal horizon N_2. The change in the input signal $\Delta u(k)$ influences computations from the current

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

time to the control horizon N_u. Due to possible inaccuracies in calculations and presence of noise, optimization repeats in every sampling step to minimize the influence of errors. This repetitive approach called receding horizon strategy is visualized on Figure 1.

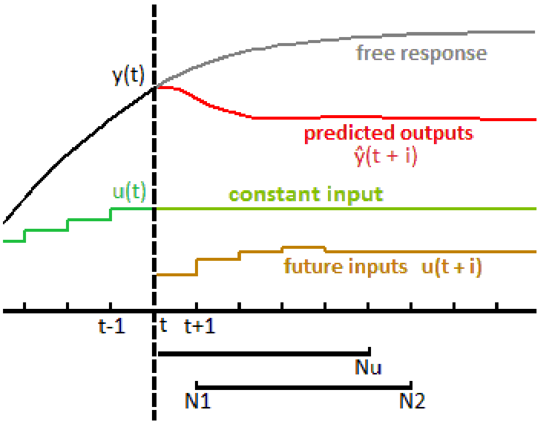


Figure 1: Receding horizon strategy

The estimation of the future outputs is based on the superposition principle and contains a sum of two calculated vectors. The first is a free response derived from the momentary system state and estimation that the input value will remain the same for the length of the horizon. The second is a forced response, a calculated outcome of the series of inputs suggested by the optimization applied on the internal model. Exchange of individual measurements and estimations is depicted in Figure 2.

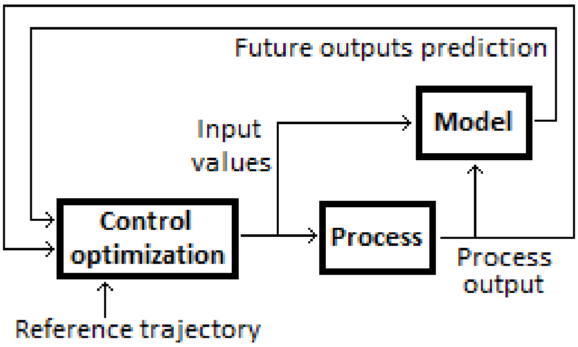


Figure 2: Basic structure of model predictive control

Conditions of the optimization procedure are stated in an objective function representing significance of individual signals involved in the control process. The algorithm aims to find a series of values $\Delta u(k)$ to $\Delta u(k+N_u-1)$ that results in the lowest result J.

$$J = \sum_{i=N_1}^{N_2} \delta(i)\left[\hat{y}(k+i) - w(k+i)\right]^2 + \sum_{i=1}^{N_u} \lambda(i)\left[\Delta u(k+i-1)\right]^2, \tag{1}$$

where $\delta(i)$ and $\lambda(i)$ are weighting values, usually constants representing a ratio of the minimization between a divergence of output from the desired value and a change of the action value.

In case of time delay the control algorithm is extended with shifting the computed interval limited by the minimal and the maximal horizon by the size of time-delay and system estimations necessary for calculation of the free-response compute future sampling steps up to the size of delay (Normey-Rico & Camacho, 2007).

System description

The controlled system is formed by eight identical water tanks connected in series, so the liquid flows directly from one tank to another. The goal of regulation is to control the height of water in the last tank by changing the inflow into the first tank.

Liquid levels are considered near an operating point and as such the physical relations between inflow, outflow and accumulation inside the tank can be described by the following equations:

$$F\frac{dh}{dt} = q_{IN} - q_{OUT} \tag{2}$$

$$q_{OUT} = K \cdot h \tag{3}$$

where F represents the surface area, h is the surface height, K is constant from the tank characteristics, q_{IN} is the flow of the liquid into the tank and q_{OUT} is the flow of the liquid out of the tank.

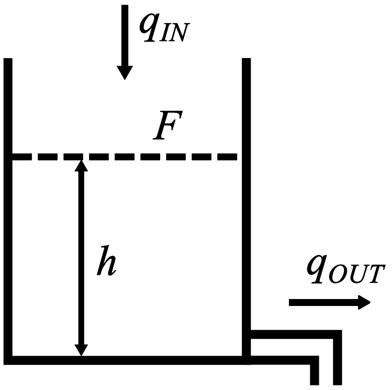


Figure 3: Illustration of a water tank

The transfer function describing the dependence of the surface height on the inflow is

$$h(s) = \frac{\frac{1}{K}}{\frac{F}{K}s+1} q_{IN}(s) \tag{4}$$

If we consider that the parameter K is identical for every tank, then based on the equation (3) it is possible to deduce the surface height from the following relation:

$$h_n = \frac{K_g}{\left(\frac{F}{K}s+1\right)^n} q_{IN}(s) = G(s) q_{IN}(s) \tag{5}$$

After performing of a substitution $1/K = K_g$ and $F/K = T$, the transfer function of n tanks is described as

$$G(s) = \frac{K_g}{(Ts+1)^n}. \tag{6}$$

Individual tanks are connected in such way that outflow from one tank is inflow into another.

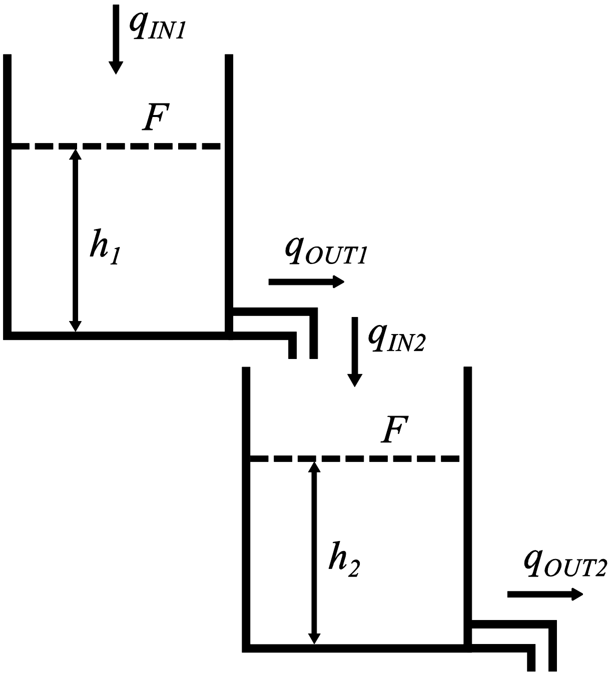


Figure 4: Illustration of two water tanks connected in series

In case of this experiment we consider the following:

Maximal tank height $h_{\max} = 1.5\ \mathrm{m}$,

tank diameter $d_T = 1\ \mathrm{m}$,

water surface area $F = \frac{\pi \cdot d_T}{4} = 0.785\ \mathrm{m}^2$,

time constant $T = 2\ \mathrm{min}$,

constant $K = \frac{F}{T} = 0.3925\ \mathrm{m}^2\ \mathrm{min}^{-1}$,

system gain $K_g = 1/K = 3.08\ \mathrm{m}^{-2}\ \mathrm{min}$.

And the system description is therefore given as

$$G(s) = \frac{3.08}{(2s+1)^8} \tag{7}$$

After a transformation of (7) into the numerical description, this system would have a general form expressed as

$$G(z^{-1}) = \frac{b_1 z^{-1} + b_2 z^{-2} + \ldots + b_8 z^{-8}}{1 + a_1 z^{-1} + \ldots + a_8 z^{-8}} \tag{8}$$

Applying an 8th order internal model into the predictive controller as in form (8) would increase the already abnormal computation time of the applied predictive controller (Bobál et al., 2016).

System approximation

In order to decrease the complexity of the system control, a simplification was performed to express its dynamic as a 2nd order system. Additionally, an accumulation of several low level dynamics may lead to a slow response of the overall system which may be interpreted as time-delay.

With aim to achieve a system description in the following form

$$G(z^{-1}) = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} z^{-d} \tag{9}$$

As an identification algorithm was selected the least square method (LSM) aiming for a system with a possible time delay value. This was achieved by repeating the identification, each time with a different value of time-delay. All system responses related to every time-delay in a defined interval were compared with the original 8th order system and the most precise outcome was determined by the integrated square error criterion.

The sampling period T_0 was set to 1 minute.

Based on results of the method the final system parameters were determined as

$$G(z^{-1}) = \frac{-0.01638 z^{-1} + 0.06371 z^{-2}}{1 - 1.815 z^{-1} + 0.8314 z^{-2}} z^{-5} \tag{10}$$

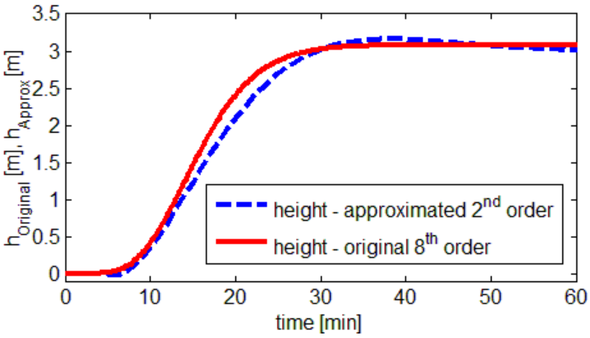


Figure 5: Comparison of transfer functions of original 8th order and approximated 2nd order systems

Figure 5 illustrates the similarity between the original 8th order system and the newly approximated 2nd order system. As can be seen the description received from the LSM manages to provide a system with a very similar behaviour with a slight difference in the area where the signal begins to settle. Furthermore, the approximated system exhibits more oscillating performance.

RESULTS

The approximated system was used as the internal model for the Generalized predictive controller, in order to provide estimated responses on the series of input signals.

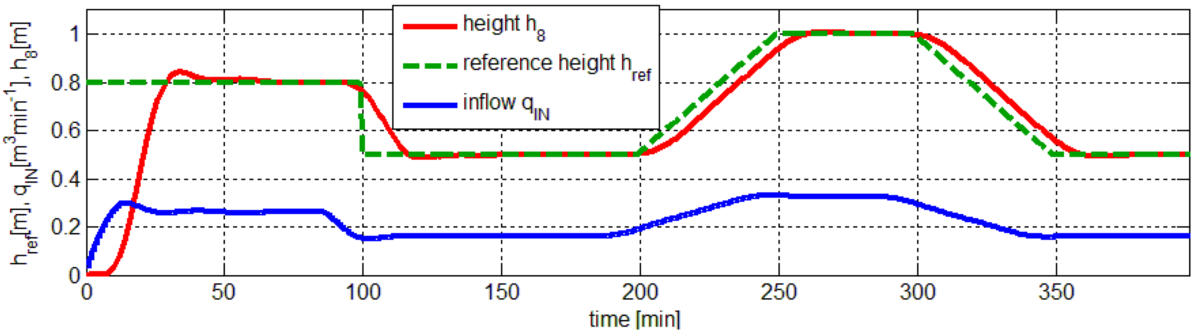


Figure 6: Predictive control of the 8th order process approximated by a 2nd order model

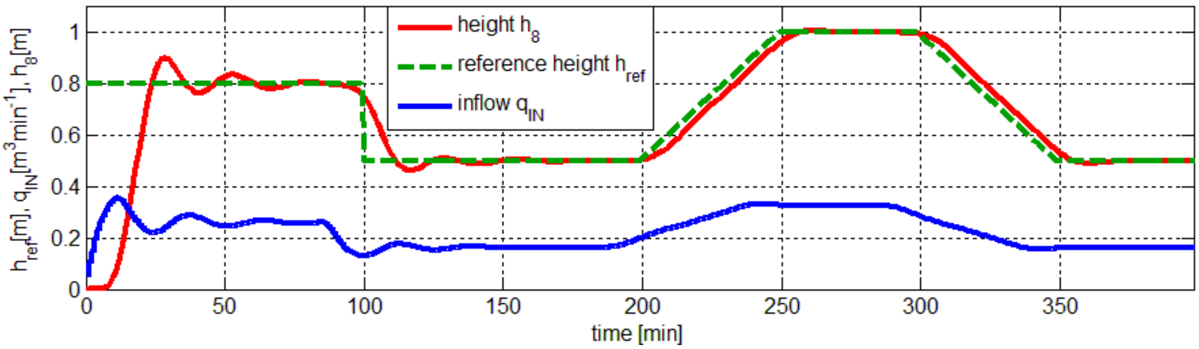


Figure 7: Predictive control of the 8th order process approximated by a 2nd order model with a decreased optimization of changes in the input signal

The reference trajectory was shaped to contain sudden steps, static section as well as gradual linear changes in positive and negative direction. Moreover, the duration of the simulation was set to enable stabilization after each change.

Figure 6 shows how the predictive controller was able to follow the reference trajectory with precision despite the simplified description of the controlled system.

Variability of the predictive control also offers an option of a faster transition in exchange for the precision of control. By decreasing the weight parameter for the optimization of changes in the input signal, the overall process gains a faster performance and it is able to follow ramp changes more closely, on the other hand the stabilization during step changes in the reference trajectory exhibit a significant increase in oscillations. Results of regulation altered in this way are illustrated in Figure7 and displays an improvement in areas of smaller but frequent changes of the desired value. The disadvantage of lost precision can be seen at around 30 minutes and 120 minutes of simulation time.

CONCLUSION

The paper presented an experiment where instead of a complex high-order system a simplified interpretation was performed, consequently decreasing demands for the control algorithm and computation time.

The original series of eight 1st order systems was tested by excitation signals. Based on gained responses, the whole process was identified as a 2nd order system. Furthermore, an accumulation of slow initial increases was replaced with time-delay effect. Consequently, this approximation managed to decrease computation demands of the applied predictive controller.

The control approach was verified by a subsequent simulation, which has demonstrated its applicability without significant errors in control.

Additional experiment has shown that it is possible to increase the speed of transitions in the output variable and its precision during linear changes by changing weighting parametrs.

A series of low-order systems, creating a single high-order system can be controlled as a low-order system with time-delay.

ACKNOWLEDGEMENT

This paper was supported by Internal Grant Agency of Tomas Bata University under the project No. IGA/FAI/2017/009.

REFERENCES

Angeli, D.; Amrit, R. and Rawlings, J. B. 2012. “On Average Performance and Stability of Economic

Model Predictive Control." *IEEE Transactions on Automatic Control*, **57** (7), 1615-1626.

Bobál, V.; Talaš, S. and Kubalčík, M. 2016. "LQ Digital Control of Coupled Liquid Level Equal Atmospheric Tanks – Design and Simulation." *WSEAS Transactions on Heat and Mass Transfer* (11), 62-71.

Kubalčík, M. and Bobál, V. 2012. "Predictive Control of Higher Order Systems Approximated by Lower Order Time-Delay Models." *WSEAS Transactions on Systems*, **10** (11), 607-616.

Kvasnica, M. 2010. *Real-Time Model Predictive Control via Multi-Parametric Programming.* Saarbrücken, VDM Verlag Dr. Müller Aktiengesellschaft & Co. KG.

Morari, M. and Lee, J. H. 1998. "Model predictive control: past, present, future." *Computers and Chemical Engineering*, **23**, 667-682.

Normey-Rico J. E. and E. F. Camacho 2007. *Control of dead-time processes.* London, Springer-Verlag.

Richard, J.P. 2003. "Time-delay systems: an overview of some recent advances and open problems." *Automatica*, **39**, 1667-1694.

Wang, Y. and Boyd, S. 2009. "Fast Model Predictive Control Using Online Optimization." *IEEE Transactions on Control Systems Technology*, **18** (2), 267-278.

AUHOR BIOGRAPHY

STANISLAV TALAŠ studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2013. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His e-mail address is talas@fai.utb.cz.

VLADIMÍR BOBÁL graduated in 1966 from the Brno University of Technology, Czech Republic. He received his Ph.D. degree in Technical Cybernetics at Institute of Technical Cybernetics, Slovak Academy of Sciences, Bratislava, Slovak Republic. He is now Professor at the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín, Czech Republic. His research interests are adaptive and predictive control, system identification and CAD for automatic control systems. You can contact him on email address bobal@fai.utb.cz.

ADAM KRHOVJÁK studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2013. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interests focus on modelling and simulation of continuous time technological processes, adaptive and nonlinear control. He is currently working on programming simulation library of technological systems.

LUKÁŠ RUŠAR studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2014. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interest focuses on model predictive control. His e-mail address is rusar@fai.utb.cz.

COMPENSATION OF VALVE DEADZONE USING MIXED INTEGER PREDICTIVE CONTROL

Jakub Novak and Petr Chalupa
Tomas Bata University, Faculty of Applied Informatics
Centre for Security, Information and Advanced Technologies
nam. TGM 5555, Zlín 76001, Czech Republic
E-mail: jnovak@fai.utb.cz

KEYWORDS
Predictive control, Mixed-integer quadratic programming, Valve deadzone, Hybrid system.

ABSTRACT

Stiction is a nonlinear friction phenomenon that causes poor performance of control loops in the process industries. In this work, we develop a mixed-integer MPC (Model Predictive Control) formulation including valve dynamics for a sticky valve in order to improve control loop performance. The introduction of the valve nonlinearity into the model prevents the MPC from requesting physically unrealistic control actions due to valve stiction. Simulation studies using a two-tank systems show that, if the deadband value is known a-priori, the Mixed-Integer Quadratic Programming (MIQP) can effectively improve the closed-loop performance in the presence of valve stiction.

INTRODUCTION

Model Predictive Control has become one of the most widespread modern control strategies which have been successfully applied in many industrial applications. The idea of MPC is to determine an optimal control at the current time instant by solving an optimal control problem on a prediction horizon. The main reason for the wide-scale adoption of MPC is its ability to handle constraints on inputs and states that arise in most applications. MPC also naturally handles the multidimensional systems.

Control valves are necessary elements in many chemical process control systems and are equipped for manipulating mass flows, energy flows or pressure. In general, they contain static and dynamic nonlinearities including saturation, backlash, stiction, deadband and hysteresis. If these nonlinearities are present, the valve is not capable of following the command signals provided by the controller thus leading to control performance degradation. A limit cycle is typically produced around the steady-state operating points if these nonlinearities are not included explicitly in the controller design. To compensate for valve malfunction several methods based on MPC has been developed. The framework developed in (Rodrıguez and Heath 2012) used the inverse of the nonlinearity in series with the original nonlinearity to overcome the problem. General compensation formulation that includes dynamics of sticky valve and additional constraints on inputs rate of change have been introduced in (Durand and Christofides 2016). A significant contribution in the area of valve nonlinearities compensation has been done in the work by (Zabiri and Samyudia 2006) which used a hybrid formulation of the input constraint within the MPC design to express the actuator backlash. The controller was used only in the only in the proximity of steady state operating points. Strategy that uses two-move stiction compensation have been revised in (Bacci di Capaci et al 2016) and successfully applied to the pilot plant.

In this paper, nonlinear model of the process with defined valve dead-zone was developed and used for optimization of the control signal via MPC strategy. The valve nonlinearity is expressed in terms of IF THEN conditions resulting in the hybrid model. Hybrid systems represent a unified framework for modelling such processes that combine continuous and discrete dynamic with logical rules and appear in various applications like robotic systems and automotives. A special class of hybrid systems called Mixed Logical Dynamical (MLD) systems has been introduced in (Bemporad and Morari 1999). Hybrid systems can effectively model a variety of systems: hybrid automata, nonlinear systems with the nonlinearity represented by the piece-wise affine functions, linear systems with constraints, etc. MPC is a general approach for control of such systems. However, the optimization problem is no longer quadratic programming (QP) problem but a Mixed-Integer Quadratic Programming problem. The inclusion of integer variables turns the easily solved QP problem, into an NP-hard problem (Borrelli et al 2006).

VALVE NONLINEARITY DESCRIPTION

A common process nonlinearity affecting the performance of control circuits with control valves (Fig. 1) is known as ‘stiction’ which exhibits a range of nonlinear behaviour including hysteresis, backslash and deadzones, both dynamic and static. In this paper, the model that includes deadband in every valve move was chosen. In Figure 1 the signal *u* is the process input, that is, the valve output, *y* is the process output, u_{MPC} is the MPC output, *w* and *v* are white Gaussian noises.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

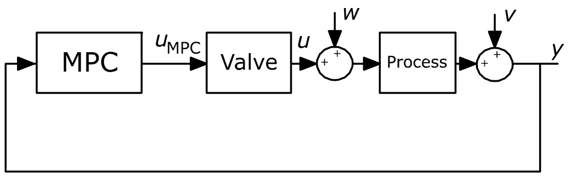


Figure 1: The Closed-loop System with the (sticky) Control Valve followed by the Process.

The sticky valve has a nonlinear dynamics expressed by the following relations:

$$u_i(k) = \begin{cases} u_{i,MPC}(k) : \triangle u_{i,MPC} > d \\ u_{i,MPC}(k) : \triangle u_{i,MPC} < -d \\ u_i(k-1) : otherwise \end{cases} \qquad (1)$$

The deadband model involves a set of logical rules that represent three regions of the deadband defined by parameter d (Figure 2).

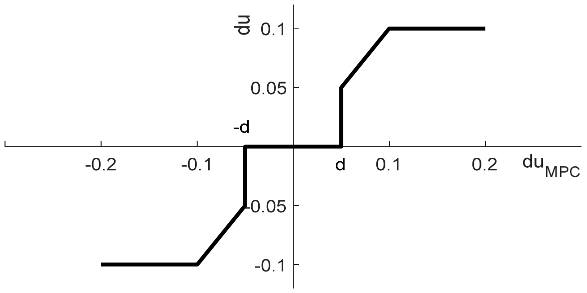


Figure 2: Sticky Valve Nonlinearity.

The input change $\triangle u_{i,MPC}(k)$ with respect to the deadband size d would determine what the output would be. It should be noted that the deadzone is active if $u(k) = u(k-1)$ i.e. the input signal is travelling within the deadband. The nonlinearity is formed by a set of three relatively simple linear relations, thus constituting a sort of switching multiple model scheme.

MODELING THE VALVE NONLINEARITY

The nonlinearity (1) of the valve can be modelled using a set of logical variables δ_{ij} :

$$\begin{aligned} &\delta_{i1}(k) \leftrightarrow \triangle u_i(k) \le -d_i \\ &\delta_{i2}(k) \leftrightarrow \triangle u_i(k) \ge d_i \\ &\delta_{i3}(k) \leftrightarrow \triangle u_i(k) = 0 \\ &u_{i,\min} \le u_i(k) \le u_{i,\max} \\ &\triangle u_{i,\min} \le \triangle u_i(k) \le \triangle u_{i,\max} \\ &\sum_{j=1}^{3} \delta_{ij}(k) = 1 \end{aligned} \qquad (2)$$

Propositional logic is translated into an equivalent linear inequalities using the strategy described in (Bemporad and Morari, 1999). For example, the first relation of (2) can be translated into two linear inequalities:

$$\begin{aligned} &\triangle u_i(k) + (M + d)\delta_{i1}(k) \le M \\ &\triangle u_i(k) - (m + d + \varepsilon)\delta_{i1}(k) \ge \varepsilon - d \end{aligned} \qquad (3)$$

where M,m are upper bounds and lower bounds of $\triangle u(k)$ and ε is a small positive scalar. This equivalence permits the assignment of binary variables to dynamical constraints which may define the different operation modes of hybrid system. The linear process dynamics is expressed by a state-space model

$$\begin{aligned} &\boldsymbol{x}(k+1) = \boldsymbol{A}_p x(k) + \boldsymbol{B}_p u(k) \\ &\boldsymbol{y}(k) = \boldsymbol{C}_p x(k) \end{aligned} \qquad (4)$$

where $\boldsymbol{A}_p \in R^{n*n}, \boldsymbol{B}_p \in R^{n*r}, \boldsymbol{C}_p \in R^{m*n}$ and n is the process order, r and m are number of input and outputs, respectively.

The resulting MLD system which includes dynamics of the process is described by the following relations:

$$\begin{aligned} \boldsymbol{x}(\mathrm{k}+1) &= \boldsymbol{A}x(\mathrm{k}) + \boldsymbol{B}_1 u(\mathrm{k}) + \boldsymbol{B}_2 \delta(\mathrm{k}) \\ &\quad + \boldsymbol{B}_3 z(\mathrm{k}) + \boldsymbol{B}_0 \\ \boldsymbol{y}(\mathrm{k}) &= \boldsymbol{C}x(\mathrm{k}) + \boldsymbol{D}_1 u(\mathrm{k}) + \boldsymbol{D}_2 \delta(\mathrm{k}) \\ &\quad + \boldsymbol{D}_3 z(\mathrm{k}) + \boldsymbol{D}_0 \\ \boldsymbol{E}_2 \delta(\mathrm{k}) + \boldsymbol{E}_3 z(\mathrm{k}) &\le +\boldsymbol{E}_1 u(\mathrm{k}) \\ &\quad + \boldsymbol{E}_4 x(\mathrm{k}) + \boldsymbol{E}_5 \end{aligned} \qquad (5)$$

where $A, B_1, B_2, B_3, B_0, C, D_1, D_2, D_3, D_0, E_1, E_2, E_3, E_4, E_5$ are matrixes of appropriate dimension and $z(k)$ are ancillary continuous variables.

PREDICTIVE CONTROL OF THE MLD SYSTEM

Bemporad and Morari introduced a model predictive control of hybrid systems using mixed logical dynamical (MLD) system description and a mixed integer linear program solver in (Bemporad and Morari, 1998). The quadratic objective function in case of control to a setpoint may be written:

$$\begin{aligned} \min J(u, \mathrm{z}, \delta) &= \left\| x(\mathrm{k}+N_P) - x_{ref} \right\|_S^2 + \\ &+ \sum_{j=0}^{N-1} \left(\left\| x(\mathrm{k}+j) - x_{ref} \right\|_Q^2 + \left\| \triangle u(\mathrm{k}+j) \right\|_R^2 \right) \end{aligned} \qquad (6)$$

Subjected to equations (5) which define the MLD system state predictions. In (6) the matrices ***Q,R,S*** indicate weighting matrices and y_{ref} is the reference signal and *Np* is the prediction horizon.

The problem (6) can be rewritten to a general MIQP programing form

$$\begin{aligned} & \min\ 0.5x^T \boldsymbol{Hx} + f^T x \\ & x = \begin{bmatrix} x_i \\ x_c \end{bmatrix} x_c \in R^{n_c}, x_i \in N^{n_i} \\ & \boldsymbol{a}_j^T \begin{bmatrix} \boldsymbol{x}_c^T \\ \boldsymbol{x}_i^T \end{bmatrix} + b_j = 0, \quad j = 1,...,m_{ec} \\ & \boldsymbol{c}_j^T \begin{bmatrix} \boldsymbol{x}_c^T \\ \boldsymbol{x}_i^T \end{bmatrix} + d_j \geq 0, \quad j = 1,...,m_{ic} \end{aligned} \tag{7}$$

where n_c and n_i define the numbers of continuous and integer variables, $\boldsymbol{H}$ is a positive definite matrix, $\boldsymbol{f}$ is the n-dimensional vector. The n-dimensional vectors a_j and c_j and vectors b and d are used to set up the linearity and nonlinearity constraints. The numbers of equality and inequality constraints are specified with m_{ec} and m_{ic}, respectively. The equality and inequality constraints define a feasible region in which the solution to the problem must be located in order for the constraints to be satisfied. The problem can be solved either by brute force or using the existing tools for MIQP programming based on branch-and-bound or branch-and-cut strategies. For MLD systems by definition of the vector

$$\boldsymbol{V}(k) = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_{n_i} \\ u(k) \\ \vdots \\ u(k+N-1) \\ z(k+1) \\ \vdots \\ z(k+N) \end{bmatrix} \tag{8}$$

the problem (7) can be rewritten in the compact form:

$$\begin{aligned} & \min\ 0.5\boldsymbol{V}(k)^T \boldsymbol{HV(k)} + \boldsymbol{f}^T \boldsymbol{V}(k) \\ & \boldsymbol{A}_c \boldsymbol{V}(k) \leq \boldsymbol{b}_c \\ & \boldsymbol{C}_c \boldsymbol{V}(k) = \boldsymbol{d}_c \end{aligned} \tag{9}$$

The MPC algorithm in the paper uses incremental form which is insensitive to slowly varying system and measurement trends and therefore has integral action (Di Ruscio 2013). Given the process model

$$\begin{aligned} x(k+1) &= \boldsymbol{A}x(k) + \boldsymbol{B}u(k) + v \\ y(k) &= \boldsymbol{C}x(k) + w \end{aligned} \tag{10}$$

where $x(k) \in R^n$ is the state vector, $u(k) \in R^r$ is the control input vector, $y(k) \in R^m$ is the output vector, $\boldsymbol{A},\boldsymbol{B}$ and $\boldsymbol{C}$ are system matrices of appropriate dimensions. The augmented model which is independent of the unknown disturbances is then given

$$\begin{bmatrix} \triangle\boldsymbol{x}(k+1) \\ \boldsymbol{y}(k) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & 0 \\ \boldsymbol{C} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \triangle\boldsymbol{x}(k) \\ \boldsymbol{y}(k-1) \end{bmatrix} + \begin{bmatrix} \boldsymbol{B} \\ 0 \end{bmatrix} \triangle\boldsymbol{u}(k) \tag{11}$$

SIMULATION EXAMPLE

In other to evaluate the presented approach the simulation example of the two-tank system is used. Figure 3 shows the scheme of the system.

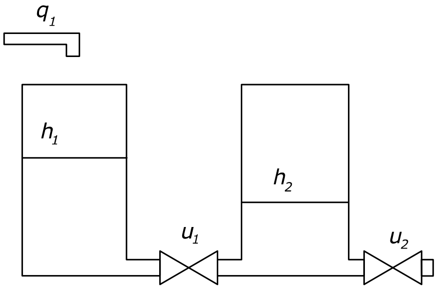


Figure 3: Two-tank System Scheme

The changes of water levels can be described by the set of differential balance equations of the two tanks:

$$\begin{aligned} S_T \frac{dh_1}{dt} &= q_1 - u_1 S_v \sqrt{2g}\sqrt{h_1 - h_2} \\ S_T \frac{dh_2}{dt} &= u_1 S_v \sqrt{2g}\sqrt{h_1 - h_2} - u_2 S_v \sqrt{2g}\sqrt{h_2} \end{aligned} \tag{12}$$

where S_T is cross-section area of the tanks
S_v is cross-section area of the valves
g is gravity constant
q_1 is volume flow through pump
u_1 is inputs signal to valve 1
u_2 is inputs signal to valve 2
h_1 is water level in tank 1
h_2 is water level in tank 2.

The parameters of the simulation system (Table 1) were taken from (Chalupa and Novak 2013).

Table 1 Parameters of the System

Par.	Value
S_T	0.0154m^2
S_v	5e-5m^2
g	9.81m/s^2
q_1	1e-4m^3/s

The levels in both tanks are controlled using two control valves. The second control valve u_2 has a deadband nonlinearity described by the relations (1) of the size $d = 0.05$. There are also limits on both control signals and control increments:

$$\begin{aligned} 0 \le u(k) \le 1 \\ -0.1 \le \Delta u(k) \le 0.1 \end{aligned} \qquad (13)$$

The system was linearized with a sampling time of 20s in steady-state conditions $h_1 = 0.3\text{m}, h_2 = 0.15\text{m}$. The performance of the MPC control with (*Np*=10, *Q=I, R*=0, *S*=100*I*) that is unaware of the nonlinearity is presented in Figure 4 and 5.

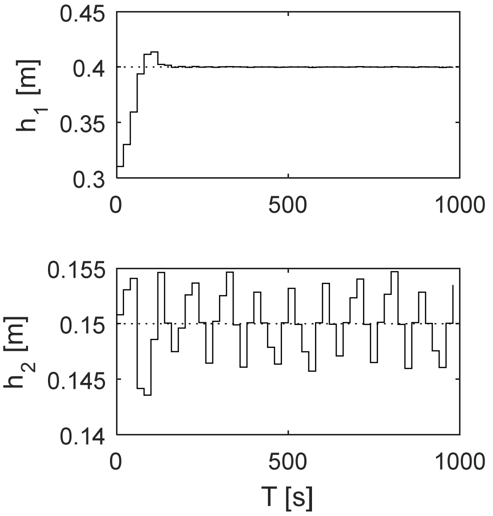


Figure 4: Output Response for Step-change of the Level in the First Tank (dotted line - reference)

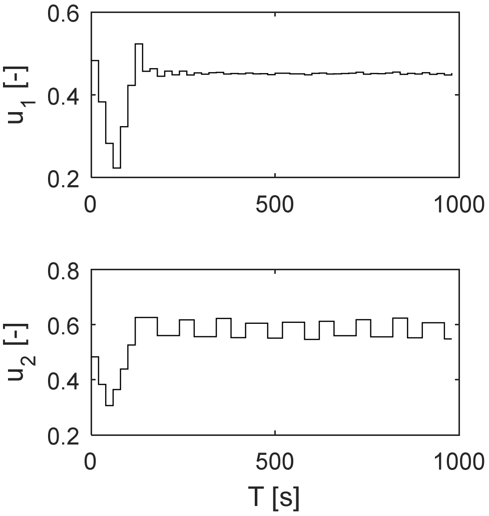


Figure 5: Controller Output Response for Step-change of the Level in the First Tank

The nonlinearities have negative effects on set-point tracking and result in oscillations in control loops. When a valve has deadband, the valve output does not change in response to changes in the control signal to the valve until the control signal overcomes the deadband.

In order to improve the performance the nonlinearity (1) is translated into the set linear inequalities using (2) that constrain the optimization problem and introduce binary variables. Thus originally quadratic optimization problem with continuous variables is replaced by a mixed-integer quadratic optimization problem. In order to reduce the number of binary variables the control horizon *Np* is reduced to 3 steps which represents 9 binary variables δ. The resulting quadratic problem has *Np**3 binary variables and *Np*r+N*(m+n)* continuous variables (inputs and predicted states of the augmented model). The problem is constrained by limits for input variables, binary variables δ and relations defining the evolution of the predicted states. The problem is solved using the branch-and-bound strategy and interior point method is used for solution of the relaxed problems (Novak and Chalupa, 2015). Using the successive linearization strategy the parameters of augmented model are obtained at each sampling period using the current state $x(k)$ and linearization of (12). The steady-state target for each step of the reference signal is given:

$$y_{ref} = \begin{bmatrix} 0 \\ 0 \\ h_{1ref} \\ h_{2ref} \end{bmatrix} \qquad (14)$$

The resulting control courses for step-changes of both outputs with the same weighting matrices are presented in Figure 6 and 7.

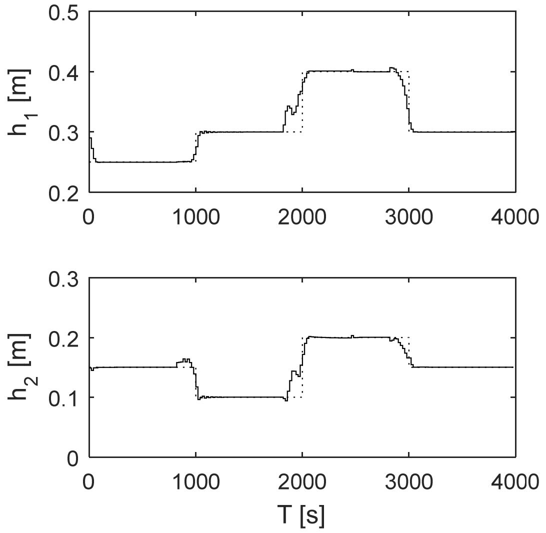


Figure 6: Output Response for MIQP-based MPC Formulation (dotted line - reference)

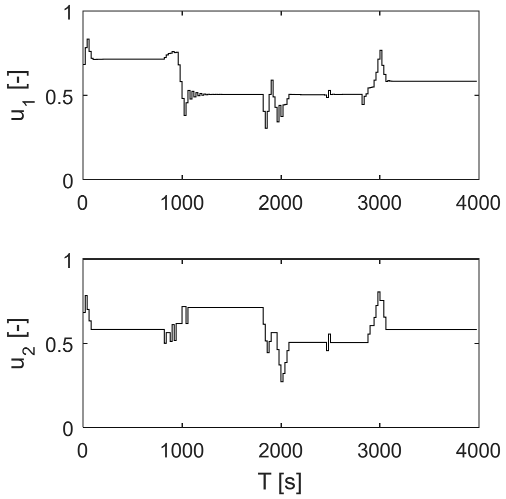


Figure 7: Input Response for MIQP-based MPC Formulation

It can be observed that stiction embedding MPC guarantees very good tracking performance and also an effective stiction compensation. The number of relaxed quadratic problems that are solved during solution of MIQP problem are presented in Figure 8. Solution of the MIQP problem with 9 binary variables δ by brute-force would require solution of 2^9 relaxed QP problems. Because of the efficiency of the branch-and-bound algorithm less than 40 QP problems are solved at each sampling point.

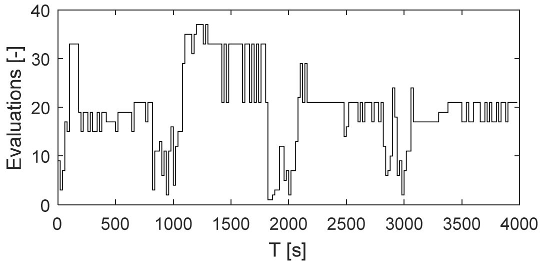


Figure 8: Number of Evaluations of the Relaxed QP Problems

CONCLUSION

This paper has presented a possible formulation of MPC to face deadband nonlinearity in control valves for industrial processes. The improved performance of MIQP-based MPC controller is possible at the expense of using more complex and computationally more demanding mixed-integer optimization algorithm. However, the MIQP-based MPC requires the knowledge of the deadband a-priori.

ACKNOWLEDGMENT

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme project No. LO1303 (MSMT-7778/2014).

REFERENCES

Bacci di Capaci, R.; Scali, C.; and B. Huang. 2016. “A revised technique of stiction compensation for control valves”. In *Proceedings of the 11th IFAC DYCOPS*, (Trondheim, Norway) 1038–1043.

Bemporad, A. and M. Morari. 1999. “Control of systems integrating logic, dynamics, and constraints”. *Automatica*, Vol. 35, 407-427.

Bemporad, A. and M. Morari. 1998. “Predictive control of constrained hybrid systems”. In *Proceedings of Int. Symposium on Nonlinear Model Predictive Control*, (Ascona, Switzerland), 108-127.

Borrelli, F.; A. Bemporad; M. Fodor; D. Hrovat, 2006. “An MPC/hybrid system approach to traction control”, *IEEE Transactions on Control Systems Technology* 14, 541-552.

Chalupa, P. and J. Novak. 2013. “Modeling and Model Predictive Control of Nonlinear Hydraulic System”. *Computers & Mathematics with Applications* ,66(2), 155–164.

Di Ruscio, D. 2013. “Model Predictive Control with Integral Action: A simple MPC algorithm”. *Modeling, Identification and Control*, Vol. 34, No. 3, 119-129.

Durand, H. and P. Christofides, 2016. “Actuator stiction compensation via model predictive control for nonlinear processes”. *AIChE Journal*, Vol. 62, 2004–2023.

Novak, J. and P. Chalupa. 2015. “Implementation of Mixed-integer Programming on Embedded System”. *Procedia Engineering*, Vol. 100, 1649-1656.

Rodrıguez, M. and W. Heath. 2012. “MPC for plants subject to saturation and deadzone, backlash or stiction”. In *Proceedings of the 4th IFAC Nonlinear Model Predictive Control Conference* (Noordwijkerhout, The Netherlands). 418–423.

Zabiri, H., and Y. Samyudia. 2006. “A hybrid formulation and design of model predictive control for systems under actuator saturation and backlash”. *Journal of Procces Control*, Vol 16, 693-709.

AUTHOR BIOGRAPHIES

JAKUB NOVAK was born in Zlin, Czech Republic and received the Ph.D. degree from the Tomas Bata University, Zlin, Czech Republic in 2007. He is now a researcher at the CEBIA-TECH research center at Faculty of Applied Informatics at the university. His research interests are multiple model strategies and predictive control. His e-mail address is : jnovak@fai.utb.cz

PETR CHALUPA was born in Zlin, Czech Republic in 1976. He graduated from Brno University of Technology in 1999. He obtained his Ph.D. in Technical Cybernetics at Tomas Bata University in Zlin in 2003. He works as a researcher at CEBIA-TECH research center at Tomas Bata University in Zlin. His research interests are adaptive and predictive control of real-time systems. You can contact him on email address chalupa@fai.utb.cz

STATE-SPACE PREDICTIVE CONTROL OF INVERTED PENDULUM MODEL

Lukáš Rušar, Adam Krhovják, Stanislav Talaš and Vladimír Bobál
Department of process control
Faculty of applied informatics, Tomas Bata university in Zlin
Nad Stráněmi 4511, Zlin 76005, Czech Republic
E-mail: rusar@fai.utb.cz

KEYWORDS
predictive control, state-space, inverted pendulum, predictor-corrector.

ABSTRACT

This paper presents a possible way to control the a very fast nonlinear systems. The system of the inverted pendulum was chosen as an exemplar process. This is an example of the nonlinear single-input multi-output process with a sampling period in order of milliseconds. The state-space predictive control was chosen as a control method and the system is described by CARIMA model. The whole process of the controller design is described in this paper. That includes a description of the inverted pendulum nonlinear mathematical model and its linearization, the inference of the output values prediction and the control signal calculation. The control signal is calculated by predictor-corrector method. The results compare several optimization methods to achieve the fastest calculation of the control signal. All of the simulation was done in Matlab.

INTRODUCTION

In real life we can come across with many types of processes. Many of them are nonlinear and their mathematical models are very complex. Even the sampling period can be very different. This paper focuses on the very fast processes with a sampling period in the order of milliseconds. The basic control methods may not handle with this situation with required precision so we need a more advanced method. The predictive control is a great example of the modern control method that can be used to solve the complex control problems (Bobál 2008).

This method belongs to the model based control methods and the mathematical model is used for the output values prediction. This prediction is determine on the chosen time horizon that should be long enough to cover the step response of the controlled system. The model of the inverted pendulum is described by the state-space CARIMA mathematical model for the single-input multi-output (SIMO) system (Bars et al. 2011; Wang 2009).

The control signal calculated by the predictive control ensures the desired output values in the near future time horizon. This is achieved by minimization of the cost function that usually has a quadratic form and it minimize the differences between the reference value and the output value and the control signal increments. If the process require some kind of the process variable constraints, several method such as quadratic programming method, fast-gradient method, predictor-corrector method etc. can be used to minimize the cost function (Camacho and Bordons 2004; Maciejowski 2002; Rossiter 2003).

However, the chosen CARIMA mathematical model used to the prediction of the output values works only for the linear models so the nonlinear mathematical model of the inverted pendulum needs to be linearized.

This paper is divided into the following sections. The model of the inverted pendulum is described in the first section. The predictive control and the calculation of the control signal are described next. The final sections shows the results of the research and the conclusion (Albertos Peréz and Sala 20014; Hangos et al. 2004).

MATHEMATICAL MODEL OF THE CONTROLLED SYSTEM

The Amira PS600 inverted pendulum system was used as the exemplar model. The photo of this system is shown at figure 1. The main parts of the system are cart driven by servo amplifier and the pendulum rod attached to the cart (Amira 2000; Chalupa and Bobál 2008).

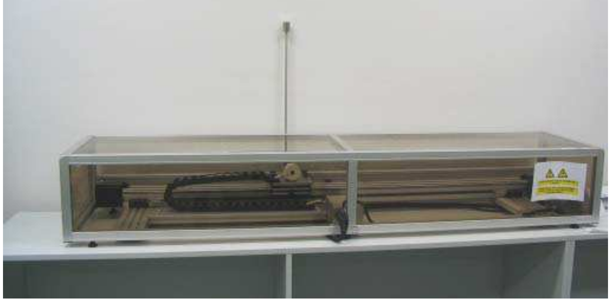


Figure 1 : Amira PS600 Inverted Pendulum system

The inverted pendulum system is an example of the single-input two-output system. The force produced by the DC motor that moves with the cart is the input variable and the cart position and the angle of the pendulum rod are the output variables. The figure 2 shows the analysis of the forces acting in the system (Amira 2000; Chalupa and Bobál 2008).

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

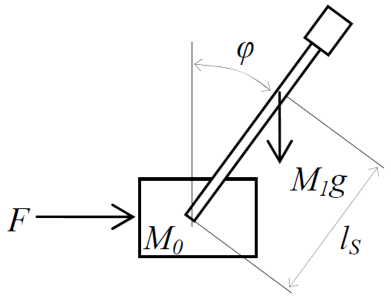


Figure 2 : Analysis of the inverted pendulum

The variables in the figure 2 are following. The angle of pendulum rod is φ, M_0 and M_1 stands for the weight of the cart and pendulum respectively, *ls* is a distance between centre of gravity of the pendulum and the centre of rotation of the pendulum and *g* is the gravity acceleration constant. Symbol *F* represents the force produced by the DC motor.

The affect of the pendulum on the cart can be expressed as a horizontal and a vertical forces described by the equations (1) and (2)

$$H = M_1 \frac{d^2\left(r + l_s \sin\varphi\right)}{dt^2} \tag{1}$$

$$V = M_1 \frac{d^2\left(l_s \cos\varphi\right)}{dt^2} \tag{2}$$

where *r* is the position of the cart.

The equation (3) describe a motion equation of the cart.

$$M_0 \frac{d^2 r}{dt^2} = F - H - F_r \frac{dr}{dt} \tag{3}$$

where F_r is the constant of a velocity proportional friction of the cart. The rotary motion of the rod about its centre is derived according to the angular conversation law and described by the equation (4).

$$\Theta_s \frac{d^2\varphi}{dt^2} = V l_s \sin\varphi - H l_s \cos\varphi - C\frac{d\varphi}{dt} \tag{4}$$

where Θ_s represents the inertia moment of the pendulum rod with respect to the centre of gravity and *C* denotes the friction constant of the pendulum.

If we substitute the equations (1) and (2) into the equations (3) and (4) we get the nonlinear equations (5) and (6) describing the behavior of the inverted pendulum system.

$$Mr'' + F_r r' + M_1 l_s \varphi'' \cos\varphi - M_1 l_s \left(\varphi'\right)^2 \sin\varphi = F \tag{5}$$

$$\Theta\varphi'' + C\varphi' - M_1 l_s g \sin\varphi + M_1 l_s r'' \cos\varphi = 0 \tag{6}$$

where following abbreviations were used:

$$\Theta = \Theta_s + M_1 l_s^2 \tag{7}$$

$$M = M_0 + M_1 \tag{8}$$

The nonlinear state-space model of this system can be obtain by choosing the state vector as shown in the equation (9).

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} r \\ r' \\ \varphi \\ \varphi' \end{bmatrix} \tag{9}$$

The dynamics of the system in the state-space model representation can be described as the equation (10)

$$\boldsymbol{x}' = \begin{bmatrix} x_1' \\ x_2' \\ x_3' \\ x_4' \end{bmatrix} = \begin{bmatrix} x_2 \\ f_1(\boldsymbol{x}.u) \\ x_4 \\ f_2(\boldsymbol{x}.u) \end{bmatrix} \tag{10}$$

where the functions f_1 and f_2 are derived from the equations (5) and (6).

$$f_1(\boldsymbol{x}.u) = \frac{1}{M_1^2 l_s^2 \cos^2 x_3 - M\Theta}\Big[-CM_1 l_s x_4 \cos x_3 + \\ +M_1^2 l_s^2 g \sin x_3 \cos x_3 + \Theta F_r x_2 - \Theta M_1 l_s x_4^2 \sin x_3 - \Theta u\Big] \tag{11}$$

$$f_2(\boldsymbol{x}.u) = \frac{1}{M_1^2 l_s^2 \cos^2 x_3 - M\Theta}\Big[MCx_4 - MM_1 l_s g \sin x_3 - \\ -M_1 l_s F_r x_2 \cos x_3 + M_1^2 l_s^2 x_4^2 \sin x_3 \cos x_3 + M_1 l_s u \cos x_3\Big] \tag{12}$$

The cart position *r* and the angle of the pendulum φ are the output variables (Chalupa and Bobál 2008).

$$\boldsymbol{y} = \begin{bmatrix} x_1 \\ x_3 \end{bmatrix}\begin{bmatrix} r \\ \varphi \end{bmatrix} \tag{13}$$

The described nonlinear model has to be linearized around some operating point. The linearization about the operating point means substitution of the absolute value of the input, output and state variables by its divergence from the steady state.

$$\begin{aligned} \boldsymbol{x}_\delta(t) &= \boldsymbol{x}(t) - \overline{\boldsymbol{x}} \\ \boldsymbol{u}_\delta(t) &= \boldsymbol{u}(t) - \overline{\boldsymbol{u}} \\ \boldsymbol{y}_\delta(t) &= \boldsymbol{y}(t) - \overline{\boldsymbol{y}} \end{aligned} \tag{14}$$

Where $\overline{\boldsymbol{x}}$ is a vector of the equilibrium state variables, $\overline{\boldsymbol{u}}$ is a vector of the equilibrium input variables, $\overline{\boldsymbol{y}}$ is a vector of the equilibrium output variables, $\boldsymbol{x}_\delta, \boldsymbol{u}_\delta, \boldsymbol{y}_\delta$ are divergences from equilibrium values.

The linearized state-space model can be expressed in form

$$\begin{aligned} \dot{\boldsymbol{x}}_\delta &= \boldsymbol{A}\boldsymbol{x}_\delta + \boldsymbol{B}\boldsymbol{u}_\delta \\ \boldsymbol{y}_\delta &= \boldsymbol{C}\boldsymbol{x}_\delta \end{aligned} \tag{15}$$

where matrices ***A***, ***B*** are

$$\boldsymbol{A} = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}=\overline{\boldsymbol{x}},\boldsymbol{u}=\overline{\boldsymbol{u}}} \\ \boldsymbol{B} = \left.\frac{\partial f}{\partial \boldsymbol{u}}\right|_{\boldsymbol{x}=\overline{\boldsymbol{x}},\boldsymbol{u}=\overline{\boldsymbol{u}}} \tag{16}$$

The precise form of the ***A***, ***B*** and ***C*** matrices are shown in the equation (17)

$$
\boldsymbol{A}=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & a_{22} & a_{23} & a_{24}\\ 0 & 0 & 0 & 1\\ 0 & a_{42} & a_{43} & a_{44}\end{bmatrix}
$$

$$
\boldsymbol{B}=\begin{bmatrix}0\\ b_2\\ 0\\ b_4\end{bmatrix}
$$

$$
\boldsymbol{C}=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix} \tag{17}
$$

where the inner components of the matrices are derived for the equilibrium point when the pendulum is in the upright position (Chalupa and Bobál 2008).

$$
a_{22}=\frac{\Theta F_r}{R}, a_{23}=\frac{M_1^2 l_s^2 g}{R}, a_{24}=\frac{-CM_1 l_s}{R}
$$

$$
a_{42}=\frac{-M_1 l_s F_r}{R}, a_{43}=\frac{-MM_1 l_s g}{R}, a_{44}=\frac{CM}{R}
$$

$$
b_2=\frac{-\Theta}{R}, b_4=\frac{M_1 l_s}{R}, R=M_1^2 l_s^2-M\Theta \tag{18}
$$

However, this is still a continuous-time model and it needs to be transferred into a discrete-time form suitable for the chosen predictive control method. It can be done by transferring the state-space model into the input-output model

$$
A(s)\,y(t)=B(s)\,u(t) \tag{19}
$$

and then into its discrete representation

$$
\tilde{A}(z^{-1})\,y(k)=B(z^{-1})\,\Delta u(k) \tag{20}
$$

where the polynom $\tilde{A}(z^{-1})$ is

$$
\tilde{A}(z^{-1})=(1-z^{-1})\,A(z^{-1}) \tag{21}
$$

STATE-SPACE PREDICTIVE CONTROL

The chosen predictive control method uses the state-space CARIMA (Controlled Auto-Regresive and Integrated Moving Average) model for prediction of the output values. This model is described by equation (22).

$$
\begin{aligned}\boldsymbol{x}(k+1)&=\tilde{\boldsymbol{A}}\boldsymbol{x}(k)+\boldsymbol{B}\Delta u(k)\\ y(k)&=\boldsymbol{C}\boldsymbol{x}(k)\end{aligned} \tag{22}
$$

Where the vector of state variables has form

$$
\begin{aligned}\boldsymbol{x}(k)=[&y(k),y(k-1),\cdots,y(k-na),\\ &\Delta u(k-1),\cdots,\Delta u(k-nb+1)]^T\end{aligned} \tag{23}
$$

The matrices $\tilde{\boldsymbol{A}}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ from the model (22) can be expressed as

$$
\tilde{\boldsymbol{A}}=\begin{bmatrix}-\tilde{a}_1 & \cdots & -\tilde{a}_{na} & -\tilde{a}_{na+1} & b_2 & \cdots & b_{nb-1} & b_{nb}\\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0\\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0\\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0\\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0\\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0\end{bmatrix}
$$

$$
\boldsymbol{B}=\begin{bmatrix}b_1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0\end{bmatrix}^T
$$

$$
\boldsymbol{C}=\begin{bmatrix}1 & 0 & \cdots & 0 & 0\end{bmatrix} \tag{24}
$$

The values $-\tilde{a}_i$ for $i=1,\ldots,na+1$ and b_j for $j=1,\ldots,nb$ consist of the coefficients of the polynoms $\tilde{A}(z^{-1})$ and $B(z^{-1})$ from the equation (21) (Bars et al. 2011; Camacho and Bordons 2004).

The prediction of the output values is obtained recursively using the CARIMA model represented by equation (22). The final matrix form of this prediction is

$$
\hat{\boldsymbol{y}}=\boldsymbol{F}\boldsymbol{x}+\boldsymbol{H}_f\Delta\boldsymbol{u}_f \tag{25}
$$

where $\hat{\boldsymbol{y}}$ is the vector of the predicted output values and $\Delta\boldsymbol{u}_f$ is the vector of the future control increments

$$
\hat{\boldsymbol{y}}=\begin{bmatrix}\hat{y}(k+1)\\ \hat{y}(k+2)\\ \vdots\\ \hat{y}(k+N)\end{bmatrix}
$$

$$
\Delta\boldsymbol{u}_f=\begin{bmatrix}\Delta u(k)\\ \Delta u(k+1)\\ \vdots\\ \Delta u(k+N)\end{bmatrix} \tag{26}
$$

where N is the chosen time horizon for prediction.

The aim of the predictive control is minimize the difference between the future reference values and the predicted output values and also minimize the control signal demand. The quadratic cost function in the equation (27) is used to this optimization problem.

$$
J=(\boldsymbol{w}-\hat{\boldsymbol{y}})^T\boldsymbol{Q}_\delta(\boldsymbol{w}-\hat{\boldsymbol{y}})+\Delta\boldsymbol{u}_f^T\boldsymbol{Q}_\lambda\Delta\boldsymbol{u}_f \tag{27}
$$

where $\boldsymbol{w}$ is a vector of the future reference values, $\hat{\boldsymbol{y}}$ is the vector of the predicted outputs values, $\boldsymbol{Q}_\lambda$ and $\boldsymbol{Q}_\delta$ are the diagonal weighting matrices containing the weighting coeffitients λ and δ. The vector $\Delta\boldsymbol{u}_f$ is unknown vector of the future control increments (Camacho and Bordons 2004; Fikar and Mikleš 2008).

Because of the chosen optimization method, this cost function needs to be modified into the form of the equation (28).

$$
J=\frac{1}{2}\boldsymbol{u}^T\boldsymbol{H}_c\boldsymbol{u}+\boldsymbol{g}^T\boldsymbol{u} \tag{28}
$$

where

$$\begin{aligned} \boldsymbol{H}_c &= 2\left(\boldsymbol{Q}_\lambda + \boldsymbol{H}_f^T \boldsymbol{Q}_\delta \boldsymbol{H}_f\right) \\ \boldsymbol{g}^T &= 2\left(\boldsymbol{F}\boldsymbol{x} - \boldsymbol{w}\right)^T \boldsymbol{Q}_\delta \boldsymbol{H}_f \end{aligned} \tag{29}$$

PREDICTOR-CORRECTOR METHOD

The predictor-corrector method is one of the primal-dual interior-point methods using to solve the inequality constrained convex quadratic problems

$$\begin{aligned} f(x) &= \frac{1}{2}\boldsymbol{x}^T \boldsymbol{G}\boldsymbol{x} + \boldsymbol{g}^T \boldsymbol{x} \\ \boldsymbol{A}^T \boldsymbol{x} &\geq \boldsymbol{b} \end{aligned} \tag{30}$$

which is exactly the problem that the predictive control solves. The equation (30) represents the general formulation of the constrained quadratic problem. The aim is to find the unknown vector $\boldsymbol{x}$ with respect to the chosen constrains representing the future values of the control signal increments (Nocedal and Wright 2000; Wright 1997).

This is the iterative method and we have to set the starting points of the unknown vector $\boldsymbol{x}_0$, the vector of the Lagrange multipliers $\boldsymbol{\lambda}_0$ and the slackvector $\boldsymbol{s}_0$ where $\boldsymbol{s} = \boldsymbol{A}^T\boldsymbol{x} - \boldsymbol{b}, \boldsymbol{s} \geq 0$. These starting points serves to calculate the initial residual vectors $\boldsymbol{r}_d$, $\boldsymbol{r}_s$ and $\boldsymbol{r}_{s\lambda}$

$$\begin{aligned} \boldsymbol{r}_d &= \boldsymbol{G}\boldsymbol{x}_0 + \boldsymbol{g} - \boldsymbol{A}\boldsymbol{\lambda}_0 \\ \boldsymbol{r}_p &= \boldsymbol{s}_0 - \boldsymbol{A}^T\boldsymbol{x}_0 + \boldsymbol{b} \\ \boldsymbol{r}_{s\lambda} &= \boldsymbol{S}_0 \boldsymbol{\Lambda}_0 \boldsymbol{e} \end{aligned} \tag{31}$$

where $\boldsymbol{S}_0$ and $\boldsymbol{\Lambda}_0$ are the diagonal matrices containing the elements of the $\boldsymbol{s}_0$ and $\boldsymbol{\lambda}_0$. The $\boldsymbol{e}$ is vector of ones (Nocedal and Wright 2000; Wright 1997).

There is also need to calculate the initial complementarity measure μ which is need for centering parameter σ

$$\mu = \frac{\boldsymbol{s}_0^T \boldsymbol{\lambda}_0}{m} \tag{32}$$

where m is the number of the inequality constraints.

The whole algorithm can be divided into two parts. The first is the calculation of the predictor step and the second is the calculation of the corrector step. The predictor step is calculated by applying the Newton's method around the current point on the equations (31).

$$\begin{bmatrix} \boldsymbol{G} & -\boldsymbol{A} & \boldsymbol{0} \\ -\boldsymbol{A}^T & \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{S} & \boldsymbol{\Lambda} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{x}^{aff} \\ \Delta\boldsymbol{\lambda}^{aff} \\ \Delta\boldsymbol{s}^{aff} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{r}_d \\ \boldsymbol{r}_p \\ \boldsymbol{r}_{s\lambda} \end{bmatrix} \tag{33}$$

The affine scaling direction $\left(\Delta\boldsymbol{x}^{aff}, \Delta\boldsymbol{\lambda}^{aff}, \Delta\boldsymbol{s}^{aff}\right)$ is obtained by solving these equations. Then the scaling parameter α^{aff} for the predictor step is chosen. This parameter have to satisfy the conditions in the equations (34).

$$\begin{aligned} \boldsymbol{\lambda} + \alpha_\lambda^{aff} \Delta\boldsymbol{\lambda}^{aff} &\geq 0 \\ \boldsymbol{s} + \alpha_s^{aff} \Delta\boldsymbol{s}^{aff} &\geq 0 \end{aligned} \tag{34}$$

The final scaling parameter is chosen in the following way:

$$\begin{aligned} \alpha_\lambda^{aff} &= \min_{i:\Delta\lambda_i<0}\left(1, \min \frac{-\lambda_i}{\Delta\lambda_i^{aff}}\right) \\ \alpha_s^{aff} &= \min_{i:\Delta s_i<0}\left(1, \min \frac{-s_i}{\Delta s_i^{aff}}\right) \\ \alpha^{aff} &= \min\left(\alpha_\lambda^{aff}, \alpha_s^{aff}\right) \end{aligned} \tag{35}$$

Now the complementarity measure μ^{aff} of the predictor step and the centering parameter σ can be calculated.

$$\mu^{aff} = \frac{\left(\boldsymbol{s} + \alpha^{aff}\Delta\boldsymbol{s}^{aff}\right)^T \left(\boldsymbol{\lambda} + \alpha^{aff}\Delta\boldsymbol{\lambda}^{aff}\right)}{m} \tag{36}$$

$$\sigma = \left(\frac{\mu^{aff}}{\mu}\right)^3 \tag{37}$$

Now we can move to the calculation of the corrector step. This is done by adjusting the right hand side of the equation (33) by computed affine scaling direction and the centering parameter. The resulting equation system is shown as equation (38) (Nocedal and Wright 2000; Wright 1997).

$$\begin{bmatrix} \boldsymbol{G} & -\boldsymbol{A} & \boldsymbol{0} \\ -\boldsymbol{A}^T & \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{S} & \boldsymbol{\Lambda} \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{\lambda} \\ \Delta\boldsymbol{s} \end{bmatrix} = - \begin{bmatrix} \boldsymbol{r}_d \\ \boldsymbol{r}_p \\ \boldsymbol{r}_{s\lambda} + \Delta\boldsymbol{S}^{aff}\Delta\boldsymbol{\Lambda}^{aff}\boldsymbol{e} - \sigma\mu\boldsymbol{e} \end{bmatrix} \tag{38}$$

Solving this system gives us the final scaling direction $\left(\Delta\boldsymbol{x}, \Delta\boldsymbol{\lambda}, \Delta\boldsymbol{s}\right)$. The step length is chosen in the same way it was in the predictor step calculation in the equations (35).

$$\begin{aligned} \boldsymbol{\lambda} + \alpha_\lambda \Delta\boldsymbol{\lambda} &\geq 0 \\ \boldsymbol{s} + \alpha_s \Delta\boldsymbol{s} &\geq 0 \end{aligned} \tag{39}$$

Now we can update the unknown vector $\boldsymbol{x}$, the vector of the Lagrange multipliers $\boldsymbol{\lambda}$ and the slackvector $\boldsymbol{s}$.

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \alpha\Delta\boldsymbol{x} \\ \boldsymbol{\lambda}_{k+1} &= \boldsymbol{\lambda}_k + \alpha\Delta\boldsymbol{\lambda} \\ \boldsymbol{s}_{k+1} &= \boldsymbol{s}_k + \alpha\Delta\boldsymbol{s} \end{aligned} \tag{40}$$

The final step of this algorithm is updating the residuals vectors $\boldsymbol{r}_d$, $\boldsymbol{r}_s$ and $\boldsymbol{r}_{s\lambda}$ and the complementarity measure μ.

$$\begin{aligned} \boldsymbol{r}_d &= \boldsymbol{G}\boldsymbol{x} + \boldsymbol{g} - \boldsymbol{A}\boldsymbol{\lambda} \\ \boldsymbol{r}_p &= \boldsymbol{s} - \boldsymbol{A}^T\boldsymbol{x} + \boldsymbol{b} \\ \boldsymbol{r}_{s\lambda} &= \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{e} \end{aligned} \tag{41}$$

$$\mu = \frac{\boldsymbol{s}^T\boldsymbol{\lambda}}{m} \tag{42}$$

RESULTS

This section shows the results of the process simulation for the step reference signal and the ramp reference signal. Both of the simulation were done with same controller parameters N = 20 steps, λ = 0.001, δ = 10 and the sampling period T_0 = 40 ms. The only difference between done simulations is in the control signal calculation method. The first method is quadprog function built-in in the Matlab and the second one is the presented predictor-corrector method. The results compare the computation time of one control step. The time was measured by the Matlab tic() ... toc() function. The simulations were also compared by two quadratic criterions for analysis of the control quality. The first criterion, described in equation (43), compares the control increments made in every step and the second criterion, described in equation (44), compares a difference between the reference value and the output value.

$$S_u = \frac{1}{N}\sum_{k=1}^{N}\Delta u^2(k) \tag{43}$$

$$S_e = \frac{1}{N}\sum_{k=1}^{N}\left[w(k)-y(k)\right]^2 \tag{44}$$

Table 1 shows the system parameters used for the mathematical model of the system.

Table 1 : System parameters

Symbol	Value	Meaning
M_0 [kg]	4	Cart weight
M_1 [kg]	0.36	Pendulum weight
l_s [m]	0.42	Pendulum length
Θ [kg.m^2]	0.08433	Pendulum inertia moment
F_r [kg/s]	6.5	Cart friction
C [Kg.m^2/s]	0.00652	Pendulum friction
k_a [N/V]	7.5	Servo amplifier gain
g [m/s^2]	9.81	Gravity constant

The figure 3 shows the pulse response of the system for the input pulse F = 2 N for 1s.

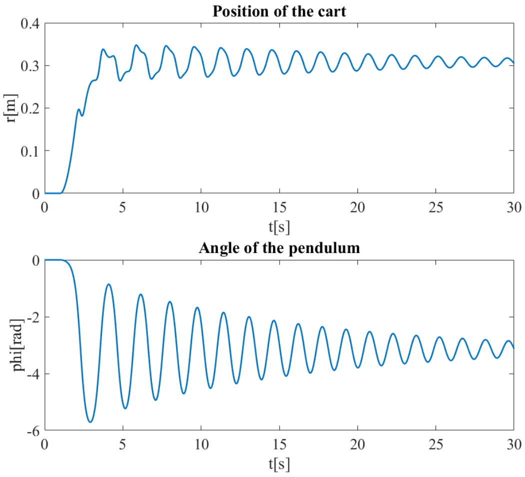


Figure 3 : Pulse response

As we can see, the angle of the pendulum dropped on the value of -π rad. That means the zero value of the pendulum angle is the upward position of the pendulum. Figures 5 and 6 represents the simulations result for the step reference signal.

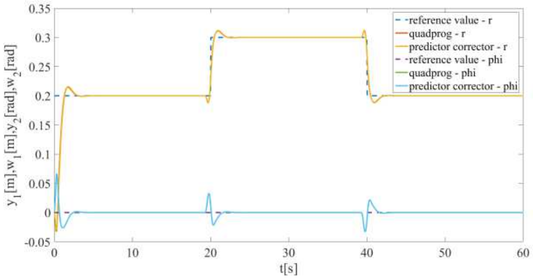


Figure 4 : Simulation outputs

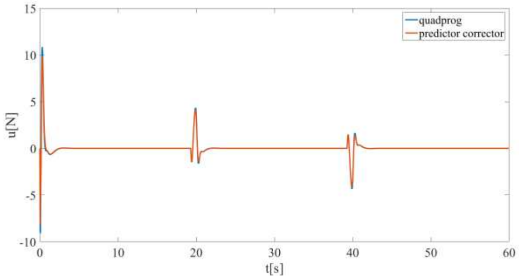


Figure 5 : Simulation inputs

Table 2 :Simulation results

	S_{e1}[m^2]	S_{e2}[rad^2]	S_u[N^2]
Predictor-corrector	5.28 . 10^{-4}	4.68 . 10^{-5}	0.102
quadprog	5.13 . 10^{-4}	4.95 . 10^{-5}	0.130

Computation time:

- Predictor-corrector - 7.73 ms
- Quadprog - 24 ms

Figures 6 and 7 represents the simulation result for the ramp reference signal.

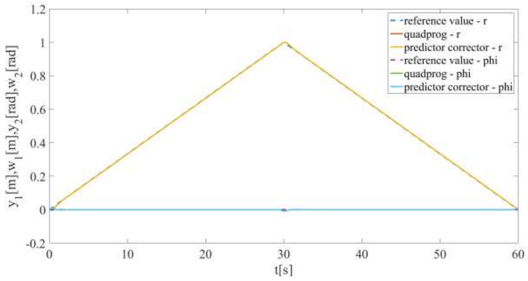


Figure 6 : Simulation outputs

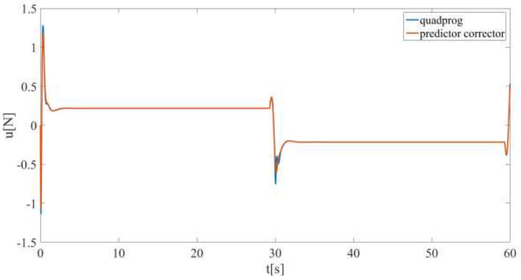


Figure 7 : Simulation inputs

Table 3 : Simulation results

	S_{e1}[m^2]	S_{e2}[rad^2]	S_u[N^2]
Predictor-corrector	$4.94 \,.\, 10^{-6}$	$1.63 \,.\, 10^{-6}$	$1.51 \,.\, 10^{-3}$
quadprog	$4.87 \,.\, 10^{-6}$	$1.66 \,.\, 10^{-6}$	$1.78 \,.\, 10^{-3}$

Computation time:

- Predictor-corrector - 7.99 ms
- Quadprog - 24.5 ms

CONCLUSION

In this paper, the predictive controller based on the state-space CARIMA was presented. The controller was tested on the inverted pendulum system which is an example of the nonlinear single-input two-output system. The goal of the control of this system is to keep the pendulum rod at the upward position while the cart is moving. The movement of the pendulum acting like a disturbance in the system. This system is also relatively fast with chosen sampling period T_0 = 40 ms. The mathematical model of the inverted pendulum was made according to the real laboratory model of the inverted pendulum Amira PS600. The aim of this paper is to present a complex procedure of the creation of the predictive controller which is capable to control such process. The presented predictive controller works with a linear models while the inverted pendulum model is nonlinear. Therefore the linearization of the nonlinear model is also presented besides only the inverted pendulum mathematical model explanation. The calculation of the input signal is done by the minimization of the cost function that minimize the differences between the output and the reference signals and the control signal increments. This minimization is achieved by two different methods. The first one is the quadprog function built-in the Matlab and the second one is the presented predictor-corrector method. The result section compares these two methods. Both methods have almost the same results according to the examined criterions Se and Su. The major difference between them is in the computation time. The mean computation time of the quadprog function is three times longer than the time of the predictor-corrector method.

REFERENCES

Albertos Pérez P. and Sala A. 2004. *Multivariable Control Systems: an Engineering Approach*. Springer. London.

Amira. 2000. *PS600 Laboratory Experiment Inverted Pendulum.* Amira GmbH, Duisburg.

Bars R.; R. Haber and U. Schmitz. 2011. *Predictive control in process engineering: From the basics to the applications.* Weinhaim: Willey-VCH Verlag.

Bobál, V. 2008, *Adaptive and predictive control.* vol. 1. Zlin, Tomas Bata University in Zlín.

Camacho E.F. and C. Bordons. 2004. *Model predictive control*, Springer Verlag, London.

Chalupa P. and V. Bobál. 2008. "Modelling and Predictive Control of Inverted Pendulum". In: Proceedings 22nd European Conference on Modelling and Simulation. pp. 531-537.

Fikar M. and J. Mikleš. 2008. *Process modelling, optimisation and control*, Springer-Verlag, Berlin.

Hangos K.M.; Bokor J. and Szederkényi G. 2004. *Analysis and Control of Nonlinear Process Systems*. Springer. London.

Maciejowski J.M. 2002. *Predictive control with constraints*, Prentice Hall, London.

Nocedal J. and S. Wright. 2000. *Numerical optimisation second edition*. Springer, New York.

Rossiter J.A. 2003. *Model based predictive control: a practical approach*, CRC Press.

Wang L. 2009. *Model predictive control system design and implementation using MATLAB*, Springer Verlag, London.

Wright S. 1997 *Primal-dual interior point methods.* Philadelphia: Society for Industrial and Applied Mathematics.

ACKNOWLEDGMENT

This article was created with support of the Ministry of Education of the Czech Republic under grant IGA reg. n. IGA/FAI/2016/009.

AUTHOR BIOGRAPHIES

LUKÁŠ RUŠAR studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2014. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interests focus on model predictive control. His e-mail address is rusar@fai.utb.cz.

ADAM KRHOVJÁK studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2013. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interests focus on modeling and simulation of continuous time technological processes, adaptive and nonlinear control.

STANISLAV TALAŠ studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2013. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His e-mail address is talas@fai.utb.cz.

VLADIMÍR BOBÁL graduated in 1966 from the Brno University of Technology, Czech Republic. He received his Ph.D. degree in Technical Cybernetics at Institute of Technical Cybernetics, Slovak Academy of Sciences, Bratislava, Slovak Republic. He is now Professor at the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín, Czech Republic. His research interests are adaptive and predictive control, system identification, time-delay systems and CAD for automatic control systems. You can contact him on email address bobal@fai.utb.cz.

1DOF GAIN SCHEDULED PH CONTROL OF CSTR

Adam Krhovják, Stanislav Talaš and Lukáš Rušar
Department of Process Control
Faculty of Applied Informatics, Tomas Bata University in Zlín,
Nad Stráněmi 4511, 760 05 Zlín, Czech Republic
krhovjak@fai.utb.cz

KEYWORDS
Neutralization, continuously stirred tank reactor, nonlinear model, parametrized linear model, scheduling variable, gain scheduled controller.

ABSTRACT

Motivated by the rich dynamics of chemical processes, we present a gain scheduled control strategy for a pH neutralization occurring inside continuously stirred tank reactor built on a linearization of a nonlinear state equation about selected operating points. Firstly, we address the problem of a selection of scheduling variable. Based on this, an extra scheduling mechanism is presented to simulate the behavior of a nonlinear process using a linear model. Specifically, the proposed step aims at extending the region of validity of linearization by introducing a parametrized linear model, which enables to construct linear controller at each point. Finally, the parameters of resulting family of linear controllers are scheduled as functions of the reference variable, resulting in a single scheduling controller.

INTRODUCTION

Today's world is full of complex systems that try to facilitate our live. More or less complex nonlinear control algorithms play an integral role in advanced flight systems, adaptive strategies has been successfully implemented in robotic manipulators. One can even say that we are heading towards a new industrial revolution of intelligent systems interacting with each other. It is evident that the more complex system is the more challenging control task it brings. From this perspective, the crucial point of the design is to find a balanced way between effectivity and complexity.

When moving from linear to nonlinear spheres superposition principle known from linear systems does not hold any longer and we are offered richer scenarios that are unfortunately difficult to control. However since linear system theory is so much more traceable we can view its mathematics as an effective tool to simplify our design problem. There is no question that whenever possible, we should utilize this useful trick. Basically, there is only one limitation we must bear in mind. The fact that such a controller can only operate in the neighborhood of a single operating point, predicting the local behavior of the nonlinear system. Interestingly enough, in many cases, it is possible to capture how the dynamics of a system change in its equilibrium. Repeating this idea in multiple operating points we are given a whole family of linearized systems that act as parameterized linear model (PLM). Moreover, it may be even possible to find one or more variables that parameterize this equilibrium points. In such cases, it is intuitively reasonable to linearize the nonlinear model about selected operating points, capturing key states of a system, design a linear controller at each point, and interpolate the resulting family of linear controllers by monitoring scheduling variables.

Early scheduling theories revealed a crucial point associated with selection of scheduling variables. Although there has been a great deal of research discussion about scheduling guidelines it did not bring any general clue. We refer the interest reader to (Rugh 1991; Shamma and Athans 1991; Shamma and Athans 1992; Lawrence and Rugh 1995) for deeper and more insightful into problematics. Moreover, most efforts have been concentrated on nonlinear equations of motion of airplanes or missiles (Khalil 2002), except few applications in car engines (Jiang 1994; Kaminer et al. 1995).

However, nonlinear phenomena are more than typical for chemical processes that may even occur in living cells. We have heard of studies, claiming that metabolism of cancer cells is significantly affected by pH value (Kroemer and Pouyssegur 2008).

From this view point, it is very important to maintain its value inside the boundaries.

In order to reflex richness of pH processes, we have stressed to illustrate gain scheduling strategy for pH neutralization running inside continuously stirred tank reactor (CSTR). Throughout the paper we gradually reveal the scheduling procedure satisfying a tracking problem as well as the design of a control trajectory for the CSTR. We will see that it allows us to smoothly move from one design to another, even though the process exhibits significant nonlinearities.

MODEL OF THE CSTR

A simplified model of the CSTR (Corriou 2004) shown in Figure 1 illustrates neutralization of waste water. As can be seen, process consists of two inlet streams that are perfectly mixed inside reactor.

The first stream represents waste water that enters the reactor with the constant feedrate of q_w at $\mathrm{pH_w}$. Pump with a flow rate q_a that discharges acid serve as actuator to neutralize waste water.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Both parameters of the CSTR and initial conditions of the neutralization process are captured in Table 1.

Table 1: Model parameters

Waste water inlet stream	q_w= 2000 l/h
Waste water OH$^-$ concentration	$C_w^{OH^-}$ = 10^{-13} mol/l
Waste water H$^+$ concentration	$C_w^{H^+}$ = 10^{-1} mol/l
Acid pump OH$^-$ concentration	$C_a^{OH^-}$ = 1.29·10^{-15} mol/l
Acid pump H$^+$ concentration	$C_a^{H^+}$ = 7.7371 mol/l
Initial pH	pH(t = 0) = 13
Reactor volume	V = 4000 l

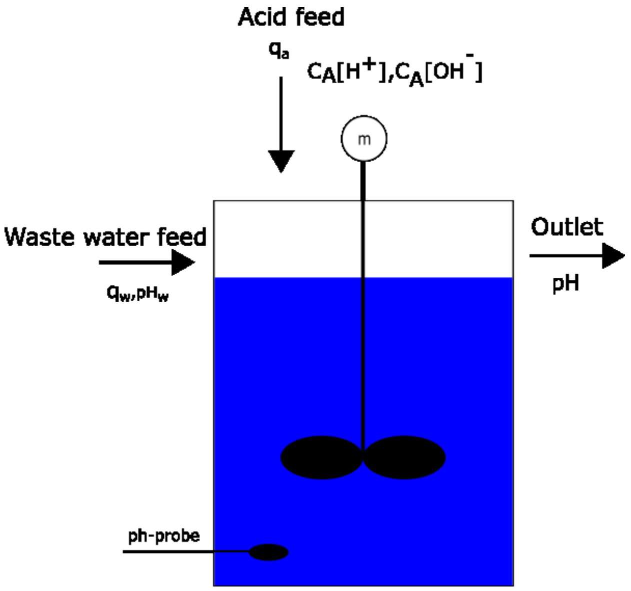


Figure 1: Continuously stirred tank reactor

The only step needed to develop the model of CSTR is to write conservation equation (Richardson 1989), representing material balance for a single material. Recall that the general form of a mass balance is given.

INPUT = OUTPUT + ACCUMULATION

It is easy to see that the simplified dynamic behavior of the CSTR can be modeled by

$$V\frac{d}{dt}C^{OH^-} = q_w C_w^{OH^-} - (q_w + q_a)C^{OH^-} + q_a C_a^{OH^-} \quad (1)$$

$$V\frac{d}{dt}C^{H^+} = q_w C_w^{H^+} - (q_w + q_a)C^{H^+} + q_a C_a^{H^+} \quad (2)$$

where V represents reactor volume, C stands for the concentration of H^+ and OH^- ions q_w and q_a are waste and acid feedrates, respectively.

The main problem of neutralization is that autoprotolysis of water occurs.

$$H^+ + OH^- \Leftrightarrow H_2 0 \quad (3)$$

Assuming the constant temperature, this reaction has the approximate equilibrium constant

$$C^{H^+} \cdot C^{OH^-} = 10^{-14} \quad (4)$$

In order to simplify our task, let us suppose that this reaction does not take place and concentration of both ions remain in the solution. Indeed, since our control task is to track pH value we will find the equilibrium of the autoprotolysis.

To put it in another way, we known that reaction (4) consumes or produces a certain amount of ΔH, until it reaches steady state.

$$(C^{H^+} + \Delta H)\cdot(C^{OH^-} + \Delta H) = 10^{-14} \quad (5)$$

Substitution of $\tilde{C}^{H^+}$ for $(C^{H^+} + \Delta H)$ in (5) yields

$$(\tilde{C}^{H^+})\cdot(C^{OH^-} + \tilde{C}^{H^+} - C^{H^+}) = 10^{-14}, \quad (6)$$

which can be easily rewritten into the standard form as

$$0 = \underbrace{-1}_{a}(\tilde{C}^{H^+})^2 + \underbrace{(C^{H^+} - C^{OH^-})}_{b}\cdot\tilde{C}^{H^+} + \underbrace{10^{-14}}_{c} \quad (7)$$

It can be easily seen that the above equation must be solved for unknown $\tilde{C}^{H^+}$. Fortunatelly, (7) is quadratic equation, producing straight forward solution.

Having calculated (7), the pH value is given by

$$pH = -\log(\tilde{C}^{H^+}) \quad (8)$$

We can now see, that the equilibrium concentration $\tilde{C}^{H^+}$ depends only on the term b.

This leads us to the simple modification. By substracting (1) from (2) we obtain neutralization model

$$V\frac{d}{dt}b = q_w b_w - (q_w + q_a)b + q_a b_a \quad (9)$$

where b represents correponding difference between hydrogen and hydroxyl ions.

MODEL STRUCTURE FOR GS DESIGN

In the process of designing and implementing a gain scheduled controller for a nonlinear system, we have to find its approximations about the family of operating (equilibrium) points. Thus, the nonlinear first-order ordinary differential equation (9) capturing the dynamics of the neutralization process has to be transformed into its linearized form.

In view of our example, we shall deal with single-input single-output linearizable nonlinear system represented by

$$\dot{x} = f(x,u) \quad (10)$$

$$y = g(x) \quad (11)$$

where $\dot{x}$ denotes derivative of x with respect to time variable and u are specified input variables. We call the variable x the state variable and y the output variable. We shall refer to (10) and (11) together as the state-space model.

To obtain a state-space model of the CSTR, let us take $x = b$ as a state variable and $u = q_a$ as a control input. Then the state equation is

$$\frac{dx(t)}{dt} = \frac{1}{V}\left[q_w b_w - q_w x(t) - x(t)u(t) + b_a u(t)\right] \tag{12}$$

Because the output of the process is pH value, the output equation takes the form

$$y = -\log\left\{\frac{1}{2}\left[x(t) + \sqrt{x(t)^2 - 4\cdot 10^{-14}}\right]\right\} \tag{13}$$

One can easily sketch the trajectory of steady-state characteristic by setting $\dot{x} = 0$ and solving for unknown x.

Therefore the equilibrium points correspond to the solution of

$$0 = \frac{1}{V}\left[q_w b_w - q_w x(t) - x(t)u(t) + b_a u(t)\right] \tag{14}$$

Having calculated equilibrium points of state equation, our goal now is to approximate (12) about selected single operating point. Suppose $x \neq 0$ and $u \neq 0$, and consider the change of variables

$$x_\delta(t) = x(t) - \bar{x} \tag{15}$$

$$u_\delta(t) = u(t) - \bar{u} \tag{16}$$

$$y_\delta(t) = y(t) - \bar{y} \tag{17}$$

It should be noted that in the new variables system has equilibrium in origin.
Expanding the right hand side of (12) about point $(\bar{x},\bar{u})$, we obtain

$$f(x,u) \approx f(\bar{x},\bar{u}) + \frac{\partial f(\bar{x},\bar{u})}{\partial x} + \frac{\partial f(\bar{x},\bar{u})}{\partial u} + \text{H.O.T.} \tag{18}$$

If we restrict our attention to a sufficiently small neighborhood of the equilibrium point such that the Higher-Order Terms are negligible, then we may drop these terms and approximate the nonlinear state equation by the linear state equation.

$$\dot{x}_\delta = Ax_\delta + Bu_\delta \tag{19}$$

where

$$A = \left.\frac{\partial f}{\partial x}\right|_{x=\bar{x},u=\bar{u}} \qquad B = \left.\frac{\partial f}{\partial u}\right|_{x=\bar{x},u=\bar{u}} \tag{20}$$

PARAMETRIZATION OF LINEAR MODELS

Before we present a parametrization via scheduling variable, let us first examine configuration of the gain scheduled control system captured in Figure 2. From the figure, it can be easily seen that controller parameters are automatically changed in open loop fashion by monitoring operating conditions. From this point of view, presented gain scheduled control system can be understand as a feedback control system in which the feedback gains are adjusted using feedforward gain scheduler.

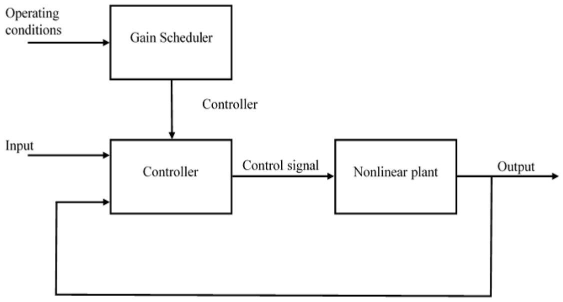


Figure 2: Gain scheduled control

Then it comes as no surprise that first and the most important step in designing a controller is to find an appropriate scheduling strategy. Once the strategy is found, it can be directly embedded into the controller design.

In order to understand the idea behind the gain scheduling let us first consider the nonlinear system

$$\dot{x} = f(x,u,\alpha) \tag{21}$$

$$y = g(x,\alpha) \tag{22}$$

We can see that the nonlinear system is basically same as the system that we have introduced in the previous section by equations (10) and (11). The only difference here is that both state and output equations are parameterized by a new *scheduling variable α* representing the operating conditions.

To illustrate this motivating discussion let us consider this crucial point in the context of our example.

Suppose the system is operating at steady state and we want to design controller such that *y* tracks a reference signal *w*. In order to maintain the output of the plant at the value $\bar{y}$, we have to generate the corresponding input signal to the system at $\bar{u} = q_w \frac{\bar{x} - b_w}{b_a - \bar{x}}$. This implies that for every value of *w* in the operating range, we can define the desired operating point by

$$\bar{y} = w \tag{23}$$

$$\bar{u} = q_w \frac{\bar{x}(w) - b_w}{b_a - \bar{x}(w)} \tag{24}$$

In other words, this leads us to the simply conclusion that we can directly schedule on a reference pH trajectory.

Having identified a scheduling variable, the common scheduling scenario takes this form

$$\dot{x}_\delta = A(\alpha)x_\delta + B(\alpha)u_\delta \tag{25}$$

Intuitively speaking, the parameters of (19) are scheduled as functions of the scheduling variable *α*. Since our model is simple nonlinear SISO system, we need to calculate constants of *A*, *B*. In other words, the key how to move from one operating point to another is given by

$$\begin{aligned} A(\alpha) &= \left.\frac{\partial f}{\partial x}\right|_{x=\bar{x},u=\bar{u}} = -\frac{q_w + u(\alpha)}{V} \\ B(\alpha) &= \left.\frac{\partial f}{\partial u}\right|_{x=\bar{x},u=\bar{u}} = -\frac{b_a - x(\alpha)}{V} \end{aligned} \qquad (26)$$

An important feature of our analysis is that even if *α* represents reference vector, the equations (24) still capture the behavior of the system around equilibria. We can also observe that both *x* and *u* are functions of *α*. This is no problem since we have defined our scheduling variable as a desired pH value. Interested reader has certainly noticed that we can easily calculate them from (23) and (24).

GAIN SCHEDULED CONTROLLER DESIGN

Since the basis for the construction of family of parametrized linear models has been previously explored, we would like to look more closely at the derivation of the linear controller at each operating point that is a prescription for designing *u* such that *y* asymptotically tracks *w* with all generated signals remaining bounded. In order to achieve this control objective, we analyze one degree of freedom configuration. The configuration arrangement is shown in Figure 3 and includes one degree of freedom that is represented by feedback controller G_Q.

In this configuration, *w* represents the reference signal, *v* is the load disturbance, *y* is the controlled output and *u* is the control input

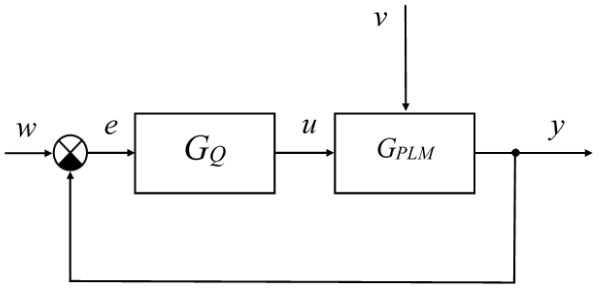


Figure 3: 1DOF control system configuration

Consider now the single-input single output linearized system described by equation (19) or, equivalently, by the transfer function model

$$Y(s) = G_{PLM}(s)U(s) = \frac{b(s)}{a(s)}U(s) \qquad (27)$$

where *U*(*s*) and *Y*(*s*) are Laplace transforms of the control input *u*(*t*) and measured output *y*(*t*)**,** respectively.

From the equation (19) it can be seen that the polynomials *a* and *b* are monic having following structure

$$a(s) = s + a_0 \qquad (28)$$

$$b(s) = b_0 \qquad (29)$$

To aid insight into controller

$$G_Q(s) = \frac{q(s)}{p(s)}, \qquad (30)$$

where *q* and *p* represents polynomials in *s*, we will work with both reference signal and disturbance signal as follows

$$W(s) = \frac{w_0}{s},\ V(s) = \frac{v_0}{s} \qquad (31)$$

As is well known we have to use integral control such that polynomial *p* takes the form

$$p(s) = s\tilde{p}(s) \qquad (32)$$

To proceed with the design of the controllers, we leave it as an exercise for the reader to verify that both closed-loop linear systems has the characteristic equation. The reader may consult (Kučera 1993).

$$a(s)\,p(s) + b(s)\,q(s) = d(s) \qquad (33)$$

One can intuitively expect that the control system is stable if we design *d* to be Hurwitz polynomial.

Toward the goal suppose we have succeeded in finding polynomials of the transfer functions

$$G_Q(s) = \frac{q(s)}{s\tilde{p}(s)} = \frac{q_1 s + q_0}{s p_0} \qquad (34)$$

that satisfy (33) for the stable polynomial

$$d(s) = (s + \beta_1)(s + \beta_2) \qquad (35)$$

Recalling the basis of linear algebra, we can obtain the controller parameters from the solution of the matrix equation

$$\begin{bmatrix} 1 & 0 & 0 \\ a_0 & b_0 & 0 \\ 0 & 0 & b_0 \end{bmatrix} \begin{bmatrix} p_0 \\ q_1 \\ q_0 \end{bmatrix} = \begin{bmatrix} d_2 \\ d_1 \\ d_0 \end{bmatrix} \qquad (36)$$

where the coefficients of polynomial *d* are given by

$$d_2 = 1, d_2 = \beta_1 + \beta_2, d_0 = \beta_1 \beta_2 \qquad (37)$$

It is important to emphasize that selectable poles β_1 and β_2 are the only parameters through which the controller parameters can be adjusted.

The resulting gain scheduled controller can be obtained by scheduling coefficients of *q*(s) as functions of *α*; that is, *α* is replaced by *w*, so that the gains vary directly with the desired pH value.

From this, the linear control law, which is prescribed by controller (34) can be rewritten in terms of scheduling variables as

$$u = q_1(\alpha)e + q_0(\alpha)\sigma \qquad (38)$$

where

$$e = \dot{\sigma}$$

So far, we have formed the basic idea of construction of gain scheduled control law. All that remains now is to show, that for a desired Hurwitz polynomial *d*, the gains are taken as

$$\begin{aligned} q_1(\alpha) &= \frac{\beta_1 + \beta_2 - a_0(\alpha)}{b_0(\alpha)} \\ q_0(\alpha) &= \frac{\beta_1 \beta_2}{b_0(\alpha)} \end{aligned} \tag{39}$$

When the control (38) is applied to the nonlinear state equation (12) it results in the closed-loop system

$$\dot{x} = \frac{1}{V}\left[q_w b_w - q_w x + q_1 (b_a - x) q_1 \left(e + \frac{q_0}{q_1} \sigma \right) \right] \tag{40}$$

In view of the procedure that we have just described, one can notice that three main issues are involved in the development of gain scheduled controller; namely linearization of neutralization process about the family of operating regions, design of a parametrized family of linear controllers for the parametrized family of linear systems and construction of gain scheduled controller.

SIMULATIONS AND RESULTS

In this section, we simulate the gain scheduled control of CSTR. We have developed a custom MATLAB function based on the simulator introduced by (Krhovják et al. 2015) that simulates adequately the behavior of CSTR. Idealistic model has been implemented according to equation (8) and (9). The popular ODE solver using based on Runge-Kutta methods (Hairer et al. 1993) was considered to calculate numerical solution.

The simulation results of gain scheduled control are presented in Figures 4-6. Figure 4 clearly illustrates how the linearized plant dynamics vary with the operating conditions that are given by scheduling variable *α*.

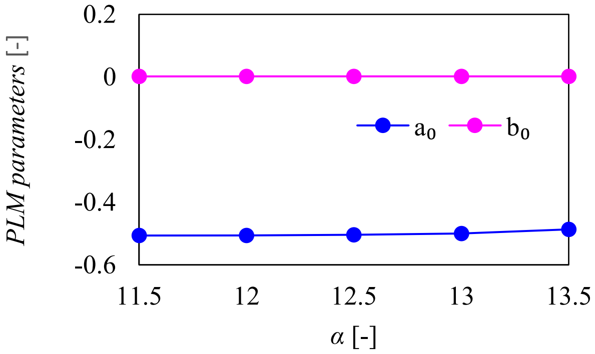


Figure 4 Parameter evolution

Figure 5 shows the responses of the control system to the sequence of step changes in reference signal. As can be seen from Figure 6, we have found such a combination of parameters *β*1 and *β*2 that results in reasonably good responses.

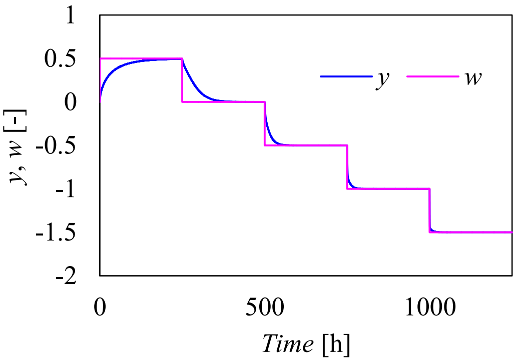


Figure 5: The responses of the closed-loop system to a sequence of step changes

From a gain-scheduling viewpoint, a step change in reference signals causes a new calculation of the equilibrium point of the system. This claim is also supported by Figure 6 in which the gain adjustment is captured. It is important to notice that the change of controller parameters occurs with the step change in reference trajectory.

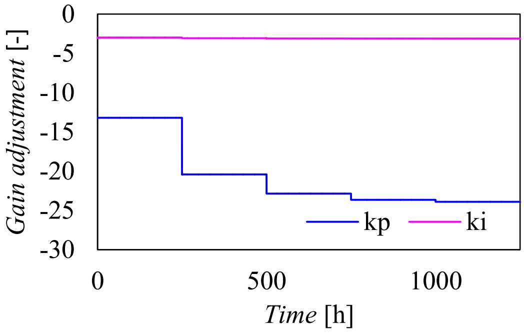


Figure 6: Gain adjustment during control

CONCLUSION

This paper addressed the control problem of neutralization processes. Excited by their dynamics, we presented a promising gain scheduling strategy that overcomes a highly nonlinear behavior. First, we have detailed studied the simplified model of the neutralization process. Based on the model, we have followed a general analytical framework for gain scheduling. The most importantly, we outlined a possible way how to select an appropriate scheduling variable. The main advantage of this approach is that linear design methods can be applied to the linearized system at each operating point. Thanks to this feature, the presented procedure leaves room for many linear control methods We have demonstrated that a gain scheduled control system has the potential to respond rapidly changing operating conditions

ACKNOWLEDGEMENT

This article was created with support of the Ministry of Education of the Czech Republic under grant IGA reg. n. IGA/FAI/2017/009.

REFERENCES

Corriou, J.P. 2004. Process control: theory and applications, London. Springer.

Hairer, E; S.P. Norsett; and G. Wanner. 1993. *Solving ordinary differential equations*. 2nd revised ed. Berlin: Springer.

Jiang, J. 1994. "Optimal gain scheduling controllers for a diesel engine*". IEEE Control Systems Magazine*, 14(4), 42-48.

Kaminer, I; A. M. Paswal; P. P. Khargonekar; and E. E. Coleman. 1995. "A velocity algorithm for the implementation of gain scheduled controllers". *Automatica*, 31, 1185-1191.

Khalil, H. K. "Nonlinear systems". 2002. Upper Saddle River, N.J.: *Prentice Hall.*

Krhovják, A.; P. Dostál; S. Talaš. 2015; and L.Rušar. "Multivariale gain scheduled control of two funnel liquid tanks in series". *in Process Control (PC), 2015 20th International Conference on*, pp. 60-65.

Kroemer, G. and J. Pouyssegur.2008."Tumor Cell Metabolism: Cancer's Achilles' Heel", *Cancer Cell*, Volume 13, Issue 6,472-482

Kučera, V. 1993. "Diophantine equations in control – A survey". *Automatica*, 29, 1361-1375.

Lawrence, D. A. and W. J. Rugh. 1995. "Gain scheduling dynamic linear controllers for a nonlinear plant". *Automatica*, 31, 381-390.

Shamma, J.S.; M. Athans. 1990. "Analysis of gain scheduled control for nonlinear plants. (1990) *IEEE Transactions on Automatic Control*, 35 (8), pp. 898-907.

Shamma, J.S and M. Athans. 1992. "Gain scheduling: potential hazards and possible remedies". *IEEE Control Systems Magazine*, 12(3), 101-107.

Shamma, J.S. and M.Athans. 1991. "Guaranteed properties of gain scheduled control of linear parameter-varying plants". *Automatica*, vol. 27, no. 4, 559-564.

Rugh, W.J. 1991 "Analytical framework for gain scheduling". *IEEE Control Systems Magazine*, 11(1), pp. 79-84.

Richardson, S.M. 1989. Fluid mechanics, New York, Hemisphere Pub. Corp.

AUTHOR BIOGRAPHIES

ADAM KRHOVJÁK studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2013. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interests focus on modeling and simulation of continuous time technological processes, adaptive and nonlinear control.

STANISLAV TALAŠ studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2013. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His e-mail address is talas@fai.utb.cz.

LUKÁŠ RUŠAR studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master degree in Automatic Control and Informatics in 2014. He now attends PhD. study in the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interests focus on model predictive control. His e-mail address is rusar@fai.utb.cz.

DESIGN OF A SIMPLE BANDPASS FILTER OF A THIRD OCTAVE EQUALIZER

Martin Pospisilik
Faculty of Applied Informatics
Tomas Bata University in Zlin
Nad Stranemi 4511, 760 05 Zlin, Czech Republic
E-mail: pospisilik@fai.utb.cz

KEYWORDS
Band Pass, Filter, Transistor Amplifier, Simulation

ABSTRACT

As they are spread across almost all applications based on electrical circuits, frequency filters form one of the most complex issues on electronic devices. There exist a great number of different design methods based on various approaches. A large group of these filters is based on active components, namely amplifiers, that are equipped with the appropriate feedback to achieve the desired transfer function. In terms of applications operating at audible frequencies it became to be a rule that one or two operational amplifiers are applied, taking into consideration that their input impedance and amplification factor are infinitely high, whilst their output impedance is infinitely low. The simulation results presented in this paper show that these considerations may not be so strict. That means the input impedance may be in tens of kiloohms and the amplification factor of 40 dB without the feedback loop can do enough. Moreover, the operational amplifiers may be replaced by a simple circuit consisting of two transistors, as it is shown in this paper. This should be considered when calculating manufacturing costs of several designs. In this paper, a detailed description of a band pass filter for a third octave equalizer is presented. However, it is obvious, that the same approach can be applied in other designs.

INTRODUCTION

The third octave equalizer is a device used for adjusting of a transfer function of various electro-acoustical chains. According to its name, the frequency range of the electro acoustic chain is divided into several bands the width of whose is 1/3 octave. That means there are 28 to 32 bands the gain of which can be set independently. These bands are splitted by means of bank of filters, usually 2nd order band passes.

Usual constructions

Usually, one of the approaches described in the following subchapters is employed in the construction of the third octave equalizers.

One block with serial impedances
The basic configuration of this circuit is depicted in the figure 1. This construction requires only one inverting amplifier and a set of tuned RLC filters. The inductors are usually replaced by synthetic inductances (gyrators), as depicted in the figure 2. The main advantage of this solution is the fact that it is simple and cheap. As all bands are implemented within one block, the circuit shows good noise parameters and there is no need to compensate gain of individual blocks. However, this solution brings a lot of drawbacks. The quality factor (Q) varies according to the setting of the appropriate band gain. The neighbouring bands affect each other.

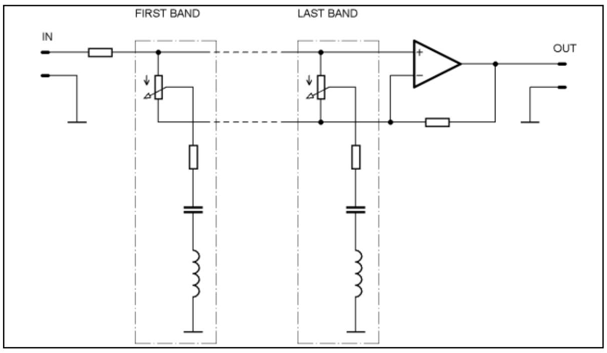


Figure 1: The Basis of an Equalizer Employing Serial RLC Filters

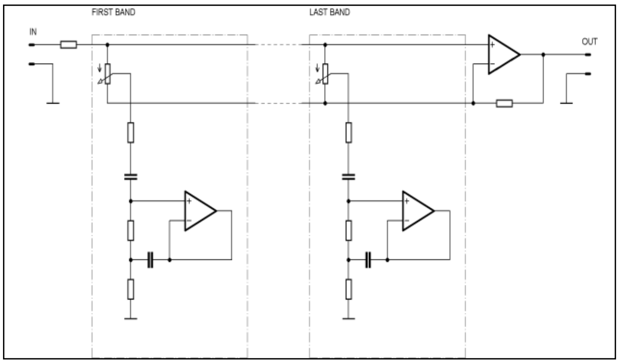


Figure 2: The Basis of an Equalizer Employing Serial RLC Filters with Synthetic Inductances

Cascade of blocks with serial impedances
This solution divides the circuit from Fig. 1 into a cascade of such blocks, each for one or several bands. The advantage of this solution consists in suppression of

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

mutual influences between the setting of the band potentiometers but, on the other hand, the noise of the blocks is cumulated and in order to achieve flat frequency response when all the potentiometers are set into the central position, the gain of all blocks must be set to 1 with no deviation.

Bank of parallel filters

Different approach to the construction is depicted in Fig. 3. It is based on a multiple input amplifier, the inputs of which are fed by means of band passes connected in parallel. The potentiometers are then connected in the negative feedback of the amplifier, allowing controlling of the contribution of each of the frequency bands to the transfer function of the circuit. For this purpose, a bank of parallel filters must be established. An important condition must be fulfilled as follows: the phase shift at the point of overlap of the transfer functions of the relevant band passes must be ±45° in order to combine the outputs of the band pass filters in parallel. This condition implies that second order filters should be employed.

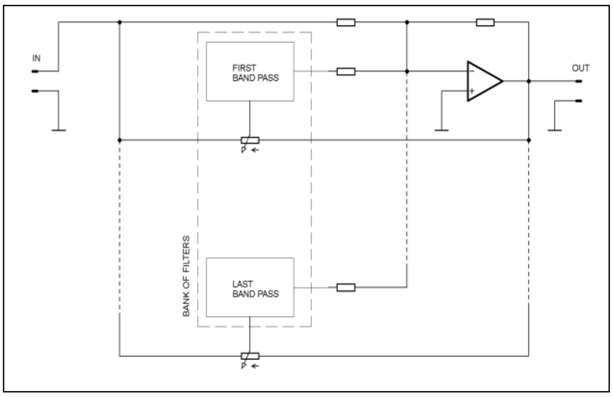


Figure 3: The Basis of an Equalizer Employing the Bank of Filters

The main advantage of this solution is the fact that the Q factor of each band pass remains independent on the setting of the appropriate potentiometer. As the band pass filters embody precisely defined transfer functions at all settings of the appropriate potentiometers, the risk of overexciting of some circuit's blocks at certain frequencies is eliminated. The main disadvantage consists in summation of the noise generated by the band passes.

It is worth mentioning that there exist modifications of the circuit depicted in Fig. 3 that allow to achieve better performance, nevertheless their complexity is adequately increased.

Motivation

The hereby described filter represents one band pass filter suitable for the topology depicted in Fig. 3. This topology has been chosen due to its advantages mentioned in the previous subchapter. Because the third octave equalizer employs from 28 to 32 of such filters for each of the two channels, reliable but cheap solution has been requested as well as a simple design of the appropriate printed circuit board. As the prices of discrete transistors are considerably low when bought in great numbers, it was decided to design the bank of filters on the basis of inverting amplifiers constructed with two transistors.

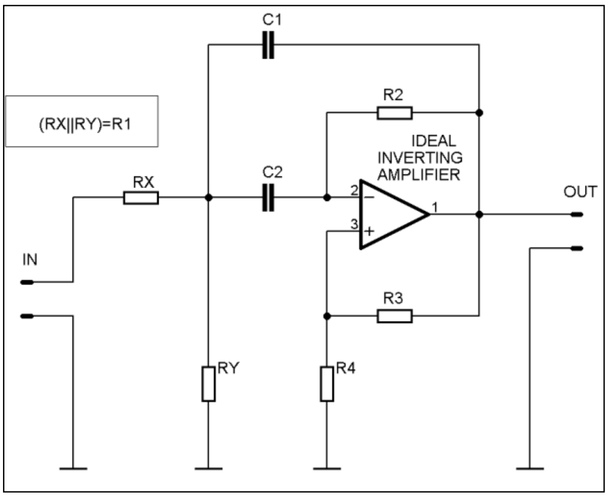


Figure 4: Band Pass Filter Topology

BAND PASS DESIGN

From the great variety of possible band pass constructions, the multiple feedback topology as described in (Smetana 2015) has been chosen. Neglecting the amplifier's characteristics, the circuit topology as depicted in Fig. 4 can be employed.

Initial requirements

In order to design the hereby mentioned band pass filter, the following requirements must be defined:

- Q factor,
- Centre frequency F_0,
- Overall gain in the pass band G_0.

According to (Hajek and Sedlacek 2002), the parameters of the filter depicted in Fig. 4 can be described by the following set of equations:

— (1)

— (2)

— (3)

— (4)

— (5)

$$G(F_0) = (1+\gamma)\sqrt{\alpha\beta}Q \qquad (6)$$

The equations (1) to (3) represent certain circuit parameters that are affected by the values of the relevant devices (see Fig. 4). The equations (4) to (6) represent the dependences of the required parameters of the circuit on the values of the circuit's devices. The overall gain of the circuit G(F_0) is rather great and therefore reduced by the voltage divider Rx, Ry to the desired value.

The hereby described filter has been designed according to the requirements enlisted in Table 1. Because the accurate capacitors are more discerning to obtain, it was decided to use the capacitors C1 a C2 of the same value. Therefore, according to (2), β = 1.

Table 1: Design Requirements

Parameter	**Value**
Centre frequency	F_0 = 1.00 kHz
Quality factor	Q = 4.32
Pass band gain	G_0 = 2.00
Capacities	$C_1 = C_2$ = 10.00 nF

The amplifier

In order to continue in the filter's design, some of the parameters of the amplifier must be known. As mentioned above, the amplifier based on two directly coupled transistors were employed. The circuit diagram of this amplifier is depicted in Figure 5.

The topology of the amplifier is quite common. The greatest advantage of this solution is the price as both transistors together can be purchased for less than 0.15 €. When designing the circuit, the following facts were assumed: The amplification of the signal at the input of the circuit is performed by the transistor T1. Therefore high beta and low noise transistor should be applied at this position. The transistor T2 is connected as an emitter follower, decreasing the output impedance of the circuit. The power supply voltage was chosen with respect to the circuit's linearity. ± 24 V symmetrical power supply is sufficient for this application. Concerning the noise issues, the collector current of the transistor T1 shall be low. Therefore the value of I_{c1} = 500 μA has been chosen. According to the facts mentioned above, it is expected that the gain factor of T1 is higher than 100. The voltage V_{R1} drop across the emitter resistor R14 is expected to be approximately 1 V. According to (7), the value of 2.2 kΩ has been chosen. Now the value of R13 can be calculated, as it as the voltage drop across the resistor is expected to be approximately 22 V. With I_{c1} = 500 μA the suitable value of R13 is 47 kΩ.

The T1 base current is expected to be lower than 5 μA. If R17 = 470 kΩ and IB1 < 5 μA, the voltage drop across R17 will be lower than 2.35 V. This has been considered when designing the voltage divider R15/R16. Provided the quiescent current through R17 is negligible and the voltage divider carries the current of 100 μA, the following values of the relevant resistors are applicable: R15 = 270 kΩ, R16 = 33 kΩ. In order to achieve low output impedance, quite great quiescent current through the transistor T2 was requested. With I_{C2} = 10 mA, the following values of R11 and R12 were applicable: R11 = 1.8 kΩ, R12 = 680 Ω. To achieve the amplification factor at least 100, the value of R18 should be lower than 1/100 of R13. The value of 330 Ω fits quite well.

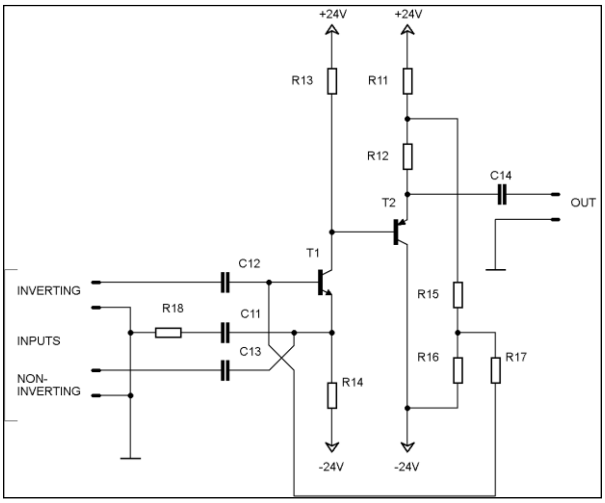


Figure 5: Amplifier Intended for the Band Pass Filter Design

Simulation results

The behavior of the above described amplifier was simulated in a software simulator based on SPICE algorithms and libraries for the following types of T1:

a) 2N5089,
b) BC550B,
c) 2N3904,
d) BC547C.

On the position of T2 always the transistor BC557B has been placed.

The simulation results are enlisted in the Table 2. According to the bias level recognized by the simulation, the correction of the R16 value has been implemented individually for each of the transistors. According to the results, the pair of BC547C and BC556B transistors has been considered as suitable for construction of the band pass filters. However, considering the total harmonic distortion (THD) and integrated output noise levels, it must be stated, that the SPICE models of the devices may not be accurate enough as the variability of the device parameters is quite large.

It should be emphasized that while the input impedance of the inverting input ranges in tens of kiloohms, the input impedance of the non-inverting input, affected mainly by the resistor R18, is considerably low. According to the simulation, the input impedance of approximately 260 Ω is expected at the non inverting input of the amplifier.

Table 2: Simulation Results of the Amplifier From Figure 5 with Different Types of T1

Simulated result	**Transistor**			
	2N5089	BC550B	2N3904	BC547C
Open loop gain	125	125	125	118
Bandwidth[1)]	1.00 MHz	1.00 MHz	0.98 MHz	1.10 MHz
Noise[2)]	45 μV	44 μV	46 μV	41 μV
Input impedance[3)]	47 kΩ	36 kΩ	21 kΩ	47 kΩ
THD	0.08 %	0.09 %	0.09 %	0.02 %
R16	27 kΩ	33 kΩ	47 kΩ	27 kΩ

[1)] Tolerance –3 dB.

[2)] Integrated output noise level for B = 10 kHz.

[3)] Input impedance measured at the inverting input.

As the values of capacitors C11 to C14 affect the transfer function of the circuit, this issue has been considered as well. The values were set in order to obtain full gain at the operating frequency of the amplifier, which is considered to be exactly 1 kHz. The simulation has shown that the capacities of 10 μF are enough.

Final circuit design

The band pass filter incorporates both the amplifier depicted in Fig. 5 and the topology depicted in Fig. 4. Its connection diagram is depicted in Fig. 6.

The feedback devices are numbered as in Fig. 4. The resistor R4 was omitted as its function is held by the low input impedance of the amplifier. The calculation of the feedback devices has been processed according to (Hajek and Sedlacek 2002) and the equations (1) to (6).

As it was mentioned above, low tolerance capacitors are needed to construct this band pass filter and therefore it is helpful to use the same capacitors at the positions C1 and C2, mainly because the quotes of such devices are quite weak.

With the requirements enlisted in the Table 2 and when R3 = 10 kΩ, the following calculations should be processed first:

$$\beta = \frac{C_2}{C_1} = \frac{10^{-8}}{10^{-8}} = 1 \qquad (7)$$

$$\gamma = \frac{R_4}{R_3} = \frac{260}{10^4} = 0.026 \qquad (8)$$

Now, the parameter α can be calculated, using the equations (1), (5), (7) and (8). By substituting, the equation (5) can be expressed as follows:

$$4.32 = \frac{\sqrt{1\cdot\alpha}}{1\cdot(1-\alpha\cdot 0.026)+1} \Rightarrow \alpha = 29.17 \qquad (9)$$

Now the values of R2, Rx and Ry must be calculated. In the following calculations, the resistors Rx and Ry can be replaced by the resistor R1 that represents their parallel combination. According to (Hajek and Sedlacek 2002), when the parameter α is known, the values of R1 and R2 can be calculated by means of the following equations with the aid of a fictive resistance R. Provided the parameter β = 1, the capacitors C1 and C2 may be replaced by a fictive capacity C.

$$C = C_1 = C_2 = 10^{-8}\ [F] \qquad (10)$$

$$R = \frac{1}{2\pi F_0 C} = \frac{1}{2\pi\cdot 1000\cdot 10^{-8}} \cong 15.92\ [k\Omega] \qquad (11)$$

$$R_1 = \frac{R}{\sqrt{\alpha}} \cong 2.95\ [k\Omega] \qquad (12)$$

$$R_2 = R\sqrt{\alpha} \cong 85.96\ [k\Omega] \qquad (13)$$

The combination of resistors Rx and Ry obtain a suitable attenuation at the input of the circuit, as its gain at the centre frequency F_0 is defined by the feedback devices and reaches quite a high value:

$$K(F_0) = (1+\gamma)\sqrt{\alpha\beta Q} \cong 24.10 \qquad (14)$$

The gain of the circuit can be now reduced by applying the resistors Rx and Ry to the desired level $G_0 = 2$:

$$R_x = \frac{R_1 K(F_0)}{G_0} \cong 35.5\ [k\Omega] \qquad (15)$$

$$R_y = \frac{R_x R}{R_x - R} \cong 3.21\ [k\Omega] \qquad (16)$$

All device values are defined now. Because the resistors must be chosen from standardized series, the values of some parts must have been realized by parallel connection of standardized devices. The list of devices used in the simulation of the circuit depicted in Fig. 6 is enlisted in the Table 3.

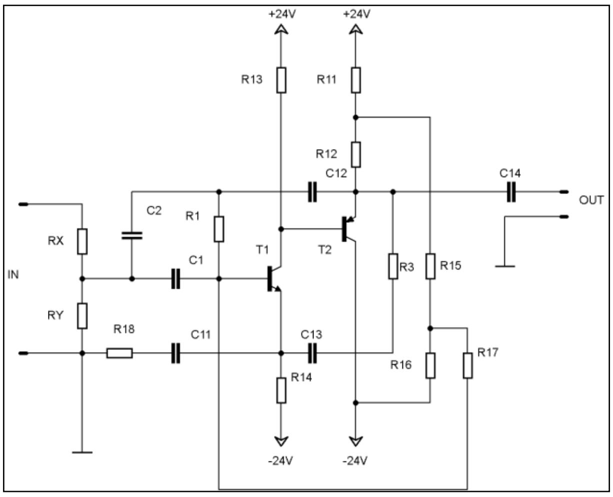


Figure 6: Final Circuit

Table 3: Calculated Values of the Devices

Devices	Values	
	Calculated	Applied
C1	10 nF	10 nF
C2	10 nF	10 nF
C11		10 µF
C12		10 µF
C13		10 µF
C14		10 µF
Rx	35.5 kΩ	47 kΩ || 150 kΩ
Ry	3.21 kΩ	3.3 kΩ || 120 kΩ
R2	85.96 kΩ	100 kΩ || 620 kΩ
R3	10 kΩ	10 kΩ
R11		1.8 kΩ
R12		680 Ω
R13		47 kΩ
R14		2.2 kΩ
R15		330 kΩ[1)]
R16		27 kΩ[1)]
R17		470 kΩ
R18		330 Ω

[1)] Selected according to simulation results of the circuit from Fig. 5.

SIMULATION

The circuit depicted in Fig. 6 has been simulated in the simulation software using SPICE libraries and algorithms. The values of the devices were set according to Table 3. The real parameters of the amplifier, mainly the limited open loop gain and inconsiderable input and output impedances reflected on the result that is included in the Table 4.

Table 4: Simulation Results for the Circuit from Fig. 6 with the Device Values from Table 3

Parameter	Value	Comment
F_0	1.0025 kHz	Fully meets the requirements
G_0	1.53	Too small; the required value is 2
Q	3.3	Too small; the required value is 4.32

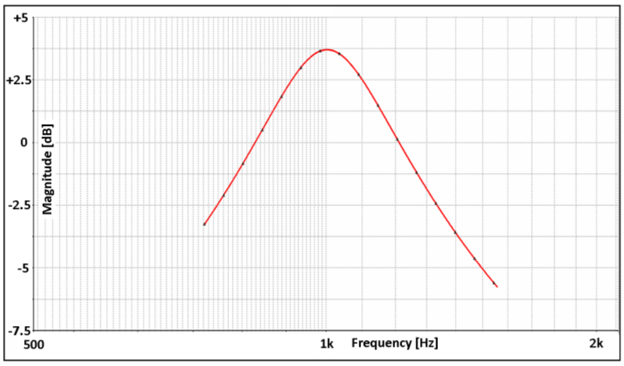


Figure 7: Frequency Response of the Simulated Circuit

Corrections

According to the simulation results it was decided not to change the values of C1, C2, Rx, Ry and R2 as the centre frequency meets the requirements as well. On the other hand, the gain of the circuit together with its quality factor can be increased by decreasing the value of R3. Therefore, for the purposes of simulation, the value of R3 has been decreased to 4.7 kΩ and the simulation was run again. With this value, the Q factor of the filter increased to the value as high as 7.4. The proper value of R3 has been estimated by means of linear interpolation. If the Q factor for R3 = 10 kΩ is 3.3 and for R3 = 4.7 kΩ is 7.4, it can roughly be estimated that the value changes with the following ratio:

$$\frac{\Delta Q}{\Delta R_3} = 1.056 \tag{17}$$

As the required value of the Q factor is 4.32, the value of 7.5 kΩ has been estimated to be used at the position of R3.

Finally obtained parameters

The parameters obtained by the final simulation using the correction on the value of R3 are enlisted in the Table 5. These parameters were found satisfying for the construction of the third octave equalizer with the topology according to Fig. 3. The frequency response of the circuit is depicted in Fig. 8. It also shows how the information on the Q-factor has been obtained: the bandwidth for the sag of -3 dB has been measured and the Q factor has been calculated according to (18), where F_0 is the centre frequency and B width of the band.

$$Q = \frac{F_0}{B} \tag{18}$$

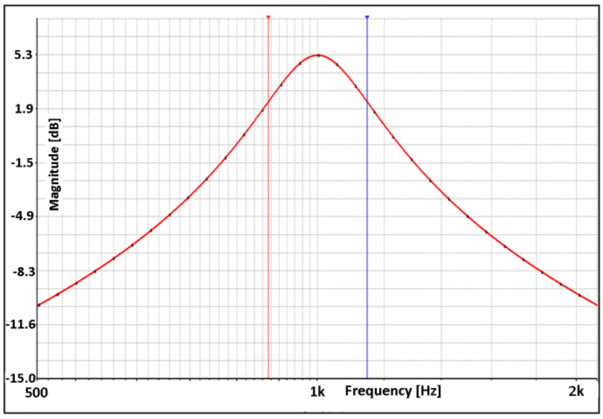


Figure 8: Frequency Response of the Simulated Circuit after Corrections

Table 5: Simulation results for the circuit from Fig. 6 with the device values from Table 3 after correction applied to the value of R3

Parameter	Value	Comment
F_0	1.0025 kHz	Fully meets the requirements
G_0	1.85	Almost meets the requirements
Q	4.38	Meets the requirements

After the final simulation it is obvious that to achieve the required gain G_0 more complex changes should have been done. On the other hand, the achieved Q factor 4.38 quite nice fits the required value 4.32.

What the simulation does not tell us

The parameters of the simulated circuit meet the requirements quite well. On the other hand as there are large differences between the discrete transistors, it is expected than in the final construction, the following parameters will be a subject of adjustment:

a) When the gain of the circuit changes due to variability of the parameters of the transistors, adjustment of R3 may be required.

b) The total bias of the circuit may vary according to the parameters of the transistors. The ratio of R11 : R12 may be required to be adjusted.

Both deviations may be adjusted by small trimming resistors, one connected in series with R3 and the second one connected between R11 and R12.

Concerning the centre frequency shift, as the open loop bandwidth of the circuit is quite large, it is not expected that there would exist a need to compensate the frequency response of the circuit.

CONCLUSIONS

This paper provides a description on the design of a band pass filter for a third octave equalizer with the aid SPICE algorithms and libraries. Because the construction of the third octave equalizer requires 64 such filters, there was an effort to decrease the price of the circuit's realization. Therefore it was decided to create the circuit on the basis of the amplifier consisting of two directly coupled transistors that can be purchased for very low prices, when ordered in bulk series. Because such amplifiers do not exhibit negligible input and output impedances and their open-loop gain is also limited, the aid of simulation software was required to tune the values of the devices used in the circuit in that way so it met the initial requirements. By repeating the design steps described in this paper, also the band pass filters for other bands can be designed. For this purpose the AC Analysis is a very helpful instrument, because not only the modular characteristics, but also the phase ones must be aligned in order to achieve the optimal performance of the whole device.

The most advantageous issue on the hereby presented design is its low price. The transistors used in this design are very cheap when ordered in bulk series. As the number of filters used in the device is considerably high, the price of the production can be significantly reduced.

The main disadvantage consists in insufficient gain at the nominal frequency, which is, according to the simulation, 5.3 dB, while the value of 6.0 dB was required. Once the shape of the transfer function is correct, this is not a critical issue as the signal level can be adjusted by front-end or back-end block of the device.

Acknowledgements

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme project No. LO1303 (MSMT-7778/2014).

REFERENCES

Biolek, D. 2003. *Solving of electrical circuits [Resime elektronicke obvody]*. BEN – Technicka literatura. ISBN 80-7300-125-X.

Bogdanowicz, A. 2011. “SPICE Circuit Simulator Named IEEE Milestone”. *The institute*. IEEE.

Hajek, K., Sedlacek, J. 2002. *Frequency filters [Kmitočtové filtry]*. BEN – Technická literatura. ISBN 80-7300-023-7

Smetana, P. 2015. *Design and construction of an analog equalizer [Návrh a konstrukce analogového ekvalizéru]*. University of West Bohemia. Diploma Theses.

Vladimirescu, A. 1994. *The SPICE book*. John Willey & Sons. ISBN 978-0471609261.

AUTHOR BIOGRAPHY

MARTIN POSPISILIK was born in Přílepy, Czech Republic. He reached his master degree at the Czech Technical University in Prague in the field of Microelectronics in 2008. Since 2013, after finishing his Ph.D. work focused on a construction of the Autonomous monitoring system, he became an assistant professor at the Tomas Bata University in Zlin, focused on communication systems and electromagnetic compatibility of electronic components. His e-mail is: pospisilik@fai.utb.cz

LQ DIGITAL CONTROL OF BALL & PLATE SYSTEM

Lubos Spacek, Vladimir Bobal and Jiri Vojtesek
Tomas Bata University in Zlín
Faculty of Applied Informatics
Nad Stráněmi 4511
760 05 Zlín
Czech Republic
E-mail: lspacek@fai.utb.cz

KEYWORDS
Ball & Plate, LQ digital control, 2DOF controller, spectral factorization.

ABSTRACT

This paper proposes the design of linear quadratic (LQ) digital controller for Ball & Plate model and 2DOF structure of the controller. Unknown parameters of the controller are determined with the help of polynomial approach to controller design. Semi-optimal solution is obtained using minimization of linear quadratic criterion. Spectral factorization with the aid of the Polynomial Toolbox for MATLAB was used for minimization of this LQ criterion. Additional poles of characteristic polynomial are placed so that the process is subtle and without sudden changes in controller output. Results have shown that the controller is able to stabilize the ball in desired position on the plate, reject external disturbances and follow reference path without much effort. Controller was designed for step changing and harmonic reference signal to further examine its capabilities.

INTRODUCTION

The Ball & Plate model is system with two inputs and two outputs. It has integrating properties, hence it can be considered unstable. This paper deals with controller design for this system using polynomial approach, because it simplifies the design problem to operations on algebraic polynomial (Diophantine) equations (Kučera 1993). Minimization of linear quadratic (LQ) criterion is used to derive controller parameters, which leads to semi-optimal solution (half of poles of characteristic polynomial have to be user-defined (Bobál et al. 2005)). This is particulary useful because it is quite challenging to place multiple user-defined poles. This process is applied to 2 degrees of freedom (2DOF) controller structure, which provides separation of feed-back part (responsible for stabilization and disturbance rejection) and feed-forward part (responsible for reference tracking) (Matušů and Prokop 2013). The PID/PSD control in closed-loop feedback structure was applied in (Jadlovská et al. 2009), where Butterworth, Graham-Lathrop and Naslin's methods were used for calculating controller parameters. A double feedback loop structure based on fuzzy logic is tested in (Wang et al. 2007). Fuzzy supervision and sliding control are proposed in (Moarref et al. 2008) and a non-linear switching is described in (Tian et al. 2006).

The paper is organized as follows. A brief description of mathematical model of the Ball & Plate structure is in Section 2. The design of LQ controller is shown in Section 3. Section 4 contains results of simulation and Section 5 concludes the paper.

BALL & PLATE MATHEMATICAL MODEL

A rough scheme of Ball & Plate model is presented in Figure 1. The derivation of system equations makes use of general form of Euler-Lagrange equation of the second kind (Rumyantsev 1994):

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} + \frac{\partial V}{\partial q_i} = Q_i \tag{1}$$

where T is kinetic energy of the system, V is potential energy, Q_i is i-th generalized force and q_i is i-th generalized coordinate. It is assumed that servomotor used for tilting the plate is described by first-order transfer function G_m with MATLAB units sent to servomotors circuit as input and actual angle of the plate as output:

$$G_m(s) = \frac{K_m}{\tau_m s + 1} \tag{2}$$

where K_m = 0.1878 and τ_m = 0.187 are gain and time constants of the motor respectively. These constants were obtained from real model's manual pages (Humusoft 2006).

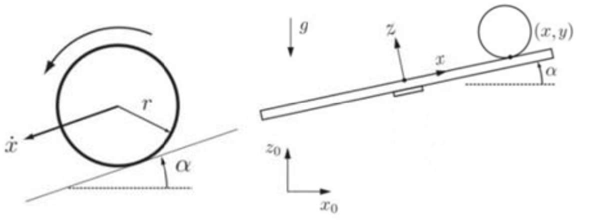


Figure 1: Ball & Plate scheme (Nokhbeh et al. 2011)

The system has only 2 generalized coordinates in total (ball position coordinates x and y), because plate angles are direct result of transfer function (2). Also the only external force acting on the system is gravitational force (friction is neglected for the sake of simplification). The

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

derivation of specific equations from (1) is not the purpose of this paper, thus only final result will be presented. This result consists of a system of 2 ordinary second-order differential equations:

$$x: \left(m + \frac{I_b}{r^2}\right)\ddot{x} - m\left(\dot{\alpha}\dot{\beta}y + \dot{\alpha}^2 x\right) + mg\sin\alpha = 0 \quad (3)$$

$$y: \left(m + \frac{I_b}{r^2}\right)\ddot{y} - m\left(\dot{\alpha}\dot{\beta}x + \dot{\beta}^2 y\right) + mg\sin\beta = 0 \quad (4)$$

where *m*, *r* and I_b are mass, radius and moment of inertia of the ball respectively, *g* is gravitational acceleration, *α* and *β* are plate angles (*α* changes *x* coordinate and *β* changes *y* coordinate), $\dot{\alpha}$ and $\dot{\beta}$ are first time derivatives of plate angles, *x* and *y* are coordinates of the ball from center of the plate and $\ddot{x}$, $\ddot{y}$ are second time derivatives of ball coordinates.

Linearized Model

For small angles of the plate, one can write $\sin\alpha \approx \alpha$ and $\sin\beta \approx \beta$. It is also assumed that the rate of change in plate inclination is small around the linearization point, thus $\dot{\alpha}\dot{\beta} \approx 0$, $\dot{\alpha}^2 \approx 0$ and $\dot{\beta}^2 \approx 0$. The moment of inertia of a hollow sphere (spherical shell) can be ideally expressed as $I_b = \frac{2}{3}mr^2$. These simplifications applied to (3) and (4) result in

$$x: \ \ddot{x} = K_b\alpha \quad (5)$$

$$y: \ \ddot{y} = K_b\beta \quad (6)$$

where K_b is constant dependent only on the gravitational acceleration *g* and the type of ball. The two dimensional problem is considered to be symmetric (see (3) and (4)), thus it is possible to express the mathematical model (by merging (2) with (5) or (6)) in one continuous transfer function *G(s)* with generalized coordinate as output *Y*(*s*) and generalized angle as input *U*(*s*):

$$G(s) = \frac{Y(s)}{U(s)} = \frac{K}{s^2\left(\tau_m s + 1\right)} = \frac{K}{\tau_m s^3 + s^2} \quad (7)$$

where $K = K_b K_m C_x$ is velocity gain of the integrating system ($C_x = 5\ m^{-1}$ is conversion coefficient from meters to normalized coordinates).
Equation (7) can be generaly discretized into:

$$G(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} = \frac{b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}} \quad (8)$$

where $B(z^{-1})$ and $A(z^{-1})$ are polynomials with unknown coefficients. Because the Ball & Plate model has double integrator, discrete transfer function (8) can be simplified as follows:

$$G(z^{-1}) = \frac{b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}}{(1 - z^{-1})^2(1 - c_1 z^{-1})} \quad (9)$$

2DOF LQ CONTROLLER DESIGN

Control Law

The controller is designed for two degree of freedom (2DOF) closed-loop control system shown in Figure 2, where *G* is controlled plant, C_f and C_b are feed-forward and feed-back parts of the controller respectively, $1/K(z^{-1}) = 1/(1 - z^{-1})$ is the summation part of the controller (it is extracted from denominators of C_f and C_b for practical reasons), *w*(*k*) is reference signal, *y*(*k*) is output of the system, *u*(*k*) is output of the controller, *n*(*k*) is load disturbance and *v*(*k*) is disturbance signal. It is assumed that no disturbances act on the system. This is obviously not true for real system, but it simplifies the design and structure of the controller.

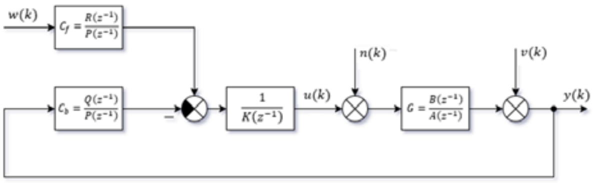

w(k)
$C_f = \frac{R(z^{-1})}{P(z^{-1})}$
n(k)
v(k)
$C_b = \frac{Q(z^{-1})}{P(z^{-1})}$
–
$\frac{1}{K(z^{-1})}$
u(k)
$G = \frac{B(z^{-1})}{A(z^{-1})}$
y(k)

Figure 2: Structure of 2DOF controller

As mentioned, the controller is designed using polynomial approach. By taking signals from Figure 2 in their discrete forms (and omitting z^{-1} in polynomials' notation), one can write a relation between reference signal and output of the system:

$$Y(z^{-1}) = \frac{BR}{AKP + BQ}W(z^{-1}) \quad (10)$$

The characteristic polynomial $D(z^{-1})$ can be extracted from (10) creating a Diophantine equation:

$$D = AKP + BQ \quad (11)$$

All polynomials in transfer functions will be called by their respective letter from now on, because omitting the term "(z^{-1})" will simplify the notation. Degree of polynomials *Q*, *R* and *P* can be obtained by determining the degree of the characteristic polynomial *D*, as described in (Bobál et al. 2005), from where it should be 6 for this specific case:

$$D = \sum_{i=0}^{6} d_i z^{-i} \quad (12)$$

Thus controllers C_b and C_f are

$$C_b(z^{-1}) = \frac{Q}{P} = \frac{q_0 + q_1 z^{-1} + q_2 z^{-2} + q_3 z^{-3}}{1 + p_1 z^{-1} + p_2 z^{-2}} \quad (13)$$

$$C_f(z^{-1}) = \frac{R}{P} = \frac{r_0}{1 + p_1 z^{-1} + p_2 z^{-2}} \quad (14)$$

where *Q* and *P* are polynomials with unknown coefficients, computed from (11) by method of undetermined coefficients. Polynomial *R* has one unknown coefficient r_0, which can be calculated for step-changing signal (see (Bobál et al. 2005)) as:

$$r_0 = \frac{d_0 + d_1 + d_2 + d_3 + d_4 + d_5 + d_6}{b_1 + b_2 + b_3} = \sum_{i=0}^{3} q_i \qquad (15)$$

In the case where reference signal is not step-changing, but harmonic, the polynomial R will be of higher degree and another Diophantine equation has to be solved:

$$SD_W + BR = D \qquad (16)$$

where S is an auxiliary polynomial not needed in controller parameters and D_w is denominator of harmonic reference signal $D_w = 1 - 2z^{-1}\cos(\omega T_0) + z^{-2}$, where ω is its angular frequency and T_0 is sampling period.

Minimization of LQ Criterion

A semi-optimal solution can be obtained by minimizing linear quadratic (LQ) criterion, which is closely described in (Bobál et al. 2005):

$$J = \sum_{k=0}^{\infty} \left\{ \left[e(k)\right]^2 + q_u \left[u(k)\right]^2 \right\} \qquad (17)$$

where $e(k) = w(k) - y(k)$ is error, $u(k)$ is controller output and q_u is penalization constant, which influences the controller output during minimization process.

According to (Bobál et al. 2005), this criterion can be minimized for input-output description of the model by applying spectral factorization on the following equation:

$$A(z^{-1}) q_u A(z) + B(z^{-1}) B(z) = D(z^{-1}) \delta D(z) \qquad (18)$$

where δ is chosen so that coefficient $d_0 = 0$ for the sake of simplification and $A(z)$, $B(z)$, $D(z)$ are conjugate polynomials of their respective counterparts. There is no analytical solution of spectral factorization for polynomials with degree 3 or higher, thus it has to be solved numerically by iterative methods (A is 3rd degree polynomial). The Polynomial Toolbox for MATLAB (Šebek 2014) contains tools for solving spectral factorization. The result of spectral factorization in this problem offers 3 roots of characteristic polynomial (12) that are optimal. Remaining 3 roots (poles) have to be user-defined. For a fully optimal solution, these poles can be placed to zero, but they are placed closer to a unit circle to make the controller more robust and its output properly bounded. Polynomial (12) can be now obtained and unknown coefficients of polynomial Q, P and R computed from (11) and (15) or (16).

RESULTS

It is important to note that controlled model in simulation was non-linear model described in (3) and (4). Its linearized form was used only for the design of the controller. Transfer function of the system is obtained after parameters K and τ_m are introduced into (7):

$$G(s) = \frac{K_b}{s^2} \frac{K_m}{\tau_m s + 1} C_x = \frac{-5.0706}{s^2 \left(0.187 s + 1\right)} \qquad (19)$$

Transfer function (19) can be discretized for the sampling period $T_s = 0.1$s:

$$G(z^{-1}) = \frac{0.00396 z^{-1} + 0.01394 z^{-2} + 0.00304 z^{-3}}{1 - 2.5871 z^{-1} + 2.1743 z^{-2} - 0.5871 z^{-3}} \qquad (20)$$

The result of spectral factorization of (18) for $q_u = 1$ are 3 optimal poles 0.8477 ± 0.1409i and 0.5821. User-defined poles were chosen to be 0.8, 0.8 and 0.8. Controller parameters for step-changing reference signal were calculated from (11) and (15) and substituted into (13) and (14):

$$C_b(z^{-1}) = \frac{-3.0311 + 7.4940 z^{-1} - 6.0558 z^{-2} + 1.5860 z^{-3}}{1 - 1.1023 z^{-1} + 0.3830 z^{-2}} \qquad (21)$$

$$C_f(z^{-1}) = \frac{-0.0069}{1 - 1.1023 z^{-1} + 0.3830 z^{-2}} \qquad (22)$$

For harmonic reference signal (with period 5s), the polynomial R is 1st degree polynomial and feedforward part of the controller (22) has the following form:

$$C_f(z^{-1}) = \frac{0.0438 - 0.0529 z^{-1}}{1 - 1.1023 z^{-1} + 0.3830 z^{-2}} \qquad (23)$$

Figure 3 and Figure 4 show step reference tracking capabilities of designed controller. Because the Ball & Plate system has integrating properties, the output of the controller is zero when the error is zero (although this would not be true if an unmeasurable load disturbance is present in the process – e.g. errors of motors).

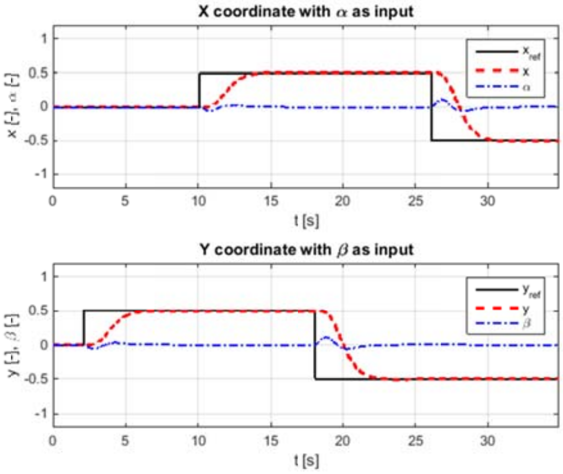


Figure 3: Step reference tracking

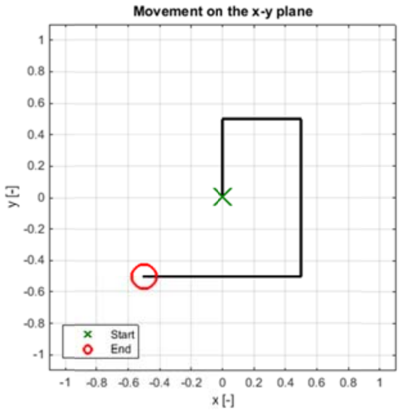


Figure 4: Step reference tracking on x-y plane

Figure 5 and Figure 6 show ability of controller to reject disturbances. Introduced disturbances were in the form of steps and it can be seen that the controller swiftly reacts to the disturbance and stabilizes the ball. Faster responses could lead to large changes in controller output, which is not appropriate in this kind of system.

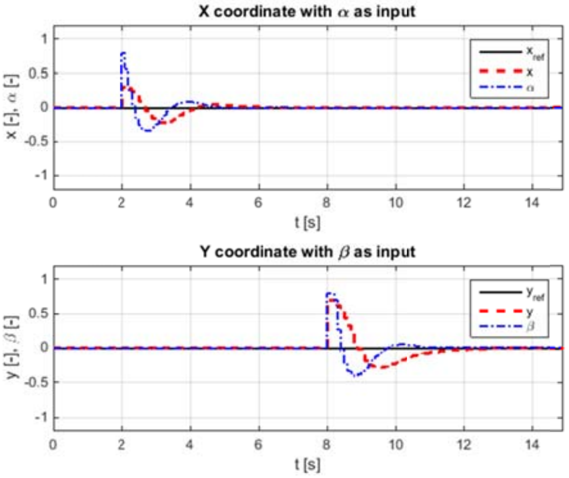


Figure 5: Step disturbance rejection

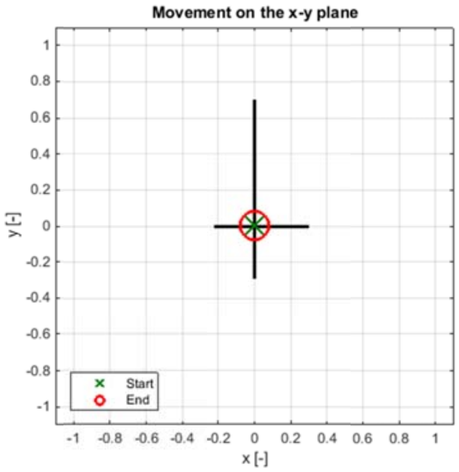


Figure 6: Step disturbance rejection on x-y plane

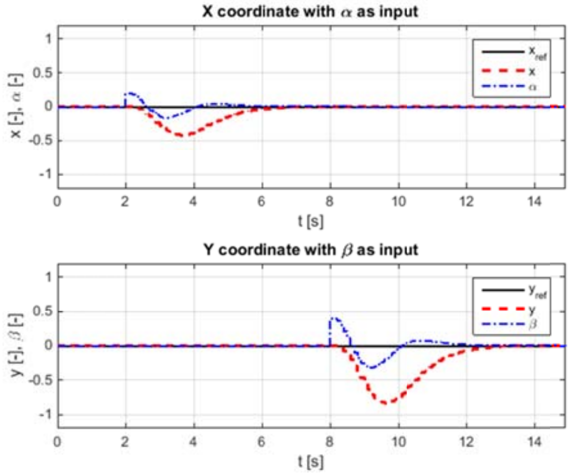


Figure 7: Step load disturbance rejection

Figure 7 and Figure 8 show rejection of disturbances applied directly to the output of the controller instead of ball's position, which simulates errors of controller. These load disturbances are also introduced as steps.

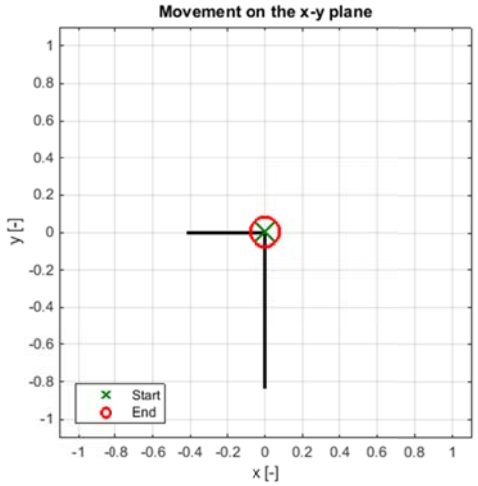


Figure 8: Step load disturbance rejection on x-y plane

Figure 9 and Figure 11 show circular reference tracking with controller designed for harmonic reference signal in (23). If the controller was designed only for step-changing reference signal, a phase lag would be present between reference harmonic signal and ball's position. Although the ball would still follow circular path for reference signal with low frequency. It would experience an amplitude reduction for higher frequencies of reference signal and the circular path would have smaller radius than desired.

A simple graphical user interface (GUI) was designed to provide more user-friendly control over the nonlinear model while testing control algorithms (Figure 10). It allows to choose the type of ball (sphere or spherical shell) and the type of reference value (manual point, circle, maze reference or custom). Relevant information is displayed in plots, which speeds up the design process and testing.

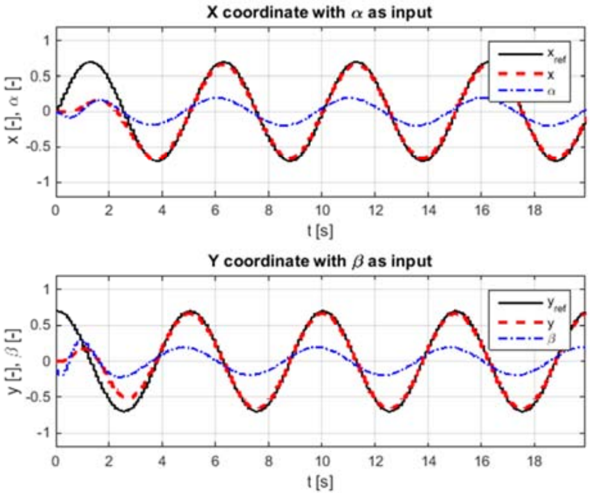
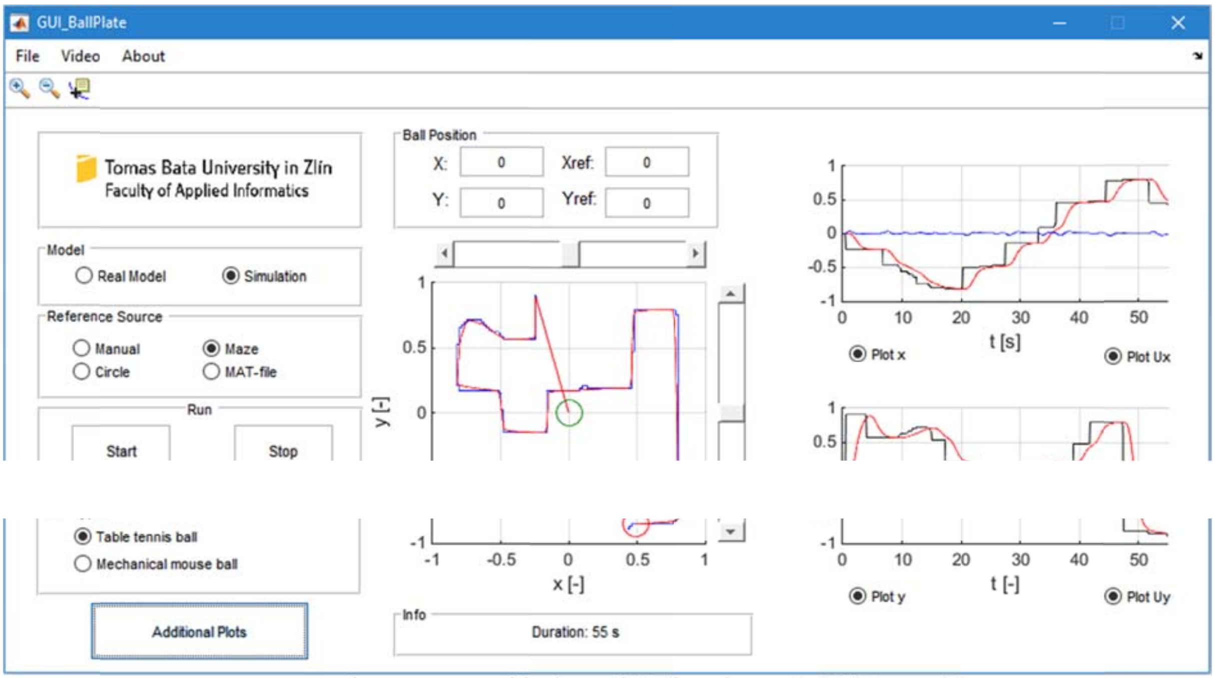


Figure 9: Circular reference tracking



Model: ○ Real Model ● Simulation

Reference Source: ○ Manual ● Maze ○ Circle ○ MAT-file

● Table tennis ball ○ Mechanical mouse ball

X: 0 Xref: 0

Y: 0 Yref: 0

● Plot x ● Plot Ux

● Plot y ● Plot Uy

Duration: 55 s

Figure 10: Graphical user interface for Ball & Plate model

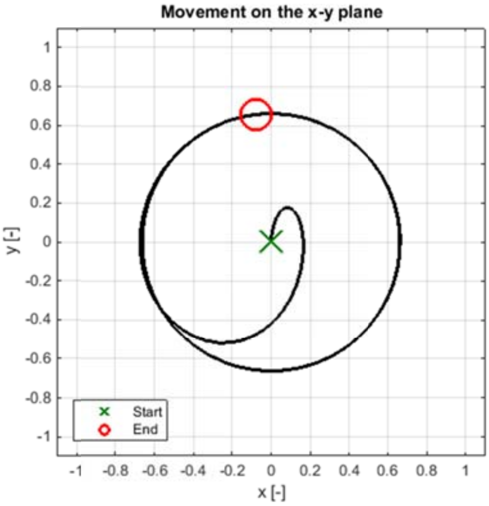


Figure 11: Circular reference tracking on x-y plane

CONCLUSION

The paper deals with design of linear quadratic (LQ) 2DOF controller for the Ball & Plate model. The controller was designed based on linearized mathematical model and polynomial approach for input/output form of the model. The presented method has been tested on computer simulation of nonlinear model of Ball & Plate structure. This model is quite sensitive to large changes in plate inclination (controller output). As a countermeasure, user-defined poles were placed in polynomial method algorithm near the unit circle, which resulted in subtle changes in plate inclination, but slowed the whole process. The minimization of LQ criterion provided rest of poles in an optimal solution, which successfully compensated system dynamics. The controller was designed for step-changing and also harmonic reference signal. It is able to reject disturbances acting on the system and successfully track desired reference value. A simple graphical user interface (GUI) was designed to act as a middlefinger between MATLAB/Simulink environment and the user.

ACKNOWLEDGMENT

This article was created with support of the Ministry of Education of the Czech Republic under grant IGA reg. n. IGA/FAI/2017/009.

REFERENCES

Bobál, V.; J. Böhm; J. Fessl; and J. Macháček. 2005. *Digital Self-tuning Controllers*. Springer-Verlag, London, 2005.

Humusoft. 2006." CE 151 Ball & Plate Apparatus User's Manual". Prague.

Jadlovská, A.; Š. Jajčišin; and R. Lonščák. 2009. "Modelling and PID Control Design of Nonlinear Educational Model Ball & Plate". In *Proceedings of the 17th International Conference on Process Control '09*. Štrbské Pleso, Slovakia, 475-483.

Kučera, V. 1993. *"Diophantine equations in control - a survey"*. Automatica, vol. 29, 1361-1375.

Matušů, R. and R. Prokop. 2013. "Algebraic design of controllers for two-degree-of-freedom control structure". In *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 7, 630-637.

Moarref, M.; M. Saadat; and G. Vossoughi. 2008. "Mechatronic design and position control of a novel ball and plate system". In *16th Mediterranean Conference on Control and Automation Congress Centre*, Ajaccio, France.

Nokhbeh, M.; D. Khashabi; and H.A. Talebi. 2011. ”Modelling and Control of Ball-Plate System”. Amirkabir University of Technology, Tehran, Iran.

Rumyantsev, V.V. 1994. ”Lagrange equations“. In *Encyclopaedia of Mathematics,* vol. 10.

Šebek, M. 2014. ”Polynomial Toolbox for MATLAB”. Version 3.0. PolyX, Prague.

Tian, Y.; M. Bai; and J. Su. 2006. ”A non-linear switching controller for ball and plate system”. In *International Journal of Modeling, Identification and Control*, vol. 1, no.3, 177–182.

Wang, H.; Y. Tian; Z. Sui; X. Zhang; and C. Ding. 2007. ”Tracking control of ball and plate system with a double feedback loop structure”. In *Proc. 2007 IEEE International Conference on Modeling and Automation*, Harbin, China, 2007.

AUTHOR BIOGRAPHIES

ĽUBOŠ SPAČEK studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master’s degree in Automatic Control and Informatics in 2016. He currently attends PhD study at the Department of Process Control. His e-mail address is `lspacek@fai.utb.cz`.

VLADIMÍR BOBÁL graduated in 1966 from the Brno University of Technology, Czech Republic. He received his Ph.D. degree in Technical Cybernetics at Institute of Technical Cybernetics, Slovak Academy of Sciences, Bratislava, Slovak Republic. He is now Professor at the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlín. His research interests are adaptive and predictive control, system identification and CAD for automatic control systems. You can contact him on e-mail address `bobal@fai.utb.cz`.

JIŘÍ VOJTĚŠEK was born in Zlin, Czech Republic in 1979. He studied at Tomas Bata University in Zlin, Czech Republic, where he received his M.Sc. degree in Automation and control in 2002. In 2007 he obtained Ph.D. degree in Technical cybernetics at Tomas Bata University in Zlin. In the year 2015 he became associate professor. He now works at the Department of Process Control, Faculty of Applied Informatics of the Tomas Bata University in Zlin, Czech Republic. His research interests are modeling and simulation of continuous-time chemical processes, polynomial methods, optimal, adaptive and nonlinear control. You can contact him on e-mail address `vojtesek@fai.utb.cz`.

AN EMBEDDED SYSTEM IMPLEMENTATION OF A PREDICTIVE ALGORITHM FOR A BIOPROCESS

Florin Stîngă*, Marius Marian**, Valentin Kese***, Lucian Bărbulescu** and Emil Petre*
*Department of Automation and Electronics
**Department of Computers and Information Technologies
Faculty of Automation, Computers and Electronics
University of Craiova
***Softronic Group
Craiova, Romania
E-mail: [florin, epetre]@automation.ucv.ro,
E-mail: [marius.marian, lucian.barbulescu]@cs.ucv.ro,
E-mail: kesevalentin@gmail.com

KEYWORDS
Bioprocess, Model predictive control, embedded implementation.

INTRODUCTION

The modeling and control of biotechnological processes is an actual and challenging problem due to their complicated structure. It is well known that these processes are dealing with living organisms that evolve over a long time or at the smallest change in their environment can become highly sensitive. Therefore the bioprocesses are characterized by highly nonlinear and uncertain dynamics, and their mathematical model is complex (G. Bastin and D. Dochain 1990; O. Bernard et al. 2011; F. Mairet et al. 2011). One of these processes whose importance resides in strict environmental rules is the wastewater treatment process; these rules are imposed in order to limit the quantity of toxic matter released in industrial and urban effluents. Nevertheless, its main drawback is the production of carbon dioxide (CO2) and its easy destabilization to input variations. For CO_2 mitigation, a solution consist in the growth of some microalgae populations that by using light as source of energy are able to assimilate inorganic forms of carbon and to convert them into requisite organic substances for cellular functions, generating at the same time oxygen (O_2). In what concerns the control of these processes, during the last years, numerous control strategies were developed: linearizing feedback (I. Neria-González et al 2009), adaptive and robust-adaptive (D. Selisteanu et al. 2007), predictive control (S. Tebbani et al. 2014), so on. Moreover, the widespread use of the embedded systems, based on microcontrollers, offered the hardware support for testing and implementing such complex control algorithms.

The paper presents a solution for implementing a model predictive control (MPC) for a continuous photo-bioreactor used for the growth of some microalgae that have the ability to use CO_2 as carbon source and, together with the solar energy, to biosynthesize various components, generating O_2. A widely used software platform for modeling, simulating and then, getting the real-time code is MATLAB/Simulink environment with the Real-Time Workshop plug-in (a.k.a Automatic Code Generation) (MathWorks 2016). However, the use of this proprietary software, may be limiting due to their costs, the number of the necessary equivalent *embedded functions – microcontroller level* translate code, and, also by the classes of microcontrollers that are supported by this application.

The model-based predictive control has been adopted in industry as an effective control strategy due to its capabilities to generate an optimal control input, and also to tackle the constraints on states, outputs and inputs. The control strategies are based on solving on-line (or off-line), at each sampling instant, a mathematical optimization problem based on a dynamical model of the plant (M. Morari and J.H. Lee, 1999). Over the years, different predictive control strategies were proposed (E. F. Camacho and C. Bordons 2004), (Q. Mayne and. E. C. Kerrigan 2007), or for hybrid systems (M. Lazar 2006), directly related to the wider area of application were used. The complexity of the algorithm resides in a mathematical optimization problem for which the feasibility may be ensured at each sampling time. Large number of variables may be involved leading to considerable computation effort and larger times for the optimal solution. This disadvantage, in real time applications, can be overcome by efficient algorithms that provide the essential computational routines required. The implementation of the MPC strategy, consider that the optimization problem is solved by means of a Hildreth's procedure. The obtained C code is tested under the same considered hypothesis on real time-platforms.

MATHEMATICAL MODEL

We consider the following photoautotrophic process which describes the growth of the green alga *C.*

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

reinhardtii in a photobioreactor under light limiting conditions (F. Stinga and. E. Petre 2016):

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(x(t))u \\ y(t) &= h(x(t)) \end{aligned} \tag{1}$$

where $u = \begin{bmatrix} u_1 & u_2 \end{bmatrix}^T = \begin{bmatrix} D & G_{in}^{CO_2} \end{bmatrix}^T$,

$x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix}^T =$

$= \begin{bmatrix} X & C_{TIC} & C_{O_2} & y_{out}^{CO_2} & y_{out}^{O_2} \end{bmatrix}^T$,

$$f(x) = \begin{bmatrix} \langle \tau_x \rangle \\ -\langle \tau_{TIC} \rangle + N_{CO_2} \\ \langle \tau_{O_2} \rangle + N_{O_2} \\ -G_{out} - V_l N_{CO_2} \\ -G_{out} - V_l N_{O_2} \end{bmatrix}, \; g(x) = \begin{bmatrix} -x_1 & 0 \\ C_{TIC,i} - x_2 & 0 \\ -x_3 & 0 \\ 0 & \dfrac{RT_a}{PV_g} \\ 0 & \dfrac{RT_a c_3}{PV_g} \end{bmatrix},$$

$$h(x) = \begin{bmatrix} x_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_4 & 0 \end{bmatrix}, \; y = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T$$

where X is the biomass concentration, C_{TIC} is total inorganic carbon concentration, and C_{O_2} is the dissolved oxygen concentration, $y_{out}^{CO_2}$ and $y_{out}^{O_2}$ are the molar fractions of CO_2 and O_2 in the outlet gas, D is the dilution rate, $C_{TIC,i}$ is the concentrations of TIC in the feed, R is the universal gas constant, V_g and V_l are the gas and liquid volume of the photobioreactor, P is the total pressure in the gas phase, T_a is the temperature, $G_{in}^{CO_2}$ and $G_{in}^{O_2}$ are the CO_2 and O_2 feeding flow rates, respectively, and G_{out} are the total output flow rate. Also, the expressions of the global volumetric rate (τ_x), the local value of the irradiance ($I(z)$), TIC consumption rate (τ_{TIC}), the oxygen production rate (τ_{O_2}), the CO_2 and O_2 mass transfer rates to the liquid phase (N_{CO_2} and N_{O_2}), the total output flow rate, the feeding flow rate of the oxygen, are expressed as in (G.A. Ifrim et al. 2013; S. Tebbani et al. 2015; F. Stinga and. E. Petre 2016).

According with the defined outputs and inputs of the system, the control objective is to maintain biomass concentration in the photobioreactor at certain setpoints and to minimize the quantity of CO_2 in the output flux, that is $y = \begin{bmatrix} X & y_{out}^{CO_2} \end{bmatrix}^T$ using the dilution rate and the feeding flow rate of CO_2 as manipulated variables, i.e. $u = \begin{bmatrix} D & y_{out}^{CO_2} \end{bmatrix}^T$. Therefore we have a multivariable control problem with two outputs and two inputs.

The MPC strategy requires solving the following constraint optimization problem, and applying the optimal control action based on the receding horizon strategy (L. Wang 2009):

$$\begin{aligned} &\min_{\Delta U} \left(\frac{1}{2} \Delta U^T H \Delta U + f^T \Delta U \right) \\ &\textit{subject to } (M \Delta U \leq N) \\ &\begin{cases} \tilde{x}(k+1) = \tilde{A}\tilde{x}(k) + \tilde{B}\Delta u(k) \\ \tilde{y}(k) = \tilde{C}\tilde{x}(k) \end{cases} \end{aligned} \tag{2}$$

where, the model of the system is expressed in the difference of the state and control variables in order to obtain an integral action in predictive formulation, and:

$$H = \Phi^T \overline{Q} \Phi + \overline{H} \tag{3}$$

$$f = \Phi_F \tilde{x} - \Phi_R R^* \tag{4}$$

with, $\Phi_F = \Phi^T \overline{Q} F$, $\Phi_R = \Phi^T \overline{Q}$,

$$F = \begin{bmatrix} (\tilde{C}\tilde{A})^T & \left(\tilde{C}(\tilde{A})^2\right)^T & \dots & \left(\tilde{C}(\tilde{A})^{N_p}\right)^T \end{bmatrix}^T$$

$$\Phi = \begin{bmatrix} \tilde{C}\tilde{B} & o_{p\times m} & \cdots & o_{p\times m} \\ \tilde{C}\tilde{A}\tilde{B} & \tilde{C}\tilde{B} & \cdots & o_{p\times m} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{C}(\tilde{A})^{N_p-1}\tilde{B} & \tilde{C}(\tilde{A})^{N_p-2}\tilde{B} & \cdots & \tilde{C}(\tilde{A})^{N_p-N_c}\tilde{B} \end{bmatrix} \tag{5}$$

$\tilde{x}(k) = \begin{bmatrix} \Delta x(k)^T & y(k) \end{bmatrix}^T$, $\Delta x(k) = x(k) - x(k-1)$,

$\tilde{A} = \begin{bmatrix} A_d & o_{p\times n}^T \\ C_d A_d & I_{p\times p} \end{bmatrix}$, $\tilde{B} = \begin{bmatrix} B_d \\ C_d B_d \end{bmatrix}$, $\tilde{C} = \begin{bmatrix} o_{p\times n} & I_{p\times p} \end{bmatrix}$

$\Delta u(k) = u(k) - u(k-1)$,

$\Delta U = \begin{bmatrix} \Delta u(k) & \Delta u(k+1) & \dots & \Delta u(k+N_c-1) \end{bmatrix}^T$

N_c is the control horizon, R^* is column vector with $p \cdot N_p$ elements of set points, N_p is the prediction horizon, $\overline{Q}$ is positive definite error weight matrix, $\overline{H}$ is a $(m \times N_c) \times (m \times N_c)$ diagonal weight matrix used as tuning parameter for closed loop performance, m is the number of inputs, p is the number of outputs, n is the number of states, $o_{p\times n}$ is a zero matrix, $I_{p\times p}$ is the identity matrix, (A_d, B_d, C_d) are the matrices of the discrete-time state-space representation of the initial system (the discrete-time model of the system was obtained by means of Euler approximation), and $(x(k), u(k), y(k))$ are the state, input and output variables of the discrete-time model of the system,

$$M = \begin{bmatrix} -\Phi \\ -\Omega_1 \\ \Phi \\ \Omega_1 \end{bmatrix}, \; N = \begin{bmatrix} -Y_{\min} + Fx^a\left(k\right) \\ -\Delta U_{\min} + \Omega_2 u\left(k-1\right) \\ Y_{\max} - Fx^a\left(k\right) \\ \Delta U_{\max} - \Omega_2 u\left(k-1\right) \end{bmatrix}, \qquad (6)$$

Ω_1 is a $\left(m \cdot N_c\right) \times \left(m \cdot N_c\right)$ lower triangular identity matrix, Ω_2 is a column matrix with N_c identity matrix $I_{m\times m}$, $\Delta U_{\min}$, $\Delta U_{\max}$, $Y_{\min}$ and $Y_{\max}$ are column vectors with $\left(m \cdot N_c\right)$ elements of $\Delta u_{\min}$, $\Delta u_{\max}$, $y_{\min}$ and, respectively $y_{\max}$.

Remark 1: For the considered control strategy we use a linearizing model of the initial system (1), defined around certain equilibrium points (see F. Stinga and. E. Petre 2016).

The constrained programming problem (2) can be solved by using the different methods, such as: the projection method, the primal-dual method, the penalty functions, and so on.

In our implementation, we consider a primal-dual method, based on an element-by-element search which minimizes the objective function, expressed by the Hildtreth algorithm. At each complete Hildreth iteration, the minimization of the objective function (2), was obtained by adaptation of a single component λ_i of the Lagrange multiplier vector λ , so that (D.G. Luenberger 1969):

$$\lambda_i^{k+1} = \max\left(0, \omega_i^{k+1}\right) \qquad (7)$$

with

$$\omega_i^{k+1} = -\frac{1}{p_{ii}}\left[d_i + \sum_{j=1}^{i-1} p_{ij}\lambda_j^{k+1} + \sum_{j=i+1}^{n} p_{ij}\lambda_j^{k}\right] \qquad (8)$$

where p_{ij} is the *ijth* element in the matrix $P = MH^{-1}M^T$ and d_i is the *ith* element in the vector $D = N + MH^{-1}F$.

The optimal solution is expressed by the following relation, taking into account that the $\lambda_i \geq 0$, the optimal command sequence applied to the controlled system is given by:

$$\Delta U = -H^{-1}\left(F + M^T\lambda\right) \qquad (9)$$

A microcontroller implementation procedure of the above programming algorithm is described in the next section.

PREDICTIVE ALGORITHM IMPLEMENTATION

We started the effort of getting the C based implementation of the above predictive control algorithms taking into account two possible strategies.

The first strategy starts with the mathematical model of the system in MATLAB/Simulink and then the predictive control algorithm can be generated by means of: MATLAB functions, embedded MATLAB functions, and Simulink library blocks. In what concerns the actual implementation of the predictive control algorithms, one can use the computing capabilities of the associated language (matrix initialization and manipulations, plus large data manipulation). The Real Time Workshop (RTW) plug-in can then be used to generate the C language source code from the Simulink model.

There are a few important remarks concerning this first strategy. The above mentioned plug-in RTW is licensed under a commercial license therefore some limits may apply on its usage. In what concern the classes of microcontrollers for which C code can be automatically generated, the number of C compilers that are supported, how large is the integration of legacy code, what limits are concerning the model size, to what extent we have a compatibility of the model and the generated code, and many more.

A second possible strategy involves first the design and then writing down the equivalent C source code for the predictive algorithm. We have opted for this second strategy after carefully considering the above points. The first reason was portability of source code written for different microcontroller/microprocessor architectures. We wanted to have a predictive algorithm C-based implementation that could be as easily portable as possible between different hardware architectures. Second reason was the usage preference of the authors for a free compiler such as GNU C compiler (GCC). Having said all these, we have continued to use the data sets generated by MATLAB/Simulink for testing the source code that we will present further on.

In our implementation we have designed and developed a mini library of C functions for common matrix manipulations such as: addition, subtraction, multiplication, computing the inverse or the transpose of a matrix. Furthermore, additional functions were written for initializing a matrix, copying a matrix, raising to a power, multiplication with a scalar or a vector.

Getting into the implementation of the predictive algorithm, we have to discuss a few aspects. In conformity with the mathematical model, we started with the calculation of the time-independent coefficients: H - as one can see in relation (3), Φ_R, Φ_F as in relation (4), $\left(\tilde{A}, \tilde{B}, \tilde{C}, A_d, B_d, C_d\right)$ defined in the above section, $\left(\Phi, F\right)$ according to relation (5), M and N as in relation (6).

```
#define EYE(sz, m) memset(m, 0, sizeof(m)); \
                for (i = 0; i < sz; i++) \
                    m[i][i] = 1;
#define TRIL(sz, m) memset(m, 0, sizeof(m)); \
                for (i = 0; i < sz; i++) \
                  for (j = 0; j <= i; j++) \
                    m[i][j] = 1;
#define ZERO(m)    memset(m, 0 , sizeof(m));
#define INITM(sz, m1)  memset(m1, 0, sizeof(m1)); \
            for (i = 0; i < sz; i++) { \
             for (j = 0; j <= i; j++) {\
               m1[i][j] = -1; \
               m1[i+sz][j] = 1; \
             }\
            }
#define MAX(a, b)  (a > b ? (a) : (b))
double results[m], double DeltaU[Nc*m][1];
static double F[Np*p][x+p], Phi_F[Nc*m][x+p],
Phi_R[Nc*m][p], M[2*Nc*m+outcons*2*Np*p][Nc*m];
static double HInv[Nc*m][Nc*m],mHInv[Nc*m][Nc*m];
static double
mHInvMT[Nc*m][2*Nc*m+outcons*2*Np*p];
static double P[2*Nc*m + outcons*2*Np*p][2*Nc*m +
outcons* 2*Np*p];
```

```
double predc(double uLast[m], double r1[p], double
y[p], double dx[x])
{....
double A_tilda[x+p][x+p], B_tilda[x+p][m],
C_tilda[p][p+x], double C_tildaA_tildaP[p][x+p];
double Phi[np*p][nc*m], C_tildaA_tildaPB_tilda[p][m,
A_tildaP[x+p][x+p], PhiT[nc*m][np*p],
PhiTQ[nc*m][np*p], Phi_Phi[nc*m][nc*m]; dxy[n+p],
N[2*Nc*m+outcons*2*Np*p], f[Nc*m][1],
N[2*Nc*m+outcons*2*Np*p], Fdxy[Np*p][1],
eta[Nc*m][1], d[2*Nc*m+outcons*2*Np*p],
lambda[2*Nc*m + outcons*2*Np*m][1];
double lambda_p[2*Nc*m + outcons*2*Np*p][1],
lambda_m_lambda_p[2*Nc*m + outcons*2*Np*p][1],
lambda_m_lambda_pT[1][2*Nc*m + outcons*2*Np*p],
lambdaProd[1][1], a1, w;
int i, j, kk, km;
```

```
//Obtaining the vector F
copy(p, p+x, C_tilda, p, p+x, C_tildaA_tildaP, 0, 0);
   for (i = 0; i < Np; i++) {
multiply(p, p+x, p+x, p+x, C_tildaA_tildaP, A_tilda,
C_tildaA_tildaP);
copy(p, p+x, C_tildaA_tildaP, Np*p, x+p, F, i * p, 0);
copy(p, p+x, C_tildaA_tildaP, p, p+x, C_tildaA_tildaP,
0, 0); }
```

```
//Obtaining the matrix Φ
ZERO(Phi);
   for (i = 0; i < Np; i++) {
      for (j = 0; j < Nc; j++) {
         if (i < j) {
            break;
         } else if (i == j) {
multiply(p, p+x, p+x, m, C_tilda, B_tilda,
C_tildaA_tildaPB_tilda); }
else {
matPower(x+p, A_tilda, i-j, A_tildaP);
multiply(p, p+x, p+x, p+x, C_tilda, A_tildaP,
C_tildaA_tildaP);
multiply(p, p+x, p+x, m, C_tildaA_tildaP, B_tilda,
C_tildaA_tildaPB_tilda); }
copy(p, m, C_tildaA_tildaPB_tilda, Np*p, Nc*m, Phi, i *
p, j * m);}
copy(p, m, C_tildaA_tildaPB_tilda, Np*p, Nc*m, Phi, i *
p, j * m);
      }
   }
```

```
//Obtaining the matrix H
transpose(Np*p, Nc*m, Phi, PhiT);
multiply(Nc*m, Np*p, Np*p, Np*p, PhiT, Q, PhiTQ);
multiply(Nc*m, Np*p, Np*p, Nc*m, PhiTQ, Phi,
Phi_Phi);
multiply(Nc*m, Np*p, Np*p, x+p, PhiTQ, F, Phi_F);
for (i = 0; i < Nc*m; i++) {
      for (j = 0; j < p; j++) {
         Phi_R[i][j] = Phi_F[i][x + j];
      }
   }
EYE(Nc*m, eyeNcIn);
scalarMultiply(1, Nc*m, Nc*m, eyeNcIn, eyeNcInDiv);
add(Nc*m, Nc*m, Phi_Phi, eyeNcInDiv, H);
```

```
//Obtaining the matrix M
M13(Nc*m, M13);
   if (outcons) {
      for (i = 0; i < Np * p; i++) {
         for (j = 0; j < Nc * m; j++) {
M[2 * Nc * m + i][j] = -Phi[i][j];
M[2 * Nc * m + Np * p + i][j] = Phi[i][j];}}}
```

```
//Obtaining the matrix P
inverse(Nc*m, H, HInv);
scalarMultiply(-1, Nc*m, Nc*m, HInv, mHInv);
transpose(2*Nc*m + outcons*2*Np*p, Nc*m, M, MT);
multiply(Nc*m, Nc*m, Nc*m, 2*Nc*m +
outcons*2*Np*p,HInv, MT, HInvMT);
multiply(2*Nc*m + outcons*2*Np*p, Nc*m, Nc*m,
2*Nc*m + outcons*2*Np*p, M, HInvMT, P);
```

```
//Obtaining the vector f
memcpy(dxy, dx, x * sizeof(double));
memcpy(dxy+x, y, p * sizeof(double));
multiplyMatVec(Nc*m, x+p, x+p, Phi_F, dxy,
Phi_Fdxy);
multiplyMatVec(Nc*m, p, p, Phi_R, r1, Phi_Rr1);
sub(Nc*m, 1, Phi_Fdxy, Phi_Rr1, f);
multiplyMatVec(Np*p, x+p, x+p, f, dxy, Fdxy);
```

```
//Obtaining the vector N
for (j = 0; j < m; j++)
         {
      for (i = 0; i < Nc; i++)
            {
         N[i + j*Nc] = -umin[j] + uLast[j];
         N[i + j*Nc + m*Nc] = umax[j] - uLast[j]; }
         }

   if (outcons) {
      for (j = 0; j < p; j++)
         {
         for (i = 0; i < Np; i++)
         {
N[2*i + j + 2*Nc*m] = -ymin[j] + Fdxy[i*p+j][0];
N[2*i + j + (2*Nc*m+p*Np)] = ymax[j] -
Fdxy[i*p+j][0];}
         }}
```

```
//Implementation of the Hildreth's algorithm
multiply(Nc*m, Nc*m, Nc*m, 1, mHInv, f, eta);
   kk = 0;
   for (i = 0; i < 2*Nc*m + outcons*2*Np*p; i++) {
      sum = 0;
      for (j=0; j < Nc*m; j++) {
         sum += M[i][j] * eta1[j][0];
      }
      if (sum > N[i]) {
         kk = 1;
         break;
      }
   }
if (kk == 0) {
    for(i=0;i<m;i++)
    {
       results[i]=eta1[i][0];
    }}
   multiply(Nc*m, Nc*m, Nc*m, 1,HInv, f1, HInvf);
   multiply(2*Nc*m + outcons*2*Np*p, Nc*m, Nc*m, 1,
M, HInvf, d);
   for (i = 0; i < 2*Nc*m + outcons*2*Np*p; i++) {
      d[i][0] += N[i];
   }
   ZERO(lambda);
for (km = 0; km < 100; km ++) {
copy(2*Nc*m + outcons*2*Np*p, 1, lambda, 2*Nc*m +
outcons*2*Np*p, 1, lambda_p, 0, 0);
  for (i = 0; i < 2*Nc*m + outcons*2*Np*p; i++) {
        w = -P[i][i] * lambda[i][0] + d[i][0];
        for (j = 0; j < 2*Nc*m + outcons*2*Np*p ; j++)
{
            w += P[i][j] * lambda[j][0];
        }
        lambda[i][0] = MAX(0, -w / P[i][i]);
     }
sub(2*Nc*m + outcons*2*Np*p, 1, lambda, lambda_p,
lambda_m_lambda_p);
transpose(2*Nc*m + outcons*2*Np*p, 1,
lambda_m_lambda_p, lambda_m_lambda_pT);
multiply(1, 2*Nc*m + outcons*2*Np*p, 2*Nc*m +
outcons*2*Np*p, 1, lambda_m_lambda_pT,
lambda_m_lambda_p, lambdaProd);
      a1 = lambdaProd[0][0];
      if (a1 < 10e-8) {
         break;
      }}
multiply(Nc*m, 2*Nc*m + outcons*2*Np*p, 2*Nc*m +
outocns*2*Np*p, 1, mHInvMT, lambda,
mHInvMTlambda);
add(Nc*m, 1, eta1, mHInvMTlambda, DeltaU);
for (i=0; i<m;i++)
{
   results[i] = DeltaU[i][0];
}
return 0;
```

The parameters list of the ***predc*** function includes: $u(k-1), y(k), r(k)$, and $\Delta x(k)$.

This function is called iteratively in the *main()* function. It involves calculating the value of several parameters that change at every step based on state, input and output variables updates.

For defining constraints on the output variables the parameter *outcons* is used. The algorithm considers implicit constraints on the command variable. The complexity of the considered problem is influenced by two elements. First, it is the size of the two horizons: prediction and control, and second, by the number of system variables involved plus the number and type of constraints (hard or soft).

The final section of the function describes the implementation of the Hildreth's algorithm. In fact, the relations (7), (8), and (9) are translated in C level code. Also, the predictive control algorithm applying the principle of receding horizon strategy, such that only first m elements of ΔU are taken into account to form incremental optimal control. The rest of the elements are discarded, and at the next sampling instant a new control sequence is computed.

RESULTS

The considered bioprocess is a photoautotrophic growth of the green alga C. *reinhardtii* in a photobioreactor under specifying lighting conditions (see G.A. Ifrim et al 2013).

Also, for predictive algorithm the following input data are used (F. Stinga and E. Petre, 2016):

$$\tilde{A}=\begin{pmatrix} -0.03632 & 0 & 0 & 0 & 0 \\ -0.00049 & -0.09753 & 0 & 0.25117 & 0 \\ 0.00054 & 0 & -0.95000 & 0 & 1.08729 \\ 0 & 0.00969 & 0 & -4.36893 & -2.98859 \\ 0 & 0 & 0.18353 & -0.08390 & -4.14908 \end{pmatrix}$$

$$\tilde{B}=\begin{pmatrix} -0.42468 & 0 \\ 0.00036 & 0 \\ -0.00092 & 0 \\ 0 & 138.72761 \\ 0 & 38.84373 \end{pmatrix}, \quad \tilde{C}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

$n=5, \quad m=2,\ p=2 \quad , \quad T=0.3\,h. \quad ,$

$\bar{Q}=\begin{bmatrix} 100\cdot I_{N_p\times N_p} & 1\cdot I_{N_p\times N_p} \end{bmatrix} \quad , \quad \bar{H}=1\cdot I_{mNc\times mNc} \quad ,$

$\Delta u_{\min}=\{0,0\}$, $\Delta u_{\max}=\{0.1, 0.03\}$, $y_{\min}=\{0,0\}$, $y_{\max}=\{0.7, 0.06\}$.

The hardware platforms used for the implementation and testing of the described algorithm can be summarized as follows:

- PC platform: 2.4 GHz Intel i5 microprocessor and 4 GB of RAM memory, and two different microcontroller-based configurations:
- Atmel ARM Cortex M4 SAM4S16C: a 32-bit MCU, 120 MHz, 2KB of cache, 128 KB of embedded SRAM, 1 MB embedded flash memory, single-precision FPU;

- Atmel ARM Cortex M7 ATSAME70-XPLD Evaluation Board: a 32-bit MCU, 300 MHz, single- and double-precision FPU.

The software configurations include the following packages:

- 32-bit Windows 7 operating system
- MATLAB/SIMULINK 7.9.0
- Code::Blocks IDE 16.01
- GCC compiler v6.2
- AVR Atmel Studio IDE v6.2

Running the algorithm on different configurations hardware/software led to the following results. First, the simulation data obtained in MATLAB/Simulink was used as reference for all tests. The considered process outputs obtained after running the implemented algorithm on the Atmel ARM Cortex M7 ATSAME70-XPLD Evaluation Board and MATLAB/Simulink ($N_p = N_c = 5$) are presented in Figures 1 and 2.

The two figures present the evolution of the biomass concentration X and of the molar fraction of CO_2 , respectively, in two possible scenarios: with constraints on input variables (the blue line) and with constraints both on input and output variables (the red line).

Remark 2: Also, for the considered inputs of the system (the dilution rate and the feeding flow rate of CO_2), data set obtained for MATLAB environment and those generated by the microcontrollers, using ***predc*** function, coincide.

The tests yielded different execution times that are available in Table 1. The results are calculated by averaging the values returned by the real-time systems.

Obviously, the results obtained are influenced by the benchmark system. Large horizons of prediction and control, and also hard constrains both on command and output may lead to ever-increasing execution times.

These include the use of hardware-specific dedicated digital signal processing software libraries or improving the code for matrix manipulation. Another approach would be to take advantage of the compiler's levels of code optimization. Yet another approach would be to handle optimally the memory usage based on the specifics of the hardware platform used.

In these scenarios the data memory usage vary from 50.2% to 72.4%, instead the program memory usage is 8.2%. Certainly, the higher value from the data mamory usage is generated by the considered prediction and control horizons, and, also by the number of imposed constrained.

CONCLUSIONS

The paper presented an implementation of a C-based optimal control algorithm. The code was tested on two real time-platforms with different technical specifications. The obtained results are provided in order to demonstrate the effectiveness of the proposed implementation. We intend to continue working on optimizing the C implementation taking into account the specifics of target hardware platforms and also the software libraries available for the microcontrollers.

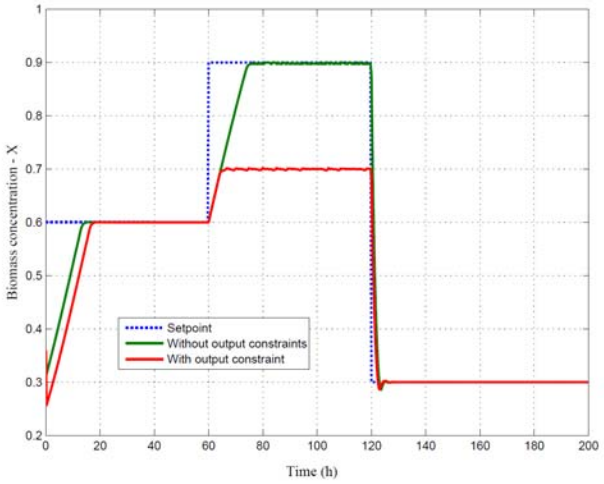


Figure 1: The time evolution of the first considered output of the system

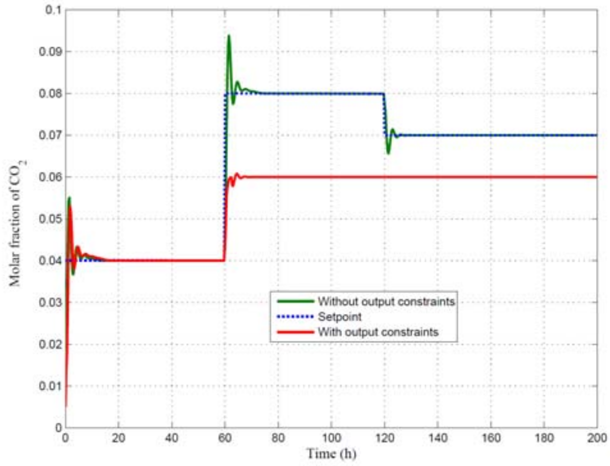


Figure 1: The time evolution of the second considered output of the system

TABLE I. THE EXECUTION TIMES

#	Implementation platform	N_p	N_c	Constraints	Measured time (sec.)
1	Atmel ARM Cortex M4	5	5	Only input	15×10^{-3}
				Both input and output	57.3×10^{-3}
2		8	5	Only input	38.2×10^{-3}
				Both input and output	133×10^{-3}
3	Atmel ARM Cortex M7	5	5	Only input	2.13×10^{-3}
				Both input and output	2.23×10^{-3}
4		8	5	Only input	2.25×10^{-3}
				Both input and output	5.27×10^{-3}

ACKNOWLEDGMENT

This work was supported by UEFISCDI, project BIOCON no. PN-II-PT-PCCA-2013-4-0070, 269/2014.

REFERENCES

G. Bastin and D. Dochain, On-line estimation and adaptive control of bioreactors, New York, NY: Elsevier, 1990.

O. Bernard, "Hurdles and challenges for modelling and control of microalgae for CO2 mitigation and biofuel production," *J. Process Control*, vol. 21, pp. 1378–1389, 2011.

S. Tebbani, F. Lopes, R. Filali, D. Dumur, and D. Pareau, "Nonlinear predictive control for maximization of CO_2 bio-fixation by microalgae in a photobioreactor," *Bioprocess Biosyst. Eng*., vol. 37, no. 12, pp. 83–97, 2014.

F. Mairet, O. Bernard, M. Ras, L. Lardon, and J.P. Steyeret, "Modeling anaerobic digestion of microalgae using ADM1," *Bioresource Technology*, vol. 102, pp 6823–6829, 2011.

I. Neria-González, A.R. Dominguez-Bocanegra, J. Torres, R. Maya-Yescas, and R. Aguilar-Lópeza, "Linearizing control based on adaptive observer for anaerobic continuous sulphate reducing bioreactors with unknown kinetics," *Chem. Biochem. Eng. Q.*, vol. 23, no. 2, pp. 179–185, 2009.

D. Selişteanu, E. Petre, and V. Răsvan, "Sliding mode and adaptive sliding mode control of a class of nonlinear bioprocesses," *Int. J. Adapt. Contr. & Signal Process*, vol. 21, no. 8-9, 2007, pp. 795-822.

M. Morari and J.H. Lee, "Model predictive control: past, present and future", in *Computers and Chemical Engineering*, 23(4-5), pp. 667-682, 1999.

MathWorks, Real Time Workshop, avaible online at https://www.mathworks.com/products/simulink-coder.html, 2016.

M. Lazar, Model predictive control of hybrid systems: stability and robustness, Phd. Thesis, 2006.

E. F. Camacho and C. Bordons, Model Predictive Control, 2nd edition, Springer, 2004.

Q. Mayne and. E. C. Kerrigan, "Tube-based robust nonlinear model predictive control", in Proceedings of the 7th IFAC Symposium – NOLCOS2007, Praetoria, 2007.

L. Wang, Model predictive control system design and implementaion using Matlab, Springer-Verlag London, 2009.

D.G. Luenberger, Optimization by vector space methods, John Wiley & Sons, Inc. , 1969.

G.A. Ifrim et al., "Multivariable feedback linearizing control of Chlamydomonas reinhardtii photoautotrophic growth process in a torus photobioreactor," *Chemical Eng. Journal*, vol. 218, pp. 191–203, 2013.

F. Stîngă and E. Petre, Predictive and feedback linearizing control of chlamydomonas reinhardtii photoautrotrophic growth process, Proceedings of the 30th European Conference on Modelling and Simulation, pp. 361-367, Regensburg, Germany, 2016.

FLORIN STÎNGĂ was born in Craiova, Romania. He received the B. Eng., M.S. and Ph.D. degrees in system engineering, all from University of Craiova, in 2000, 2003 and 2012. Currently, he is Lecturer in the Department of Automatic Control at the Faculty of Automation, Computers and Electronics, Craiova. His researches interested are in hybrid dynamical systems and predictive control.

MARIUS MARIAN received his Engineering degree in system and engineering from the University of Craiova (Romania), in 1998, and Ph.D. degree in information system security from Politecnico di Torino (Italy) in 2003. Currently, he is Lecturer in the Department of Computers and Information Technology at the Faculty of Automation, Computers and Electronics, University of Craiova. His research interests are in computer education, embedded systems and computer security.

KESE VALENTIN received the B.Eng. and M.Sc. diploma in system engineering from University of Craiova, in 1998 and 2008 respectively. Currently, he is an R&D system engineer at Softronic Group Craiova. His research interests are in microcontrollers, hardware, real-time systems, predictive control.

LUCIAN BĂRBULESCU was born in Craiova, Romania. He received the B. Eng. and M.S. degrees in computer science from the University of Craiova, in 2003 and 2005 and the Ph.D. degree in system engineering from the University "Lower Danube" of Galați in 2013. Currently, he is Lecturer in the Department of Computers and Information Technology at the Faculty of Automation, Computers and Electronics, University of Craiova. His research interests are in software systems for satellite dynamics.

EMIL PETRE received the Engineering degree in Automatic Control and Computers in 1977 and Ph.D. degree in Automatic Control in 1997 from University of Craiova, Romania. Since 1981 he is with the University of Craiova. Currently, he is Professor at the Department of Automatic Control, and from 2011, Dr. Petre serves as director of this department. His research interests include nonlinear systems, adaptive control, estimation and identification, control of bioprocesses, and neural control.

WIRELESS RADIATION MONITORING SYSTEM

Camelia Avram, Silviu Folea, Dan Radu and Adina Astilean
Department of Automation
Technical University of Cluj Napoca
Memorandumului street, no 28, Cluj Napoca, Romania
E-mail: {camelia.avram, silviu.folea, dan.radu, adina.astilean}@aut.utcluj.ro

KEYWORDS
Simulation, radiation monitoring, routing protocols, wireless communication.

ABSTRACT

A wireless sensors network to monitor the radiation level of radioactive contaminated risk areas is proposed in this paper. An experimental, low power, low cost wireless device, designed to collect and transmit radiation measurements data from radioactive waste dumps to servers was realized and tested. It becomes active and transmits information only when a given level of radiation is exceeded, offering the advantage of a reduced power consumption.

The data transmission was simulated in a mesh network. For the performance analysis was used the discrete event network simulator Qualnet 5.01 version, AODV and OLSR routing protocols being considered for the comparison purpose.

INTRODUCTION

The radioactive pollution records a progressive increase on more and more extended worldwide areas. In the situation in which to old, natural factors (radioactivity in minerals, cosmic radioactivity etc.) determining this phenomenon, new ones are added, the cumulative effect of successive irradiances becomes significant and will manifest itself (occur) during long time periods. A large quantity of radioactive tailings and wastes has been formed as a result of producing and reprocessing of uranium ores in the whole world.

The threat of radioactivity increases considerably in the vicinity of radioactive waste, due to uranium mining related activities, even after mining activities ceased. The means by which humans can be exposed to radiation include atmospheric, terrestrial and aquatic pathways. The atmospheric pathways are responsible of the inhalation of radon and its progeny as well as airborne radioactive particles (Viena 2002).

During the last years, the specific problems related to *activities* involved in *uranium mining* and processing, including perils and long term grave effects on environment, human health state and on evolution and equilibrium of ecosystems, were permanently notified and highlighted (Furuta, et al. 2002, Sainzand, C. et al. 2009).

Measurements accomplished in these zones indicate an increased level of soil and water radioactivity, often exhibiting a pronounced dependency on meteorological and seasonal influences.

Currently, the radiation management plans, worldwide imposed for uranium explorers, focus on the minimization of the human radioactive exposure. In this context, radioactive sites monitoring is the first step for the minimization of the possible radioactive contamination (Australian Government 2005, Australian Government. 2011). This process have to be continuous in time and starts from the opening of mining facilities (Vienna International Atomic Energy Agency. 2002), the measuring of workers exposure to radiation being critical during the uranium exploitation. Also, after the mining activities stop, the monitoring process can help to minimize or prevent the associated environmental and health risks for the nearby communities.

The greening work of uranium waste dumps leads to a significant exposure of workers. Consequently, it is important to control this radioactivity in different, interest sites, in order to monitor and warn against exceeding. Releases can be caused by the mass movement of the waste or the cover, geotechnical instability, erosion, human intrusion in relation to the waste. In different sites of the waste dump, radiation level may vary within wide ranges, depending on the wind speed and direction, temperature, humidity or other factors (Viena 2002). The determination of exposure pathways plays a decisive role for correct assessment of contamination of large areas, to obtain a complete image of the related radioactive pollution.

In the majority of cases, the existing data are not yet sufficient to derive relevant conclusions regarding the longtime variability of radiation and their environmental impact, especially because it is strongly influenced by local, particular conditions. Taking into account the above considerations, the design and implementation of reliable and cost effective monitoring systems, capable to offer consistent data regarding the evolution of radiation exposure levels on considerable periods of time, remains an important objective.

It must be mentioned that the development of wireless communication technologies opened a new perspective in this field, promising results being published in many research papers (Vijayashree and Rajalakshmi 2016.).

A solution based on wireless sensors networks to detect the radiation was proposed by Libelium, being awarded in the Data Acquisition category for its Radiation Sensor board. The authors connect to the Waspmote wireless sensor networking platform a Geiger Muller

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

detector to create an emergency radiation sensor network for measuring the beta and gamma radiation. These information and GPS (Global Positioning System) position are sent to a control center via ZigBee and GPRS (General Packet Radio Service) (Gascon and Yaza 2011).

Another application for radiation detection is proposed in (Kyker et al. 2004). It is based on ad-hoc wireless sensor networks, the authors aiming to realize a system which provide flexibility and adaptability for a variety of applications.

The power consumption, reduced size and convenient cost were other aspects taken into account during the designing process. For this reason, many current implementations of radioactive radiation detection are performed using the ZigBee technology (Adamu and Muazu).

Even though many of the mentioned references contribute to the performance enhancement of radioactivity monitoring networks, they do not take into account specific particularities for data collection in the proximity of uranium waste dumps. In this type of areas and in many cases there are no clear boundaries inside the contaminated surfaces (e.g. areas hosting tailing deposits, chaotically covered with vegetation, that overlap with areas where local population performs daily tasks), and the radioactive radiation reaches values that depend on varying natural and external factors.

The current research present a system capable to transmit reliable information from such areas types. A mesh network composed of fixed sensors nodes and also mobile ones (including communication devices carried by local workers during some daily activities) was considered to measure radiation levels. Furthermore, for testing and validation purposes, specific radiation measurement devices were designed and built.

The proposed radioactive radiation monitoring system was simulated, the protocols OLSRv2 and AODV being used for routing purposes (Uludag et al. 2012).

WIRELESS MONITORING SYSTEM ARCHTECTURE

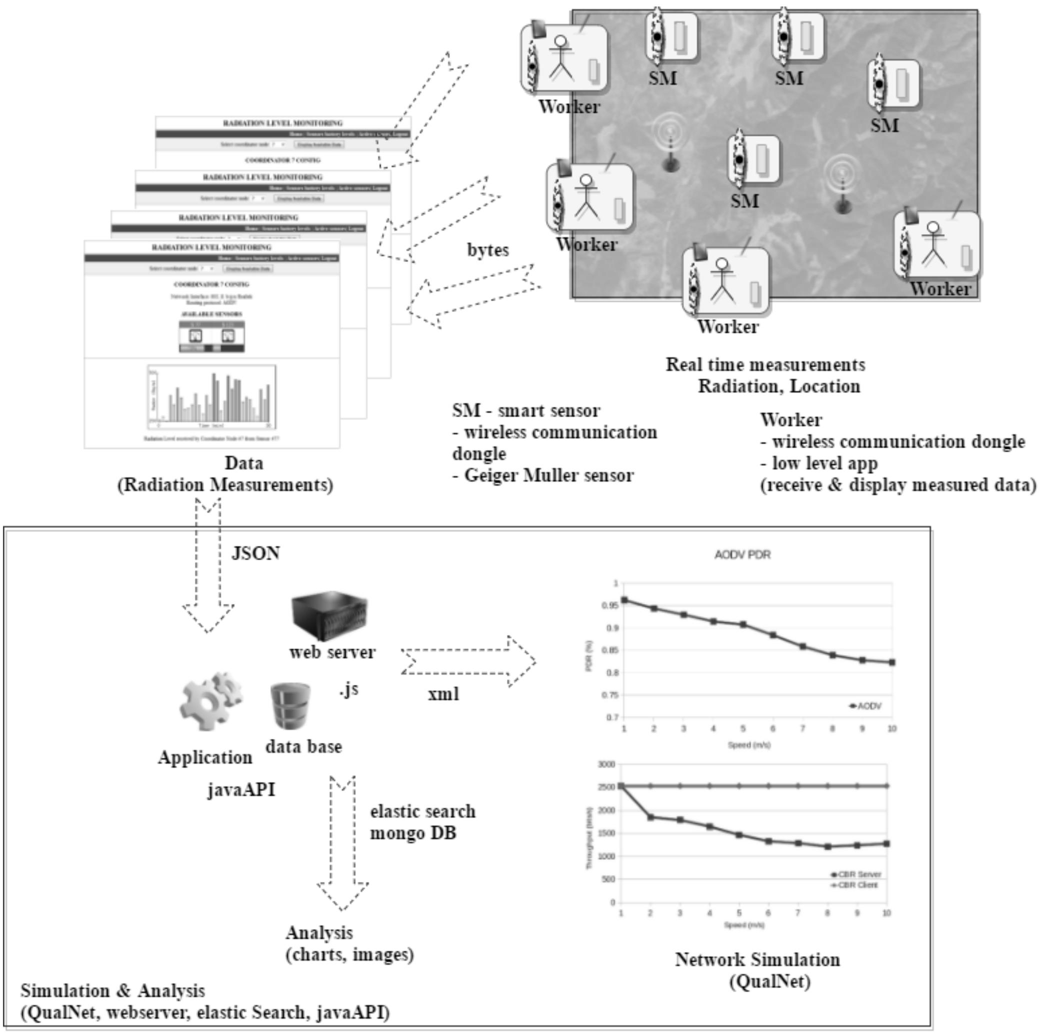


Figure 1. The architecture of the wireless monitoring system

The wireless monitoring system is composed of two parts: the real time measurements part and the simulation and analysis module.

The real time measurements are accomplished by several smart sensors (mobile and fixed, enabled with wireless communication and Geiger Muller detector) and Workers (mobile). All the measured values are sent to simulation and analysis module.

The measured data are stored and can be visualized on a web page. Java and .js api were created to data manipulation (filtering, storing in a data base, retrieving from the data base, chart creations, and analysis). For analysis purposes and realistic evaluation of the communication behavior the QualNet simulator was used as network simulator.

THE WIRELESS RADIOACTIVITY MONITORING DEVICE

The hardware architecture of the wireless radioactivity monitoring device, Figure 2, consist of: a Wi-Fi module, a CBM 20 Geiger Muller (GM) tube, a DC-DC conversion circuit which produces high voltage (HV) from a +3.0V battery and a reverse power supply protection. The device used for sending data based on Wi-Fi is the RN-131C, produced by Roving Networks.

The RN-131 module is a Wi-Fi hybrid reduced dimension circuit which can be used in mobile applications for analog and digital signal measurements.

The Wi-Fi module can scan the network for discovering the Access Points (APs), being able to associate, authenticate and connect over a Wi-Fi network to a database server using the UDP or TCP/IP protocols. The GM tube is connected to the Wi-Fi module, with a digital line and a pulse detector. Every time the tube detects a radioactive particle, the implemented system outputs a beep with an auto-oscillating buzzer, activates a blue LED and the Wi-Fi module counts the pulses and increments the value (pulses number). After a previously programmed period of time for measurements, the Wi-Fi module sends the data and passes to a sleep mode in order to extend the life time of the battery.

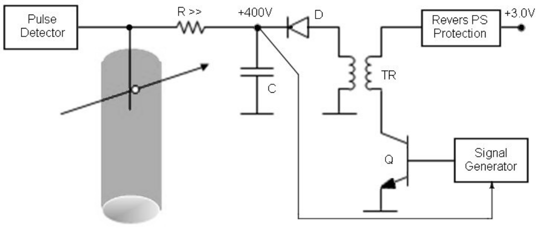


Figure 2. Hardware architecture of the radioactivity monitoring device

The Geiger-Muller tube is built differently depending on the applications in which it is used. For example, if the GM tube is used for monitoring alpha radiation, it will have a thin window at one end. Using this window, the radiations may enter the tube. The high energy electrons are generated by the photo-emission within the tube walls and can be counted. The implemented system presented in this paper is used for monitoring gamma radiation and the radioactive level in different materials. The main advantage of this system is the long lifetime which may reach up two years, due to the fact that it runs on batteries (CR123A lithium type with 3V, 1.5Ah) and that it uses Wi-Fi connectivity. The system may also warn the user by starting an alarm and sending a message when the radiation amplitude passes over a threshold limit.

Hardware implementation of the monitoring device

The design experiments were conducted using the LabVIEW™ graphical programming language. The background radiation (Figure 3) was measured, using the implemented system.

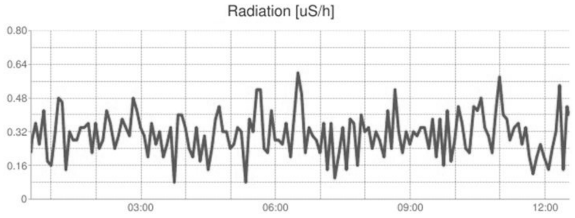


Figure 3. The background radiation

Based on the measuread values, it was observed that a permanent background gamma radiation between 0.1 and 0.5 µS/h exists in the city environment. Equation 1 is used for converting from counts per minute (CPM) to microSievert per hour (µSv/h). It is directly determined by the GM tube characteristics:

$$\mu Sv / h = CPM * 0.0057. \tag{1}$$

The electronic scheme of the DC-DC convertor to higher voltage is presented in Figure 4a. The oscillator frequency, between 1 and 3 kHz, depends on the output voltage level, which has a direct influence in reducing or increasing the power consumption. An important advantage of the implemented scheme is the current consumption, lower than 1 mA. A voltage multiplier was implemented (Figure 4b), because the power consumed by the conversion circuit is lower than the one that would be consumed in the case in which a mono alternation rectifier was used. The voltage is regulated using a feedback network implemented with the Zener diodes D11 and D12 and the resistors R21, R22, R26 and R27. The presence of particles is signaled acoustically and optically. For further data processing, the signal waveform may be acquired using the analog-to-digital converter included in the RN-131 module (with 14-bit resolution and 33 kSps rate; the input range is - 0.2 ... +0.6V). The electronic scheme implemented for connecting the RN-

131C Wi-Fi module with the particle detector is presented in Figure 5.

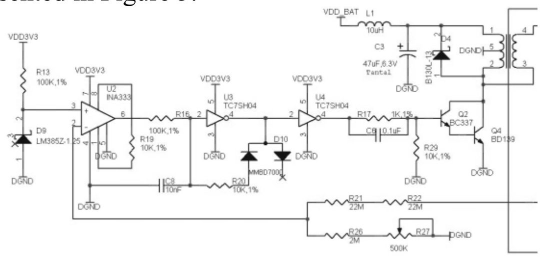


Figure 4a. DC-DC converter oscillator

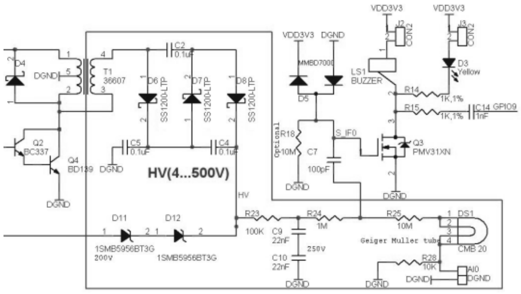


Figure 4b. DC-DC converter-voltage rectifier

The environmental temperature may also be measured using a 10 KΩs board thermistor (RT1). A green LED is turned on when the communication is performed and a red LED is used for implementing the alarms. The reverse power supply scheme is presented in Figure 6. The first version of the developed radiation detector employing the RN-131C Wi-Fi module and the CBM 20 Geiger Muller tube as a radiation detector is shown in Figure 7. The pulse shape in the Geiger Muller tube circuit is presented in Figure 8. The sample was obtained by using a Tektronix TDS1012B oscilloscope having the input probe attenuation set to 10x; the impedance is of 10 MΩs and the capacitance is of 16 pF. The comparison with a commercial system was performed for testing the proposed solution. The test system includes: a DT116 Geiger Muller detector from Fourier Systems and an USB-6009 data acquisition system from National Instruments Corp. A data logger was implemented using the Lab-VIEW^{TM} 2010 environment. The user interface is presented in Figures 9 and 10.

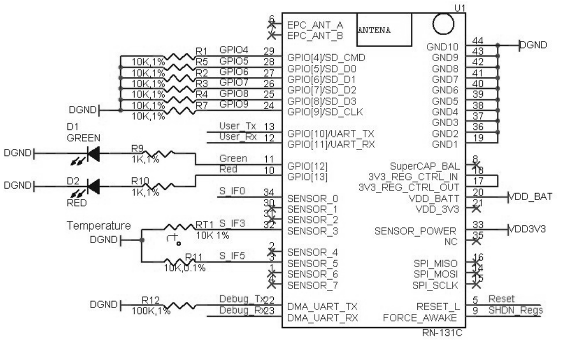


Figure 5. Wi-Fi module-connection scheme

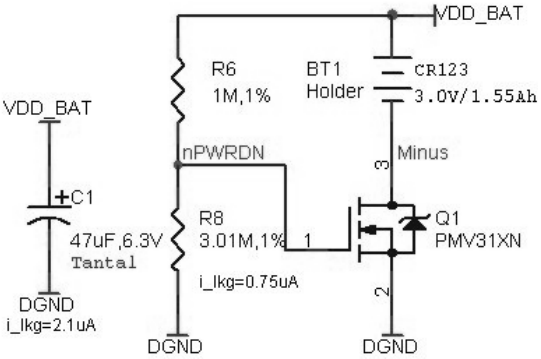


Figure 6. Reverse power supply protection

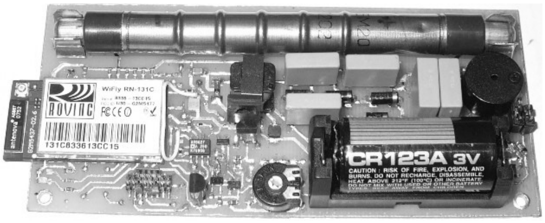


Figure 7. Radiation detector-first version

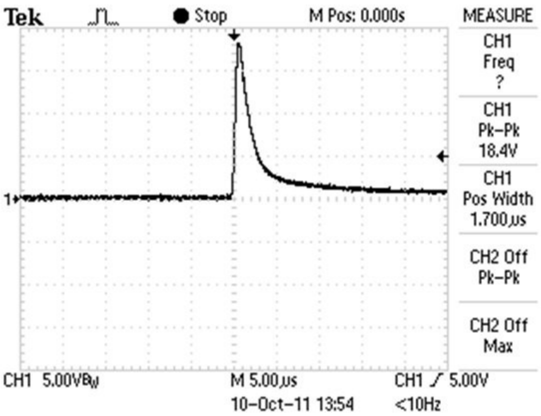


Figure 8. Geiger Muller pulse shape

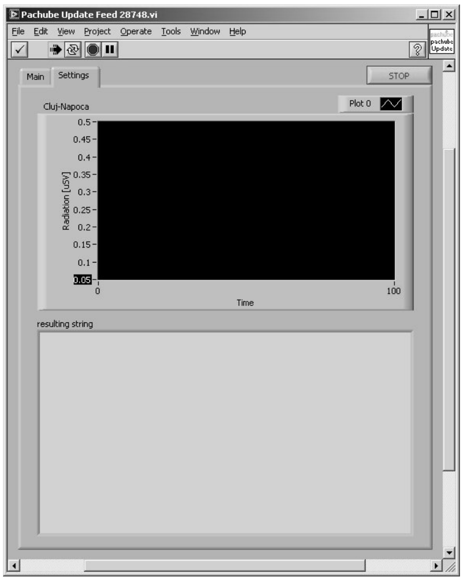
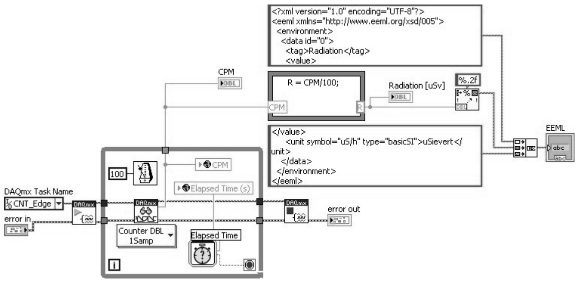


Figure 9. User interface



```
<?xml version="1.0" encoding="UTF-8"?>
<eeml xmlns="http://www.eeml.org/xsd/005">
  <environment>
    <data id="0">
      <tag>Radiation</tag>
      <value>
```

```
R = CPM/100;
```

```
</value>
      <unit symbol="uS/h" type="basicSI">uSievert</
unit>
    </data>
  </environment>
</eeml>
```

Figure 10. User interface-block diagram

PERFORMANCE ANALYSIS

Network scenario

For wireless communication network simulation we used OLSRv2 (Optimized Link State Routing) and AODV (Ad-hoc On Demand Distance Vector Routing) as routing protocols. OLSRv2 is a proactive routing protocol based on the link state algorithm. It makes use of periodic message exchange to obtain network topology information at each node. The protocol uses multipoint relays to eficiently ood its control messages. OLSRv2 provides optimal routes in terms of number of hops, available immediately when needed and it is suitable for large and dense mobile networks.

AODV is a routing protocol suitable for dynamic self-starting networks, providing loop free routes. It computes the routes on-demand and does not require periodic control messages, improving the overall bandwidth usage eficiency. The protocol scales to a large number of network nodes.

As a discrete event network simulator Qualnet 5.01 version was utilized. The detailed parameters for the network configuration are listed in Table 1.

Table 1 Parameters for the network configuration

Parameter	Values
Simulator	Qualnet 5.01
Routing Protocol	AODV and OLSRv2
Simulation area	1500 m x 1500 m
Number of nodes	41
Nodes placement	Random
Mobility	RWP, max speed 0-10 m/s
Simulation Time	100 seconds
Average Packet size	32 bytes
Transmission Interval	0.1 s
MAC Protocol	IEEE 802.11
Physical Layer Model	PHY 802.11b
Pathloss Model	Two Ray Ground
Shadowing Model	Constant
Shadowing Mean	4.0 dB
Transmission Range	270 m
Data Rate	11 Mbps

For the considered simulation we designed a network composed of 41 randomly placed nodes, in a 1500 square meters topology. The proposed topology was choosen as a good trade of between a quite dense network and computation complexity. Because the nodes speed needs to be close to urban mobility, we chosed the range of the speed from 1 m/s to 10 m/s. The data packet length is 32 bytes and contains the radiation level and the remaining battery capacity of the sensor. To be able to evaluate the worst case scenario for radiation level transmission in the network, we considered 6 CBR sources that operate in the same time interval and send data packets at a transmission interval of 100 ms.

Simulation results

Figure 11 shows the PDR (Packet Delivery Ratio) in this scenario. In Figure 12 is representing the dependency between the speed changing and PDR. AODV provides the best performances in terms of packet delivery, with more than 95% PDR for slow speeds and a minimum of 82% PDR at 10 m/s, Figure 13. In comparison, OLSR provides poor performances even at slow speeds (79% PDR at 1 m/s), Figure 14. Taking into account these results, we can state that AODV protocol is more suitable than OLSR for the proposed simulation scenario.

The web server corresponding image of the measured values sent by sensors is presented in Figure 15.

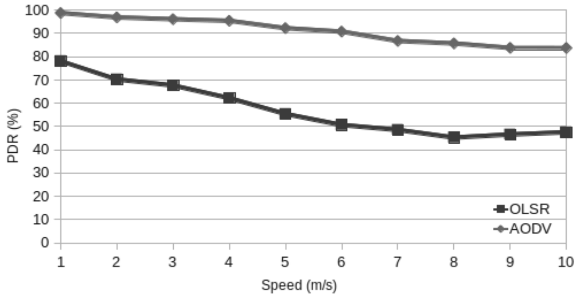


Figure 11. Packet Delivery Ratio

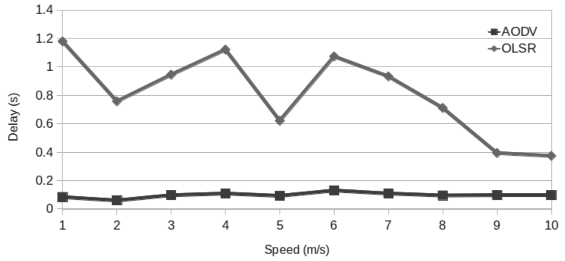


Figure 12. Average delay

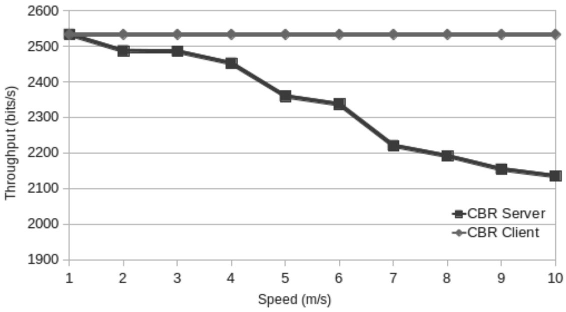


Figure 13.AODV Throughnput

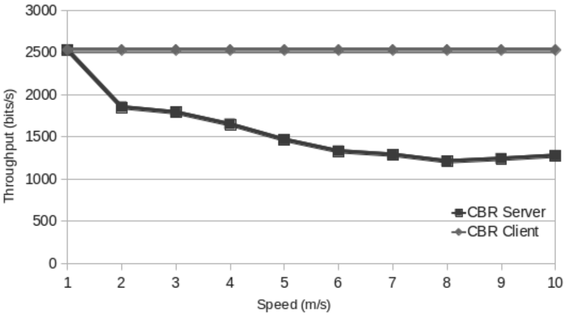
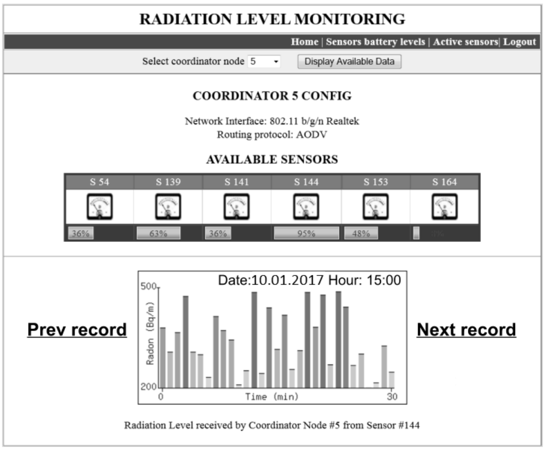


Figure 14. OLSR Throughnput



RADIATION LEVEL MONITORING

Home | Sensors battery levels | Active sensors| Logout

Select coordinator node 5 Display Available Data

COORDINATOR 5 CONFIG

Network Interface: 802.11 b/g/n Realtek
Routing protocol: AODV

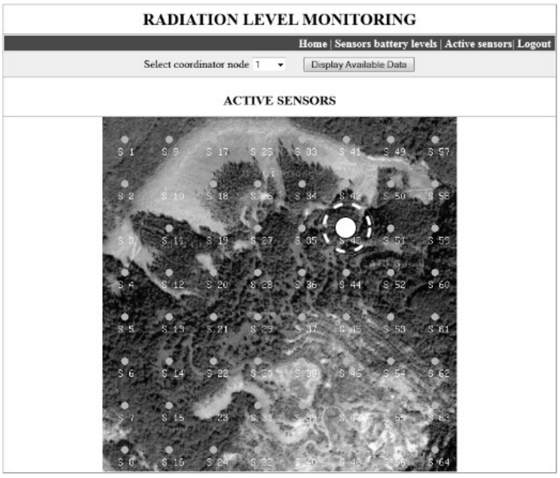
AVAILABLE SENSORS

S 54	S 139	S 141	S 144	S 153	S 164
36%	63%	36%	95%	48%	

Radiation Level received by Coordinator Node #5 from Sensor #144

Figure 15. Radiation level monitoring system

As an illustration of the final utility of this proposed system, a virtual map of radioactive contamination, generated by simulation, in the vicinity of a real radioactive waste dump is presented in Figure 16a – Figure 16d. Climatic and meteorological information was considered to determine possible spread of radioactive particles in the proximity of the site.



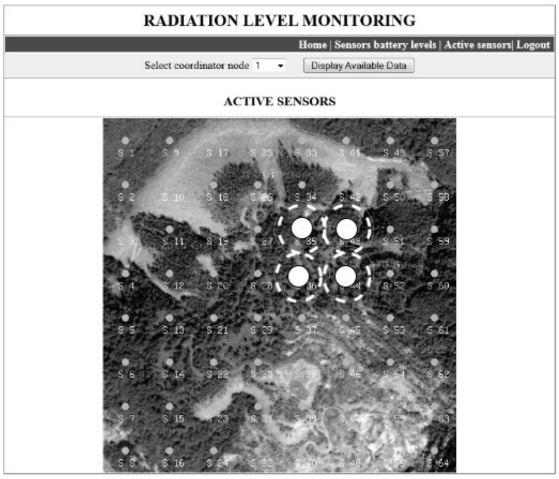
RADIATION LEVEL MONITORING

Home | Sensors battery levels | Active sensors| Logout

Select coordinator node 1 Display Available Data

ACTIVE SENSORS

Figure 16a



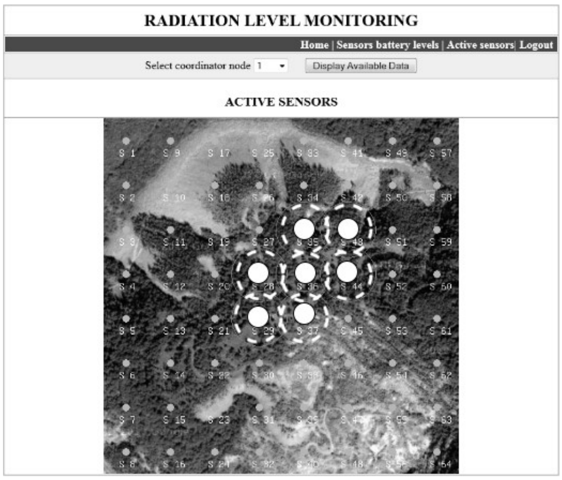
RADIATION LEVEL MONITORING

Home | Sensors battery levels | Active sensors| Logout

Select coordinator node 1 Display Available Data

ACTIVE SENSORS

Figure 16b



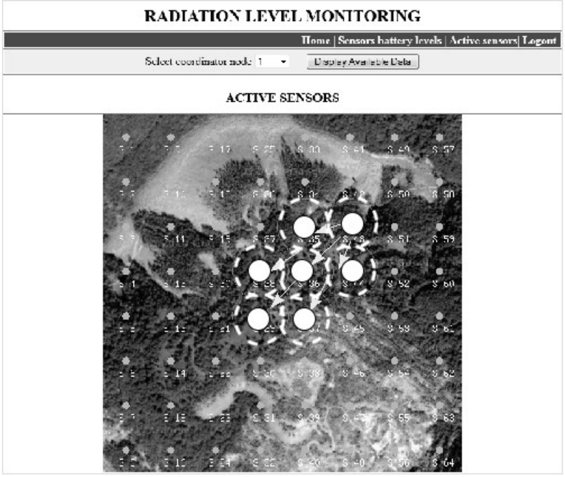
RADIATION LEVEL MONITORING

Home | Sensors battery levels | Active sensors| Logout

Select coordinator node 1 Display Available Data

ACTIVE SENSORS

Figure 16c


RADIATION LEVEL MONITORING
Home | Sensors battery levels | Active sensors | Logout
Display Available Data
ACTIVE SENSORS

Select coordinator node: 1

Figure 16d

CONCLUSIONS

The radioactivity monitoring network proposed in this paper can be used not only to transmit data related to the time evolution of the radiation level in potential radioactive contaminated areas, but also to give long-time information about background radiations and the radioactive emissions of different materials and structures. By integrating a very low power wireless device with a Geiger-Muller radiation detector, a portable device which can run on battery power was developed and tested. A mesh network topology was chosen for the considered application, the simulations indicating better results in the case in which the AODV protocol was used for routing purposes. The implemented system can improve the existing solutions presenting mobility and the capacity of storing data locally, for situations in which the wireless connectivity is unavailable for short time intervals.

REFERENCES

Adamu, H.A.; and Muazu, M.B. “Remote Background Radiation Monitoring using ZigBee Technology”, *International Journal of Electronics and Computer Science Engineering* Vol.3 Nr-2PP-148-158x

Australian Government. 2005. “Australian Radiation Protection and Nuclear Safety Agency. Radiation Protection and Radioactive Waste Management in Mining and Mineral Processing - Code of Practice and Safety Guide”, *Radiation Protection Series Publication* No. 9.

Australian Government. 2011. Australian Radiation Protection and Nuclear Safety Agency. “Monitoring, Assessing and Recording Occupational Radiation Doses in Mining and Mineral Processing - Safety Guide”, *Radiation Protection Series Publication* No. 9.1.

Furuta, S.; Ito, K.; and Ishimori. Y. 2002. “Measurements of radon around closed uranium mines”, *Journal of Environmental Radioactivity*, 62(1):97–114, 2002.

Gascon D.; and Yaza. M. 2011. “Wireless sensors network to control radiation levels”, *Libelium Comunicaciones Distribuidas S.L.*, 2011

Kyker, R.D.; Berry, N.; Stark, D.; Nachtigal N.; and Kershaw. C.; 2004. "Hybrid emergency radiation detection: a wireless sensor network application for consequence management of a radiological release", *Proc. SPIE 5440, Digital Wireless Communications* VI, 293.

Sainzand, C.; Dinu, A.; Dicu, K.; Szacsvai, T.; Cosma, C. and Quindos, L.S. 2009. “Comparative risk assesement of residential radon exposure in two radon prone areas, stei (romania) and torrelodones (spain)”, *Science of The Total Environment*, 407:4452–4460.

Viena. 2002. “Vienna International Atomic Energy Agency. Monitoring and Surveilance of Residues from the Mining and Milling of Uranium and Thorium”, *Safety Reports Series* No. 27.

Vienna International Atomic Energy Agency. 2002. “Management of Radioactive Waste from the Mining and Milling of Ores”, *Safety Standards Series* No. WS-G-1.2.

Vijayashree, S. and Rajalakshmi, S. 2016. “Survey on Deployment of WSN in Radiation Monitoring”, *International Journal of Science and Research (IJSR)*, vol.15, 2016, pp.1707-1709

Uludag, S.; Imboden, T.; and Akkaya, K. 2012. “A taxonomy and evaluation for developing 802.11-based wireless mesh network testbeds”, *International Journal of Communication Systems*, Vol.25 Issue8, Pages 963-990.

AUTHOR BIOGRAPHIES

Camelia Claudia Avram, 39 years old, PhD Eng., Faculty of Automation and Computer Science, TUCN, Romania. She has participated in many international and national research projects, with different achievements of software applications, demonstrators and one patent awarded. She has professional expertise in modelling and simulation, discrete event systems, data transmmison and real time systems. Her email address is: camelia.avram@aut.utcluj.ro.

Silviu-Corneliu Folea, 44 years old, PhD Eng., Associated Professor, Automation Department, TUCN. Him has experience in the design and creation of the incorporated systems, hardware and software instrumentations with different prototypes for National Instruments and two international patents awarded at Technical Innovation Salons. His email address is: silviu.folea@aut.utcluj.ro.

Dan Radu, 34 years old, PhD.Eng. Automation Department, Automation and Computer Science Faculty, TUCN. Experience in design and implementation of wire and wireless communication systems and mobile/Android applications.
His email address is: dan.radu@aut.utcluj.ro.

Adina Astilean, PhD. Eng. Automation Department, TUCN. She has professional expertise in modelling and simulation, discrete event systems and data transmmison. Her email address is:
adina.astilean@aut.utcluj.ro

SIMTONIA – A framework of SIMulation TOols for Nuclear Industrial Applications

József Páles, Áron Vécsi and Gábor Házi
Reactor Monitoring and Simulator Department
Centre for Energy Research, Hungarian Academy of Science
H-1525, Budapest, Hungary
E-mail: jozsef.pales@energia.mta.hu; aron.vecsi@energia.mta.hu; gabor.hazi@energia.mta.hu

KEYWORDS
nuclear industry, full-scope training simulator, modeling tools

ABSTRACT

A new set of simulation tools, developed to support the overall preparation process of full-scope simulators for nuclear reactors, is introduced. Using a uniform user interface and standardized software development platform, the tools are grouped around a framework called SIMTONIA.

Here we give a brief description of the framework and its elements, outlining the most important features of the system. The vitality of the framework is demonstrated by a simple modeling exercise.

INTRODUCTION

Centre for Energy Research of Hungarian Academy of Science (CER HAS / former KFKI AEKI) has been supported the maintenance and development of the full-scope training simulator (FSS) of Paks NPP since the late eighties. Due to these developments, the major components of the FSS (all significant models, I/O communication etc.) had been renewed by CER in the last two decades. However, witnessing rapid evolution in information technology, three years ago we decided to initiate a project focusing on the renewal of our all simulator development tools.

The result of these developments is called SIMTONIA (SIMulation TOols for Nuclear Industrial Applications) a simulator development framework, which provides a modern, comfortable, graphics environment for simulator developers. Using this framework there is no need of programming skills for model development, but developers having solid knowledge of technology can develop the simulator models. The user interface of SIMTONIA is unified, which means that independently from the technology we wish to model a unified man-machine interface can be used to produce models of

- instruments and control (I&C),
- electrical,
- single phase fluid and
- two-phase flow

networks.

The different technological networks can be coupled with each other using the same graphics environment and they can be driven by a virtual control room environment, which also can be developed using SIMTONIA. SIMTONIA's executive and instructor systems provide real-time scheduling of the models, usual simulator operations (snapshot reading and writing, backtracking, run, step, freeze functions etc) in line with the well-known ISO standards.

Technological data are stored in SIMTONIA's offline database and managed inside the framework in accordance with model variables and parameters.

In the next Section we give an overview of our framework. Then the most important models are introduced briefly. Finaly, using a simple example of water control level in a tank, we demonstrate how the models developed by the tools can be integrated. More complex application from nuclear industry can be found in (Páles et al., 2017).

OVERVIEW OF THE FRAMEWORK

SIMTONIA is a general purpose simulation environment which is designed to handle all aspects of the development and operation of full-scope simulators used for the training of nuclear power plant operators. In the present section we give a brief overview of the system focusing on its major components (Fig. 1).

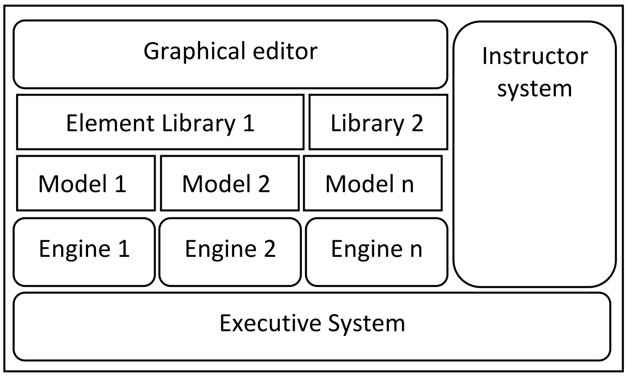


Figure 1: SIMTONIA Framework Components

The central component of the framework is an object oriented graphical editor which eneables the users to create element libraries and build up complex technological models using these libraries. The libraries contain the basic building elements of the technological

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

models, such as logic gates for I&C models or pipe sections and fluid tanks for thermohydraulic models. Each element consists of an interactive graphical representation, a set of parameters and state variables, a set of input-output connection points, and a set of callback fuctions which are used by the simulaton engine during the simulation. The user can create complex technological models by placing these building elements on the diagrams of a model, interconnecting and parametrizing them. Thus each simulation model in the framework can be represented in a uniform way by a network of interconnected elements (Fig. 2).

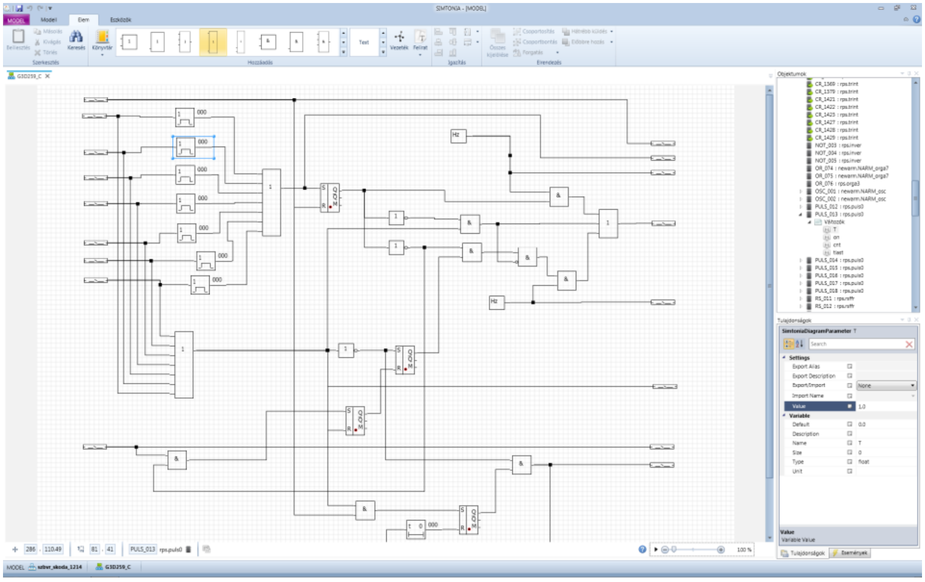
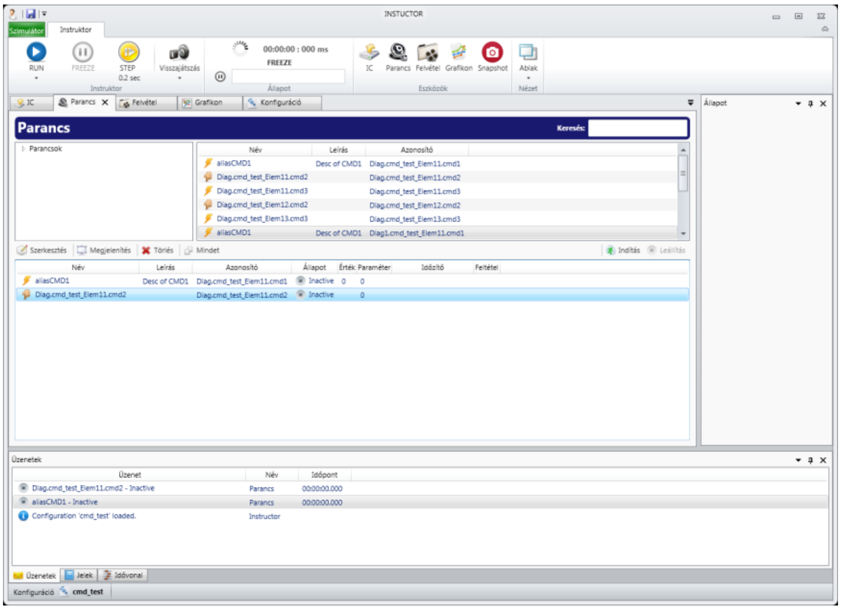


Figure 2: SIMTONIA’s Model Editor

The models can be simulated by using specific simulation engines. Since different numerical methods are required to model different physical and technological processes, each model runs on a specific simulation engine. Currently the framework contains engines to simulate reactor core physics, I&C networks, simple electrical networks, single and two phase flow networks, and human-machine interfaces. Some of them will be presented in the next section. The simulation framework can be simply extended with new simulation engines using SIMTONIA’s Model Application Programing Interface (API).

For the simulation of complex technological systems several engines have to be coupled with each other. SIMTONIA’s Executive System is responsible for the realtime scheduling of the engines, and for controlling the synchronized communication between the models. The models can communicate with each other via an export-import mechanism: a model can export variables using globally unique variable names, and other models can import these variables using the export names. For scheduling the models, the Executive System uses simulator configurations (its description can be also created in a visual manner as a kind of flowchart using a dedicated element library of SIMTONIA), which contains information about tle list of models that take part in the simulation, and the execution order of the models. The models can be scheduled to run in parallel or in sequential order. For training simulators it is important, that the simulation must run in realtime to reproduce the events in the control room in a realistic way. Beside these basic functions, the executive system also provides several builtin services to support the training of the operators and fulfill the requirements of the ANSI-ANS-3.5 standard for NPP training simulators (ANS 2009). Some of these features are given as follows:

- support for snapshot and backtrack handling,
- support for implementing and using malfunctions and local operator actions,
- support for testing and validating the models, etc.



Név	Leírás	Azonosító
aliasCMD1	Desc of CMD1	Diag.cmd_test_Elem11.cmd1
Diag.cmd_test_Elem11.cmd2		Diag.cmd_test_Elem11.cmd2
Diag.cmd_test_Elem11.cmd3		Diag.cmd_test_Elem11.cmd3
Diag.cmd_test_Elem12.cmd2		Diag.cmd_test_Elem12.cmd2
Diag.cmd_test_Elem13.cmd3		Diag.cmd_test_Elem13.cmd3
aliasCMD1	Desc of CMD1	Diag1.cmd_test_Elem11.cmd1

Név	Leírás	Azonosító	Állapot	Érték	Paraméter	Időzítő	Feltétel
aliasCMD1	Desc of CMD1	Diag.cmd_test_Elem11.cmd1	Inactive	0	0		
Diag.cmd_test_Elem11.cmd2		Diag.cmd_test_Elem11.cmd2	Inactive		0		

Üzenet	Név	Időpont
Diag.cmd_test_Elem11.cmd2 - Inactive	Parancs	00:00:00.000
aliasCMD1 - Inactive	Parancs	00:00:00.000
Configuration 'cmd_test' loaded.	Instructor	

Figure 3: SIMTONIA’s Instructor System

The Instructor System is the main user interface for the isntructors to control the operation of the simulator during the training (Fig. 3). It is connected to the Executive System, and provides a user interface for the operators to execute the following tasks:

- run, step and stop the simulation,
- saving and loading snapshots (initial conditions) of the simulation,
- initiating malfunctions and local operator actions,
- building complex simulation scenarios,
- backtracking the simulation to a previous state and and replaying it,
- plotting and saving trends of simulator variables, etc.

The SIMTONIA framework is a fully object oriented system which is built on modern software technologies: it runs over the Microsoft .NET Framework, and uses the WPF API for creating vector graphic interfaces.

MODELING TOOLS

In the following section we briefly introduce two simulation engines of the SIMTONIA framework.

Sequential engine

A common task in NPP simulation is the detailed modeling of I&C devices used in the various technological systems of the plant. The complexity of these systems moves in a wide scale and their proper modeling has great impact on the quality of simulation. Therefore it is of crucial importance to model the control systems in an NPP training simulator as accurately and efficiently as possible.

In the SIMTONIA framework we have developed a sequential network simulation engine to model the control systems. The algorithm for solving the network is an improved version of the method we used in GRASS (GRAphic Simulator System), a graphical modeling tool developed also by CER and used in the training simulator of Paks NPP (Jánosy et al. 1997).

In the element libraries the user can define input-output connection points, state variables, and a simulation routine for each basic elment (e.g. logic gates, flip-flops, comparators, analogue functions, PID controllers, etc.) used in a sequential network. The simulation routines contain the algorithms which determine how the output signals and state variables are computed from the input signals. Based on the topology of the network, the sequential engine determines an execution order of the simulation routines in a model.

The outline of the algorithm used for creating the execution order is the following:

1. A list is created containing all pairs of connected input-output connection points. This list will be sorted to determine the exeution order of the routines.
2. The connection pairs having source elements which are not driven by other elements in the network are executed. Then those connections come which have driving elements only in the already sorted part of the connection list, and so on. At the end of this sorting procedure, only those connections remain unsorted, which belong to a loop or are driven by an element of a loop.
3. The remaining part of the list can be sorted in the backward direction. All connections have to be put to the end of the list which drive only output

elements. Then the connections which drive elements already in the sorted end of the list will be listed before the tail of the list, and so on.

4. At the end of the sorting procedure the sorted list contains three sections: non-loop connections sorted in execution order, the connections taking part in loops, and the connections driven by loops sorted in execution order.

Single phase flow network engine

The SIMTONIA framework also contains a simulation engine for modeling single phase flow networks. Using the flow network element library the user can build simple pipe flow networks. The flow library contains the following basic components:

- nodes,
- branches interconnecting the nodes,
- pumps,
- valves,
- heat excangers,
- tanks and
- leaks.

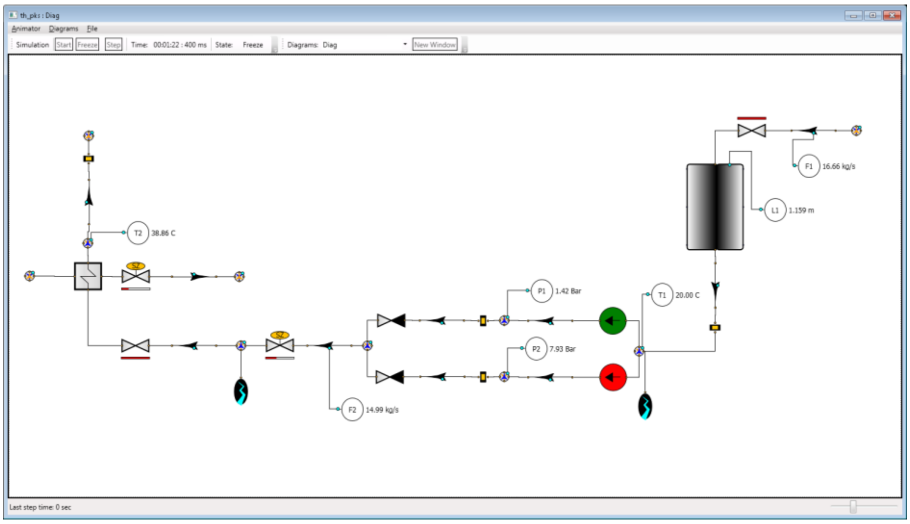
The engine simulates the dynamic behavior of the flow network by solving an equation system built up based on the structure of the network. The resulting equation system contains a mass and energy conservation equation for each node, and a momentum conservation equation for each branch. By solving the equation system the engine determines the pressure, enthalpy, and transported material concentration in each node, and the mass flow rate in the branches.

To simplifiy the solution of the equation system we assume that the fluid is incompressible, therefore the energy conservation equation can be solved independently from the mass-momentum equation system, based on the pressure and flow distribution in the network.

To further simplify the solution of the equation system the flow network can be divided into sub-networks dynamically when different parts of the network become isolated (islands) form each other by closing some valves.

EXAMPLE – SIMULATION OF WATER LEVEL CONTROL

To demonstrate the capabilities of the SIMTONIA framework we present the simulation exercise of a simple coupled thermohydraulic and I&C model. The thermohydraulic model contains an open tank which is filled up with water at the top of the tank through a manual valve (Fig. 4). The valve can be used during the simulaton to interactively set a constant mass flow rate towards the tank.



Time: 00:01:22 : 400 ms

State: Freeze

Diagrams: Diag

Last step time: 0 sec

Figure 4: Simple thermohydraulic model

The water flows out of the tank at its bottom through two parallel pipe sections, both of them containing a pump, pressure measurement devices (P1, P2), and check valves at their ends. The parallel pipe sections are joined after the check valves. The joined section contains a control valve, a manual valve, and a heat exchanger connected to the pipe by its tube side. There is a temperature measurement device at the outlet of the heat exchanger (T2), and the water leaves the simulation domain through constant pressure boundary condition, i.e. the pressure at the exit is prescribed. The shell side of the heat exchanger also contains two pressure boundary conditions, and a control valve which controls the mass flow rate of the heat exchanger.

Three control logic schemes are connected to this thermohyraulic model:

1. The control of water level (L1 measurement) in the tank with the control valve in the joined pipe section.
2. The control of the outlet temperature at the tube side of the heat exchanger using the control valve at the shell side of the heat exchanger.
3. Stopping and starting logics of the pumps based on the pressure (P1, P2) and water level (L1) measurements.

In this example we only show the details of the water level control (Fig. 5). The two simulation engines are connected with each other via export-import variables. All measurments are exported from the thermohydraulic model, thus the control logic can import the value of the water level (L1) measurement with an import element.

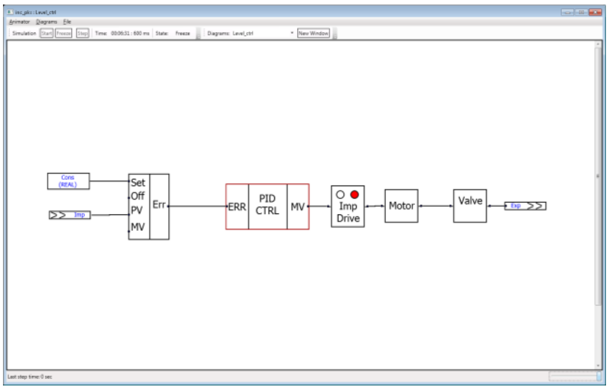


Figure 5: I&C model for level control

This value is connected to an element which calculates control error, i.e. the deviation form a constant setpoint value, which is set to 1.2 meter in this case. The calculated deviation is connected to a standard PID controller, which generates close or open commands based on the prarameters of the controller.

The output of the controller is processed by a pulse width modulator, which is often used in real technological systems to prevent valves from continuous operation (moving).

The pulses produced by the generator drive a constant speed motor, which sets the relative position (1-100%) of the valve stem. The effective position of the valve stem can be determined by the characteristics of the valve.

Finaly, the effective position of the valve is exported to the control valve of the thermohydraulic model by an export element.

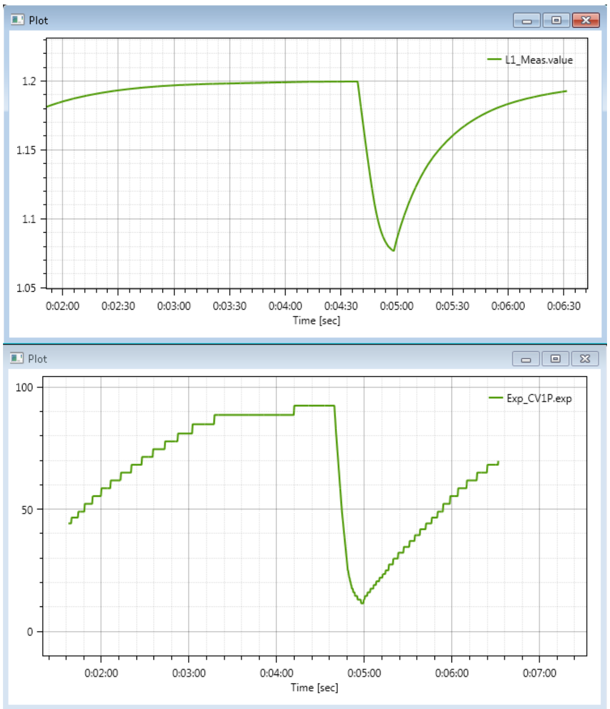


Figure 6: Plot of the water level (top) in the tank and control valve shaft position (bottom)

The evolution of the water level and the position of the control valve determined by the control logic scheme can be seen in Fig. 6. At the beginning of the transient the controller gradually opens the valve setting the water level to the proper setpoint 1.2m. When the level achieves 1.2m, we initiate a short-time leakage in the system just behind the valve. Due to the leak the level of the tank starts to decrease rapidly and accordingly the controller gradualy closes the valve, in order to stop the deviation from the setpoint. Terminating the leakage, the water level starts to increase rapidly and controller start to open the valve to restore its proper setpoint value.

CONCLUSIONS

SIMTONIA framework provides a simple way to develop complex I&C, thermohydraulic and electric network models. Its model libraries have plenty of predefined elements letting the user to develop models easily for complex thermohydraulic, I&C and electrical networks without having any programing skills. Although here just a simple modeling exercise was given to demonstrate the vitality of this framework, its more comprehensive applications are under development in the backstage: development of a compact simulator for AES-1200 and for VVER-440 type nuclear reactors. We hope to present the first simulation results of these simulators in the near future.

REFERENCES

ANS ,2009. “ANSI/ANS-3.5, Nuclear Power Plant Simulators for Use in Operator Training and Examination”, *American National Standard*, American Nuclear Society.

Jánosy J., Házi G., Seregi L., Szabó K., 1997 "GRASS – The Graphic System Simulation", *Proceedings of the 2nd CSNI Specialist Meeting on Simulators and Plant Analysers* Olli Tiihonen (Eds.), (Espoo, Finland, Sept. 29 – Oct. 2)

Páles, J.; Vécsi, Á.; Házi, G., 2017, Nuclear Industrial Applications of SIMTONIA, 31st European Conference on Modeling and Simulation, (Budapest, Hungary May 23-26).

AUTHOR BIOGRAPHIES

JÓZSEF PÁLES was born in Hungary. He studied information technology on Pannon University and obtained his M.Sc. degree in 2006. He has been working at the Reactor Monitoring and Simulator Department for the Hungarian Academy of Sciences Centre for Energy Research since 2002. His e-mail address is : pales.jozsef@energia.mta.hu.

ÁRON VÉCSI was born in Miskolc, Budapest and went to the Budapest University of Technology and Economics, where he studied physics and obtained his M.Sc. degree in 2016. He is now PhD student and working at the Reactor Monitoring and Simulator Department for the Hungarian Academy of Science Centre for Energy Research. His e-mail address is: vecsi.aron@energia.mta.hu

GÁBOR HÁZI was born in Budapest, Hungary and obtained his B.Sc. in Kando College in 1992 and M.Sc. in electrical engineering at Technical University of Budapest in 1995. He holds a PhD in nuclear engineering from the Technical University of Budapest and DSc from the Hungarian Academy of Science. He has been working for CER since 1992. Now, he is the head of Reactor Monitoring and Simulator Deparment. His e-mail address is : hazi.gabor@energia.mta.hu.

Nuclear Industrial Applications of SIMTONIA

József Páles, Áron Vécsi and Gábor Házi
Reactor Monitoring and Simulator Department
Centre for Energy Research, Hungarian Academy of Science
H-1525, Budapest, Hungary
E-mail: jozsef.pales@energia.mta.hu; aron.vecsi@energia.mta.hu; gabor.hazi@energia.mta.hu

KEYWORDS
I&C modeling, reactor power controller, core monitoring, critical safety function monitoring

ABSTRACT

In this paper we present three different nuclear industrial applications of the SIMTONIA framework developed by CER. First, we demonstrate how we could utilize SIMTONIA’s sequential engine for the verification and validation of the new control rod and reactor power controller system of Paks NPP. In our second example, the application of SIMTONIA’s visual engine for system monitoring in the VERONA core monitoring system of Paks NPP will be presented. Finaly, the critical safety monitoring system and electronic operational rules of the plant computer of Paks NPP will be briefly introduced focusing on the application of SIMTONIA’s sequential engine as an evaluation tool of logic diagrams.

INTRODUCTION

Centre for Energy Research of Hungarian Academy of Science (CER) has developed a new set of graphics tools for the development of numerical models of nuclear power plant simulators (NPP). The result of these developments is called SIMTONIA (SIMulation TOols for Nuclear Industrial Applications), which is a simulator development framework providing a modern, comfortable, graphics environment for simulator developers. The elements of this framework has been introduced, recently (Páles et. al., 2017). It might be worth noting that similar simulator development platforms were developed by other simulator suppliers (e.g. the Orchid by L3-MAPPS). However, our approach differs significantly from the one used by most of the suppliers, since we do not use old-fashioned code generation strategies, but our model engines utilize DLL based object libraries.

Using three examples, we demonstrate that the applications of the flexibe engines of SIMTONIA’s framework are not limited to simulator model developments, but they can be exploited in a number of fields of nuclear industry, too.

In the next Section we demonstrate how we could utilize SIMTONIA during the verification and validation (V&V) procedure of a new, safety related control systems of Paks NPP.

Then, the replacement of GE’s iFIX Intellution human machine interface by SIMTONIA’s visual engine used in the VERONA core monitoring system of Paks NPP will be discussed, as a cost effective, highly reliable solution.

Finaly, we discuss how SIMTONIA’s sequential engine is utilized in the plant computer system of Paks NPP for the evaluation of the so-called critical safety functions and electronic operational rules of the plant, inluding the automatic detection of the plant’s operation mode.

VALIDATION AND VERIFICATION OF NEW CONTROL SYSTEMS BY SIMTONIA

In 2015, Paks NPP initiated a public procurement procedure for reconstruction of the instrumentation and control of the Rod Control System (RCS) and the Reactor Power Controller System (RPCS) of plant. The purpose of the reconstruction has been to improve the RCS and RPCS with respect to instrumentation and control to today’s standards by applying modern methods and solutions and by ensuring that the safety and reliability criteria related to the planned extension of the nuclear plant’s lifetime are met.

SKODA JS was awarded by end user in the procurement procedure and CER as a subcontractor of SKODA JS has been taken part in the project by focusing on the simulator related activities of the project. In particular, CER has been responsible for the implementation and integration of the simulator models of the new RCS and RPC to the full-scope simulator (FSS) of Paks NPP. Beside the implementation of the new models, CER has also been supported the verification and validation procedure of the new RCS and RPCS by performing tests in the simulator in order to demonstrate the reliability of the new system.

Functions of RCS and RPCS

Without going into details, let us summarize what are the roles of RCS and RPCS in an NPP.

Basicaly, the reactor power in an NPP can be controled by introducing neutron absorbers into the reactor core where the chain reaction takes place. For rapid control of power, so called control rods are introduced into the core. Paks NPP has 37 control rods and these are organized into six groups. The groups and the individual rods can be moved up and down by the reactor operators, and in automatic mode the groups can be

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

driven by the RPCS, too, keeping the neutron flux or the turbine pressure in a certain level depending on the operation mode of RPCS. Finaly, we note that the reactor protection systems can also drive the control rods in case of abnormal operation modes, e.g. dropping down all the rods into the core by gravity, if it is necessary. During the reconstruction, the underlying principles of RCS and RPCS had to be retained, but the plant personel also requested some new functions to be introduced. For these new features some operational panels of the main and emergency control room (MCR and ECR) of the plant had to be modified, introducing new switches and displays into the panels.

Modeling RCS and RPCS in the FSS of Paks NPP

The refurbishment procedure of these systems is a long procedure in the four real units of the plant, rougly four years, because it can be done only when the reactor is shutdown between two reactor cycles, which are 15 months long. Due to the long-continued refurbishement procedure, there is a period of time when the old system is used in some units and the new one is in action in the others. Since reactor operators must be trained in the simulator for both kind of systems, therefore it had to be guarented that the models of both systems be available in the simulator and a simple way had to be provided to switch over from one model to the other. Also, the differences between MCRs and ECRs had to be represented somehow. Therefore these panels have been modified making possible to replace the control rod actuators and displays rapidly by a screw mounted new panel when the new system is used for training.

Considering the models, a software switch was introduced to jump from the old system model to the new one.

In the FSS of Paks NPP, the old RCS and RPCS were simulated by models written in Fortran subroutines. It was requested that during the reconstruction, the models of the new system be implemented by GRASS, a graphic logic diagram modeling tool developed also by CER and used earlier during the refurbishment of the reactor protection system of Paks NPP (Jánosy et al., 1994). Since GRASS has been developed by more than 20 years ago, its user interface does not suit with present standards. Therefore we have decided to apply a multi-step procedure during the implementation of the new model:

1. Implementation of the new logic elements (60 building blocks) by SIMTONIA's element editor according to the description of SKODA JS' system designers.
2. Implementation of 130 new logic schemes of the new design by the user-friendly environment of SIMTONIA's diagram editor, using the previously developed building blocks (see an example in Fig. 1).
3. Off-line testing of SIMTONIA models.
4. Automatic conversion of SIMTONIA's building blocks and diagrams to GRASS icons and pictures.
5. Coupling the GRASS models with our replica simulator of Paks NPP's FSS.
6. Model code generation by GRASS.
7. On-line testing of the new models in our FSS.

Applying this procedure, the implementation of the new RCS and RPCS system was seamless and requested only two months including both off-line and on-line testing.

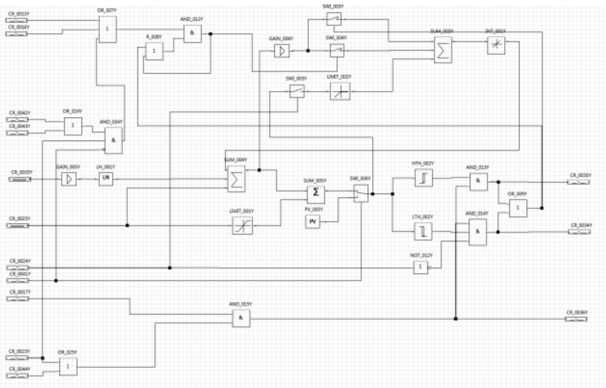


Figure 1: A diagram from the RPCS model

A priori testing of the new MCR and ECR panels

In order to *a priori* test the actuators and displays of the new systems from ergonomic point of view, we also utilized SIMTONIA. Using its visual engine, the old and the new MRC and ECR panels have been implemented as animated pictures and our touch-screen based virtual control room panel shown in Fig. 2 was coupled with the old and new simulator models, respectively.

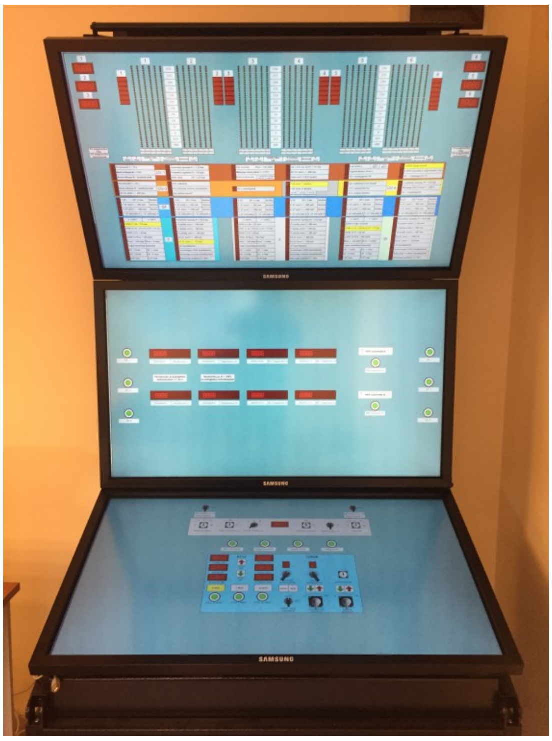


Figure 2: Touch-screen based control room panel

Using again a software switch, we provided a simple way, to switch from the old panels to the new ones (parallel with the change of models).

Verification and validation of the new models

Setting up this environment a Factory Acceptance Test (FAT) had taken part in my Institute. This V&V procedure followed a very simple scenario. First, all basic functions of the new systems, e.g. individual control rod movements and control rod group operations were tested by the operational personal of the NPP, then more complex situations (e.g. automatic reactor power control at different power levels and during normal and abnormal operations) were simulated by both the old and new systems. Parameters of the controllers were tuned by comparing the obtained relevant physical parameters with each other during these simulations.

After successful FAT, the models were introduced in the FSS of Paks NPP and Site Acceptance Test (SAT) was carried out in a similar manner than the FAT procedures. After successful SAT, a nearly one-year trial period of the new system started, using the new RCS and RPCS in the simulator by the training personel and operators of the plant one day per each week.

Finaly, the new system was first introduced in the 2nd unit of the plant at the end of 2016.

PROCESS MONITORING BY SIMTONIA

In NPPs so-called core monitoring and surveillance systems are used to extract as much information about the state of the nuclear reactor core as possible (Végh et. al., 2008). Although only a limited number of detectors can be installed in a core, using advanced numerical calculations the relevant physical quantities can be obtained with very high resolution. For instance, the core of a VVER-440/213 type NPP consists of 349 fuel assemblies, but it has only 210 thermocouples at the outlet of the assemblies and only 36 assemblies are equipped by neutron flux detectors. In spite of these limited number of measurements, utilizing the available measured data and using well-established neutron physical and thermohydraulical calculations, the temperature, the neutron flux and several other derived quantities can be obtained in more than 10000 equidistantly distributed computational points of each fuel assembly.

In Paks NPP, the VERONA on-line monitoring system has been responsible to determine the relevant reactor physical quantities and associated safety margins (e.g. distance from saturation temperature) of the reactor core since the late eighties (Végh et. al., 2008; Végh et al. 2015). This system was gradually improved in the last three decades as the performance of computational techniques drastically improved allowing more and more sophisticated numerical computations. It is also worth emphasizing that these developments were needed to establish the application of a new, more economic generation of VVER fuel assemblies while keeping or even increasing the safety level of operation.

In 2012, Paks NPP decided to change its operational practice extending the fuel cycles from 12 to 15 months increasing the enrichment of Uranium in the fuel assemblies. Some preliminary calculations revealed that the amount of on-line core monitoring calculations increases significantly due to the changes of fuel, i.e. developments were needed in the core monitoring system. It is also turned out that the increasing amount of computational work cannot be managed by the old hardware and software platforms while keeping the same high level availability (99.9%) of the system than before. Since the hardware and software platform of VERONA became also obsolete in the last decade, the NPP decided to initiate an overall reconstruction work of VERONA core monitoring system and the required major developments can be summarized as follows:

- application of a new generation but proven hardware devices (high performance servers, thin clients for monitoring, new local area network devices, network attached storage for archives),
- application of a new generation but proven software platform (Windows Server 2012 R2),
- application of a new generation of software development tools (Visual Studio 2014, Embercadero XE7, Visual Fortran Composer XE 2013),
- application of VMware virtualization technology,
- extension and acceleration of reactor physics calculation using Graphics Processing Units (GPU),
- replacement of the GE's Intellution iFIX based HMI for process monitoring.

In this paper we focus on the last item. The replacement of iFIX SCADA solution for process monitoring was motivated by the fact that the old version of iFIX used in VERONA system was not supported under the new software platform, therefore new licences would have been purchased by the plant. However, GE's iFIX licence policy basicaly did not allow the plant to initiate such a purchasing, therefore the plant management decided to replace iFIX with some other process monitoring solution. CER suggested to apply SIMTONIA's visual engine as a HMI of VERONA and this proposal was awarded.

VERONA's architecture

Paks NPP has four reactor units and each unit has its own local VERONA network. The principal users of VERONA are the reactor operators and the unit supervisor, who work together in the control room of the unit, where two displays of VERONA system are installed for on-line core monitoring purposes.

Although the local networks of VERONA are independent from each other, all of them are connected to the technological network (TN) of the NPP, which are separated from other networks (e.g. informatics network) of the plant by a data diode (Fig. 3).

The connection of local VERONA networks to TN assures external accessibility of VERONA data from the Control Centre of the plant and from some dedicated places supporting remote maintenance. Such an important place is the Computer Centre of the 3rd unit, where the VERONA-t test system has been installed. This is a kind of test bed of VERONA system, since it can be driven by the measurements of any of the units.

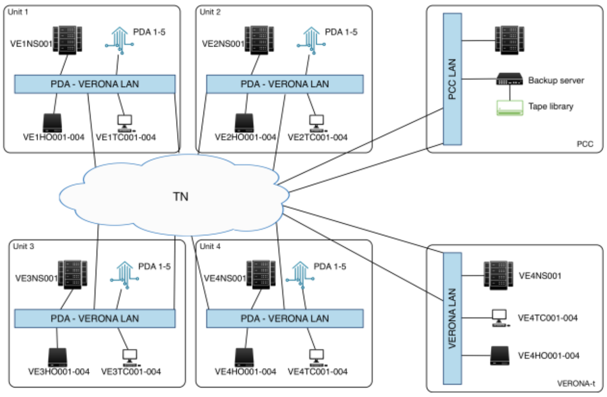


Figure 3: Connection of local VERONA networks to the technological network of the plant.

Since VERONA-t has exactly the same hardware and software components than the units, therefore the investigation of any event happening in the units can be done by VERONA-t without disturbing the normal operation of the units. It might be worth emphasizing here, that the operational regulations of the plant have strict rules for operation without the VERONA on-line core monitoring system (power must be reduced etc.), therefore any operational problem of this system can lead to significant economic loss. This is why, such events should be avoided or at least their occurrence must be reduced as much as possible. The application of VERONA-t test system is not the only way to reduce such an events. A more important approach to achieve high availability and safe operation of VERONA is the application of redundancy. In Fig. 2 one can see the architecture of the local network of VERONA for a reactor unit. PDA (Polyp Data Acquisition) is responsible to provide more than 2000 raw measurement data for VERONA. These raw data are processed in the very same way with two, redundant VDP (Verona Data Processing) servers VExHO001, VExHO003, where x={1,2,3,4} is the identifier of the reactor unit. After data processing VDP servers send the relevant reactor physical quantities to the two, redundant RPH (Reactor PHysics) servers (VExHO002, VExHO004), which are doing exactly the same time-consuming calculations by the support of built-in Tesla GPU cards. After finishing the calculations, the RPH servers send back the results to their VDP pair and the results are processed further by the VDP servers finishing one on-line data processing cycle.

A part of the measured and computed data are saved into the local archive of VDP servers and the archives are saved to NAS (VExNS001) periodically.

Data are displayed by thin clients (VExTC001…TC004), which are in connection with both physical VDP servers.

In the reactor unit 2 and 3, both local networks of VERONA have an additional server (VExHO005), which is called EXD (EXternal Display) server. Their role is to provide the possibility to connect any local VERONA network for ten external users and to display the actual processed data.

VERONA is in connection with the TN of the plant by two switches (VExSW001, VExSW002).

It is worth noting that not only the servers are redundant, but the structure of network also provides redundant connections between the servers and thin clients. Therefore, the malfunction of a system component could not lead to any degradation of high-level functionality of the system, i.e. the system is single-failure proof.

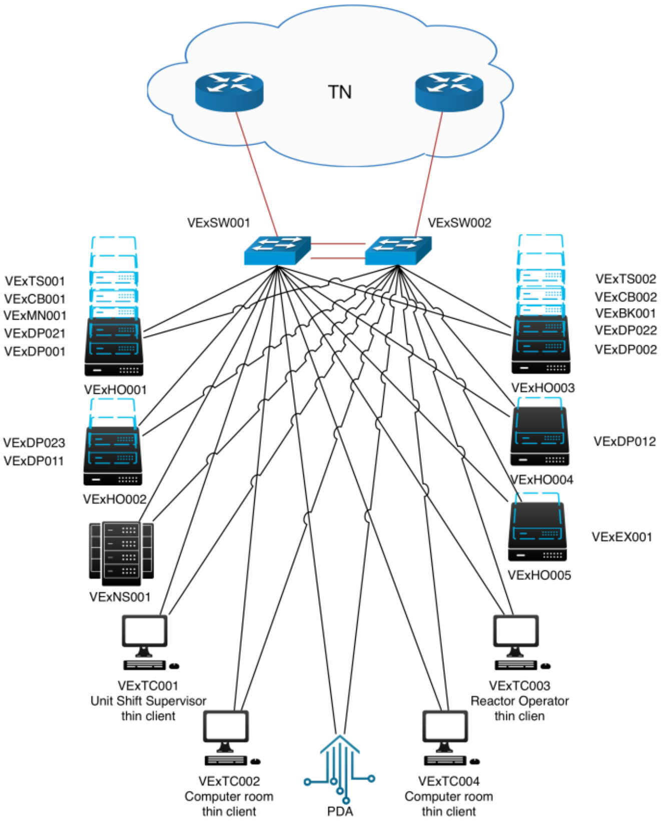


Figure 4: Local VERONA network.

In spite of its obvious benefits, virtualization has not been widely used in the nuclear industry, yet. However, during the reconstruction of VERONA, the application of virtualization was an important requirement from the plant's personal. The major motivations behind this request were to remove dependency on particular hardware vendors, to improve and speed-up disaster recovery, and to extend the lifecycle of applications.

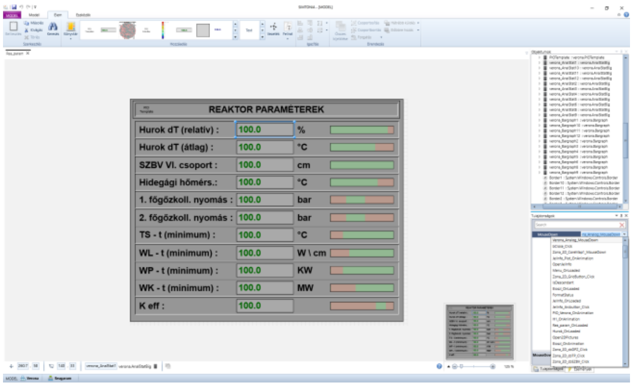
Therefore VMware's virtualization platform (ESXi 6.0) has been used in each physical server, but its advantages of virtualization were especially utilized in the real VDP servers. In the new version of VERONA, one real VDP server hosts five virtual servers integrating the functions of display (VExTS001, VExTS002), data processing (VExDP001, VExDP002), database (VExDP021, VExDP022), system management (VExMN001), backup (VExBK001) and connection broker servers (VExCB001, VExCB002) into one physical hardware (see Fig. 4). In the old system, physical display servers were used to monitor the measured and calculated data. In the new system, the results are shown by thin clients, which are connected to the virtual display servers.

Process monitoring in VERONA

In VERONA 7.0 the process monitoring of data is done by the visual engine of SIMTONIA instead of iFIX.

The visual engine is responsible for visualization of process variables, to receive and process user's interactions.

Each process picture is built up by SIMTONIA's diagram editor using some complex and several simple process elements developed by SIMTONIA's element editor. Each element has its own appearance, behaviour, animation and interaction procedure.



REAKTOR PARAMÉTEREK

Hurok dT (relatív) :	100.0	%
Hurok dT (átlag) :	100.0	°C
SZBV VI. csoport :	100.0	cm
Hidegági hőmérs.:	100.0	°C
1. főgőzkoll. nyomás :	100.0	bar
2. főgőzkoll. nyomás :	100.0	bar
TS - t (minimum) :	100.0	°C
WL - t (minimum) :	100.0	W \ cm
WP - t (minimum) :	100.0	KW
WK - t (minimum) :	100.0	MW
K eff :	100.0	

Figure 5: SIMTONIA's element editors during the development of a process monitoring display

A program called VDBIOServ runs in background in the virtual display servers. Its role is to request information from the OPCSend program running in the virtual VDP server. The visual engine of SIMTONIA communicates with VDBIOServ via the ProcessIO Application Programing Interface (API).

This API requests the data in very similar manner than an OPC DA interface using functions like AddGroup, AddItem, ReadItem, etc..). Basicaly it is a simpler version of OPC DA, but it does not use DCOM technology, simplifying the configuration of application programs. Regarding operation, its functions can be divided into two major categories:

- ProcessIO server functions,
- TCP server functions, which assure remote access to the VDB, VERONA's online database.

Functions belonging to these two major categories run in two different threads and they communicate with each other via some shared memory tables.

The ProcessIO server functions belong to SIMTONIA's ProcessIOSrv library, which provides an easy way to write programs, which can communicate with SIMTONIA's engines. The ProcessIOServ registers and provides data to the ProcessIO clients (display programs in this case) in an automatic manner.

For the proper operation of ProcessIO server, the VDBIOSrv has to create a data cache based on the data need to be displayed and some callback functions, which run when a new variable should be displayed (e.g. for checking the availability of variable in the database, or to add a new variable).

It is worth emphasizing that in contrast with the old system, where an individual set of process pictures had to be developed for each reactor unit by iFIX, SIMTONIA uses a single set of pictures and database variables can be accessed through a template resolving logic. It means that template pictures can be created by SIMTONIA's diagram editor and the variables defined according to this logic will be resolved and displayed in the picture. So, only the template has to be modified for a change in pictures, not each individual picture of each unit.

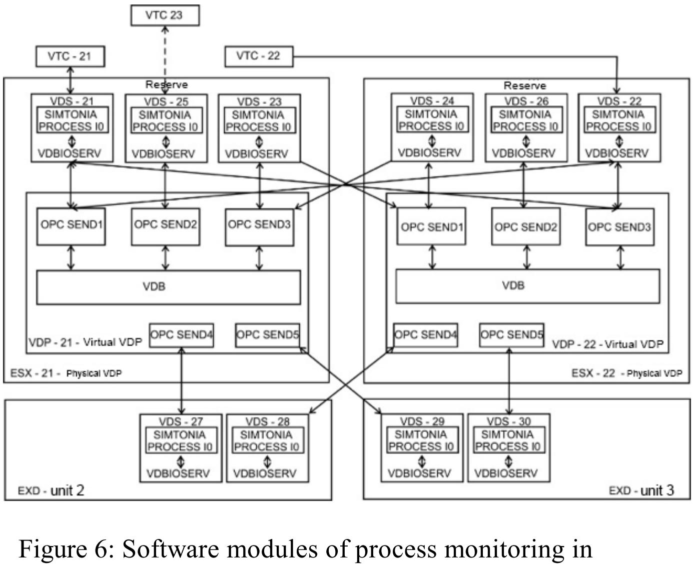
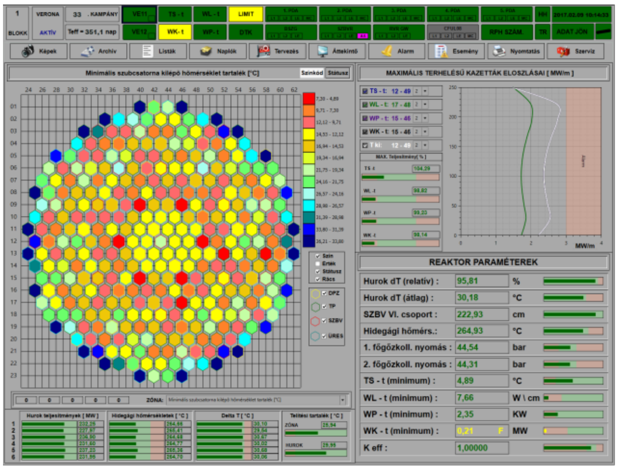


Figure 6: Software modules of process monitoring in VERONA

In Fig. 6 the connections between virtual VDP and display servers are shown focusing on the above mentioned software components. Here the connection of external display servers with the VDP servers can be seen, too.

Fig. 7 shows the main display screen of the VERONA system. The screen is made up of SIMTONIA elements, and each element has its own set of template variables. The data visualization elements placed on a picture are using these templates for the resolution of their database references. When a template variable change occurs in a picture, the visual engine automatically resolves the new references used on that picture by communicating with the VDBIOServ program.



REAKTOR PARAMÉTEREK

Hurok dT (relatív) :	95,81	%	
Hurok dT (átlag) :	30,18	°C	
SZBV VI. csoport :	222,93	cm	
Hidegági hőmérs.:	264,93	°C	
1. főgőzkoll. nyomás :	44,54	bar	
2. főgőzkoll. nyomás :	44,31	bar	
TS - t (minimum) :	4,89	°C	
WL - t (minimum) :	7,66	W \ cm	
WP - t (minimum) :	2,35	KW	
WK - t (minimum) :	0,21 F	MW	
K eff :	1,00000		

Figure 7: Main process picture of VERONA

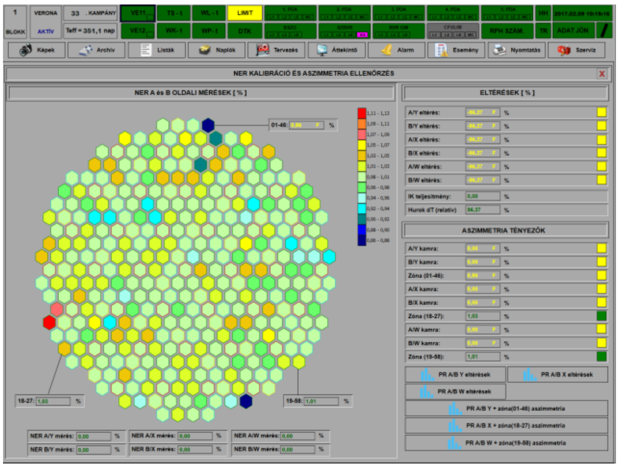

NER KALIBRÁCIÓ ÉS ASZIMMETRIA ELLENŐRZÉS
NER A és B OLDALI MÉRÉSEK [%]
ELTÉRÉSEK [%]
ASZIMMETRIA TÉNYEZŐK

Figure 8: Neutron detector calibration picture

Verification and validation of process monitoring

VERONA is a so-called safety level 3 information system in the plant, which means that it must have a very high level of availability (99.9%) during a reactor cycle. In case of its failure, there is only one hour to recover it. If the recovery is unsuccussful during this periof of time, the reactor power must be decreased and considering a longer time-scale of unavailablity the reactor nust be shutdown. Oviously, such events in an NPP have very tough ecomic consequences, therefore the replacement of VERONA's process monitoring subsystems was established with a detailed and carefuly planed test procedure. This test procedure included an important speedup test, in which we changed VERONA's 2s display cycle time to 2ms and drove VERONA by random input signals in order to study the stability and reliablity of process monitoring subsystem. After successful FAT, first the system was put into operation in the FSS of Paks NPP. The successful simulator trial period was followed by the installation of the VERONA-t test system in the 3rd unit. After VERONA-t's trial period, the new system was introduced at the end of 2015 in unit 3 and the other units in 2016.

EVALUATION OF CRITICAL SAFETY FUNCTIONS BY SIMTONIA

Nuclear reactor units has a so-called plant computer, which system gathers and evaluates all relevant physical parameters of the power plant.

The evaluation has two levels: The first level evaluation contains simple procedures, like signal conditioning, filtering etc. and focuses on individual signal processing. In the second level, more specific operational and safety parameters are derived (e.g. operation mode of the plant, complex safety margins of the unit etc.) from a set of signals processing them together. This kind of evaluation includes complex calculations, e.g. calculation of saturation temperature at the measured reactor power, and determination of relations between measured discrete signals using complex logic sequences.

In the past, this evaluation was carried out by an Excel table driven description of the evaluation algorithms in Paks NPP. Based on these tables a C source code was generated, compiled and linked to an executable file, which was run by the plant computer system.

There were several drawbacks of this approach. First, the modification of an algorithm was ponderous and time consuming. Furhermore, the resulting algorithm was not transparent for the end users, the operators, who were not able to identify easily, why a derived quantity takes its value in a certain situation.

In 2016, a refurbishment project of the old plant computer system has been started in Paks NPP and as part of this project the plant operational personal requested the introduction of a transparent second level process evaluation procedure. CER proposed to apply SIMTONIA's sequential engine for this purpose and the proposel was awarded.

Second level process evaluation by SIMTONIA

The altering of the old system to the new one included the following phases:

1. All complex calculation modules and simple logic elements used in the second level process evaluation of the old system have been implemented in a new element library of SIMTONIA.

2. The evaluation desciptor Excel tables were automaticaly converted to SIMTONIA logic schemes using the elements of the library.
3. SIMTONIA's sequence engine has been built into the new plant computer for second level process evaluation.
4. For each derived quantities (calculated in second level and shown in process monitoring displays), the corresponding evaluation scheme has been made available visualy for the users. That is, by clicking to a displayed derived quantity, SIMTONIA's corresponding logic scheme used for evaluation appears on the process monitoring screen and the animated picture makes it transparent, how the sequential engine derived the given quantity (Fig. 9).

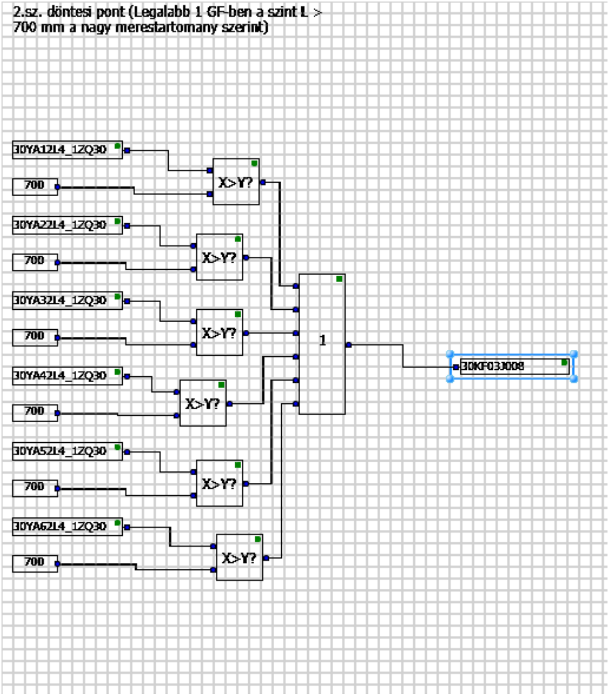


Figure 9: Evaluation of 30KF03J008 signal in SIMTONIA

CONCLUSIONS

Although SIMTONIA framework has been developed for building simulators of NPPs, its well established structure let us to apply its tools in other fields of nuclear engineering very efficiently.

In this paper, we have presented three industrial examples for its application. These successful applications and favourable operational experiences prove the stability and reliability of this system.

Since we provide a flexible licencing policy for the framework, we hope that in the near future we will find other possibilities for its application in other industrial fields.

REFERENCES

Páles, J.; Vécsi, Á.; Házi, G., 2017, SIMTONIA – A framework of SIMulation TOols for Nuclear Indsutrial Applications, 31st European Conference on Modeling and Simulation, (Budapest, Hungary May 23-26).

Jánosy J., Házi G., Seregi L., Szabó K., 1997 GRASS – The Graphic System Simulation, Proceedings of the 2nd CSNI Specialist Meeting on Simulators and Plant Analysers Olli Tiihonen (Eds.), (Espoo, Finland, Sept. 29 – Oct. 2)

Végh, J., et al., 2008. Core analysis at Paks NPP with a new generation of VERONA, Nuclear Engineering and Design 238, 1316-1331

Végh, I. Pós, Cs. Horváth, Z. Kálya, T. Parkó, M. Ignits: 2015, VERONA V6.22 – an Enhanced Reactor Analysis Tool Applied for Continuous Core Parameter Monitoring at Paks NPP, Nuclear Engineering and Design (2015), pp. 261-276

AUTHOR BIOGRAPHIES

JÓZSEF PÁLES was born in Hungary. He studied information technology on Pannon University and obtained his M.Sc. degree in 2006. He has been working at the Reactor Monitoring and Simulator Department for the Hungarian Academy of Sciences Centre for Energy Research since 2002. His e-mail address is : pales.jozsef@energia.mta.hu.

ÁRON VÉCSI was born in Miskolc, Budapest and went to the Budapest University of Technology and Economics, where he studied physics and obtained his M.Sc. degree in 2016. He is now PhD student and working at the Reactor Monitoring and Simulator Department for the Hungarian Academy of Science Centre for Energy Research. His e-mail address is: vecsi.aron@energia.mta.hu

GÁBOR HÁZI was born in Budapest, Hungary and obtained his B.Sc. in Kando College in 1992 and M.Sc. in electrical engineering at Technical University of Budapest in 1995. He holds a PhD in nuclear engineering from the Technical University of Budapest and DSc from the Hungarian Academy of Science. He has been working for CER since 1992. Now, he is the head of Reactor Monitoring and Simulator Depament. His e-mail address is : hazi.gabor@energia.mta.hu.

CAE/VR INTEGRATION – A PATH TO FOLLOW? A VALIDATION BASED ON INDUSTRIAL USE

Holger Graf and André Stork
Department of Virtual and Augmented Reality, Department of Interactive Engineering Technologies
Fraunhofer IGD
D-64823, Darmstadt, Germany
E-mail: {holger.graf|andre.stork}@igd.fraunhofer.de

KEYWORDS
Virtual Reality (VR), CAE/VR Integration, Interactive Preliminary Design Analysis, Industrial Evaluation

ABSTRACT

Numerical simulations have become crucial during the product development process (PDP) for predicting and validating different properties of new products as well as the simulation of various kinds of natural phenomena. Especially the engineering domain (CAE – Computer Aided Engineering), is seeking for new ICT solutions to cover broad ranges of physical simulations. Virtual Reality (VR) has matured in the past allowing a rapid consolidation of information and decision-making through visualization and experience. These new man machine interfaces offer advanced interaction possibilities with the digital domain and allow engineers to variate over several hypothesis. This enlightened ideas to deploy VR for "what-if-scenarios" also in the CAE domain. However, while CAD/VR integration has been sufficiently researched, the integration of CAE into VR is still facing a long road ahead. Despite recent criticism that the application of VR technology has been considered unnecessary in CAE, this paper aims to refute this by presenting methodologies for linear static FEM analysis allowing "what-if-scenarios" within interactive environments. It validates the elaborated methodologies and advantages of VR front ends by an evaluation performed within industrial engineering departments.

INTROCUCTION

Computer Aided Engineering methods have influenced the PDP significantly. Popular application areas include structural mechanics (e.g. stress analysis, dynamic analysis, modal analysis, kinematics) and computational fluid dynamics (such as the analysis and validation of fluidal behaviour). Following an analysis step, engineers have been provided with an insight into material characteristics or physical properties. Those are described by numerical values of some kind of physical quantities (e.g. pressure, stress concentration, deformations, velocities) that are available at discrete points within space. Nevertheless, computer engineering methods are complex and resource intensive, consuming significant computational power and time, thus leading to long and cost intensive analysis processes. On the other hand, the demand for shortening the time intervals within the product life cycle is crucial to remain competitive.

Here, the trend to front-loading of product life-cycle steps in order to simulate a design, validate its behaviour and derive decisions about its manufacturing has become more and more important during the last decade. Especially, tools to allow fast decision-making through an intuitive grasp of the situation have improved to reduce development cycles. One of the key driving technologies for a better communication, representation, interaction, and visualisation of design and engineering/manufacturing data has been Virtual Reality (VR). Though mainly developments in the CAD/VR domain with an emphasis on an integration into the product development process have been driving research efforts, applications demonstrated the advantage of being able to review interactively designs, conduct ergonomic studies and check feasibility of assemblies. The potential of VR for postprocessing of engineering/manufacturing data raised hopes to deploy this advanced man machine interface for *interactive conceptual simulations (ICS)* within the CAE domain (Graf and Stork 2011), (Graf and Stork 2013). While it is a valid hypothesis, still many challenges and problems remain due to the nature of the "change'n play" paradigm imposed by conceptual simulations, the real-time operations within a VR environment as well as the fragmentation of tools and workflows used especially within CAE. Choi et al. (Choi et al. 2015) presented a comprehensive literature survey of VR research within the product development since the 90's and showed, that VR technology has been mostly applied to design reviews and assembly tests of products, whereas research endeavours for an integration of VR within the computer-aided engineering (CAE) domain remained very little. Based on the number count of publications the authors' stated *"…This likely means that we can only achieve relatively little profit by applying VR to CAE. The application of VR technology has been considered unnecessary in CAE: relatively simple visualization provided enough support to make decisions.[…]"*, (Choi et al. 2015).

Aim of this paper is therefore twofold: First, it aims at presenting recent research results for CAE/VR integration, and secondly, objecting the above

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

hypothesis, by presenting an evaluation that has been conducted with automotive, aerospace, and urban planning engineering departments demonstrating the benefit and profit of using VR within engineering domains. As a conclusion it aims at answering the question "CAE/VR integration – A Path to Follow?".

To start, we should briefly recall the complications, that researchers face in order to deal with a CAE/VR process chain.

"WHAT-IF-SCENARIOS" IN CAE-VR

In typical CAE postprocessing sessions, engineers are able to get an in depth snapshot of a specific physical behaviour, and depending on the maturity of the visualisation tool, a tidy showcase. In many cases, however, analysts are much more interested in creating effects due to some physical phenomena, changing parameters, changing materials, changing constraints, changing domain, and actively wanting to drive the engineering discovery process within *"What-if-analysis"*[1] set-ups. Nevertheless, interactive conceptual changes within engineering analysis require real-time turn-around loops for model manipulation, simulation and visualisation. Depending on the complexity of the underlying models, real-time FEA simulations are not possible in complex industrial applications. This usually antagonises ideas for VR-based front-ends. Computer Graphic approaches (e.g. Georgii 2008, Weber 2016) for real-time FEA in order to simulate deformations and visualise animations rely on simplified mathematical models, that do not comply with the eningeering demand for precise validation of a physical phenomena.

Any proposed new method for an appropriate interactive exploration, suitable for engineers, requires integrated workflows. Isolated solutions lead to the bottleneck of data transfer, data preparation and interpretation, spanning different phases of the engineering analysis workflow. Here, the workflow is inherently depended on the CAD/CAE transfer process (Tierney et al. 2015), which is error prone and time consuming. De-featuring and cleaning/healing of models are typically done within dedicated tools, whereas the meshing of the domain in others. Thus, many media breaks inbetween tools lead to long design/re-design iterations for the domain.

IDEFIx

This section should provide a brief overview of the framework IDEFIx (Interactive Data Exploration for Immersive CAx), that has been developed based on a consensus on engineering requirements which have been collected throughout this research work. It summarises the results of a range of own research work, and for implementation details the reader should be referred to those papers. Thus it leaves space for an evaluation that has not been presented. Several techniques and methodologies of IDEFIx aim at faster turn-around loops within engineering tasks for *linear static analysis* and some basic simulation possibilities for CFD (Computational Fluid Dynamics). They are embedded within one sinlge point of access workplace opening engineers a seamless access to several tools required for preliminary design analysis. Main efforts have been dedicated to a design which respects the pragmatic engineering demand of a scalable integration into existing environments. Here, VR should not be used as an additional tool rather than an operational one, from which several CAE tasks can be steered if required. Tasks that cannot be performed within VR, should have been made available at the workplace also, requiring access to legacy systems for engineers. The introduced hybrid workplace, thus, is based on the capability to support traditional point-and-click applications as well as innovative VR-based postprocessing steps. It is powered by a service oriented approach, see e.g. (Graf 2011), that enables access to pre-processor functionalities, simulation module and VR postprocessor. An adequate event mechanism is capable of distributing requests from one service to other services interested in that query. As a major result several preprocessing and postprocessing steps could be integrated into one environment supporting typical CAD/CAE processes and engineering analysis steps in an immersive set-up.

Remeshing within postprocessing tools coupled to advanced, sensory based interaction possibilities and a closed re-simulation loop has not been realised up to now and provides engineers the capability to shorten their design/analysis cycles.

In view of the ICS methodology the developed techniques can be performed at different levels of interactivity combining advancements in the CAD/CAE transfer as well as VR techniques easing the variations of the underlying domain. IDEFIx offers intuitive postprocessing with preprocessing capabilities as well as coupled simulation runs, allows for conceptual changes of the basis domain by moving features or groups of mesh entities based on ensured consistencies by new element quality metrics, adaptive local refinements (see Figure 1), and offers realtime simulations by steering computation through loadcase manipulation and re-modelling of the underlying domain. Aside a coherent software engineering approach, several strategies embedded within the framework are based on formalised mathematical models that build the basis for an automation of numerical simulation processes not being available until now. New error estimates enable adaptive refinements that can be applied within user-centric and goal-oriented conceptual simulations (Larson et al. 2014) for linear static analysis, see Figure 2. This in turn enables faster

[1] "*What-if-analysis"* within this work is meant as to investigating "what-if-scenarios" within engineering analysis tasks, i.e. playing on the real world effect for various design decisions, such as feature and de-featuring of the CAD model, changing analysis mock-up, boundary conditions, acquiring a higher level of detail or interpreting results in different presentations.

engineering cycles based on numerical simulations and thus shortens the turn around cycles within the engineering analysis domain. For the evaluation we considered few basic functionalities that could be performed with IDEFIx:

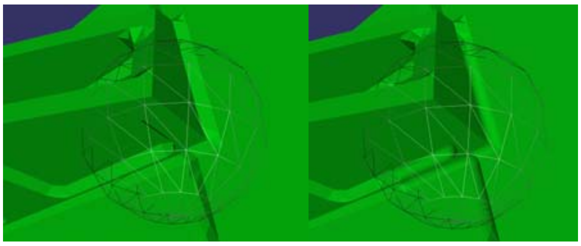


Figure 1 Adaptive meshing and local mesh refinements

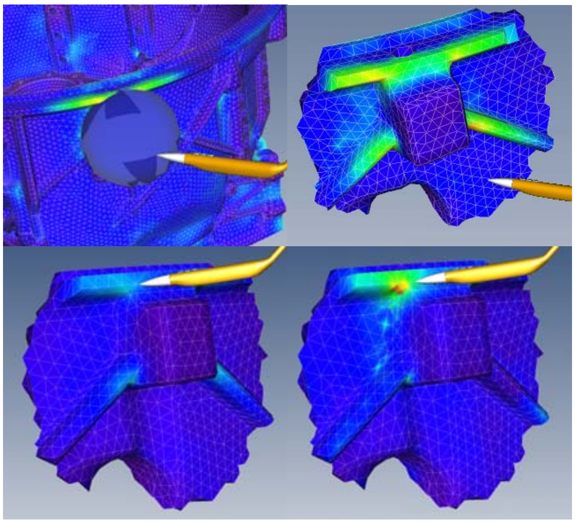


Figure 2 Conceptual simulation – local enhancement; Domain decomposition of complex domain, get deeper insight into local problems (Refinement/Re-analysis).

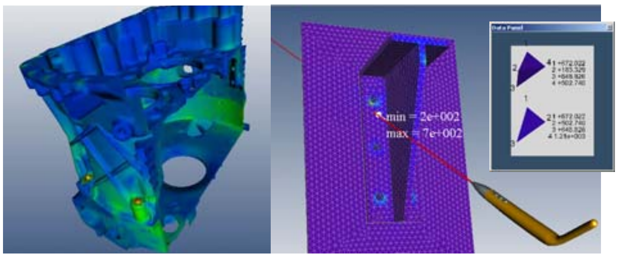


Figure 3 Interactive postprocessing of CAE data (left, mid: cross sectioning with clipping, right: element probing using laser beam metaphor)

As workplace concept we used a hybrid desktop that allowed us to integrate CAE and VR into one workplace.

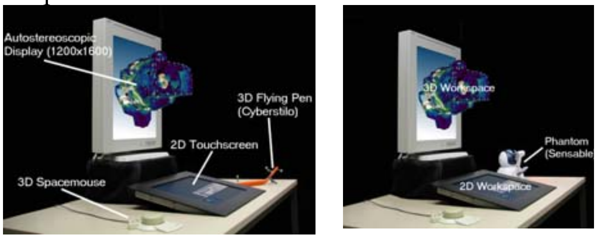


Figure 4 Hybrid Desktop combined with optical tracking system and different 6DOF input devices.

To address the shortcomings of an ergonomic use of advanced VR devices such as shutter glasses, interaction, or output devices, a system specifically designed for the desktop application of VR within an engineering environment has been proposed in earlier own work, e.g. (Graf and Stork, 2013). It is based on an autostereoscopic display paired with a second LCD touch screen in an L-shape configuration (see Figure 4). The set-up relieves the user from wearing shutter glasses or HMDs opposed to "classical" VR installations. As input device a wand like device is used that allows for 6 DOF interactions (Cyberstilo©) as well as is capable of creating 2D input for the touch screen. This allows for an easy navigation and interaction in a "*hybrid*" configuration. Interactions are possible with traditional legacy point-and-click applications inherent to the engineering domain but also 3D-based navigation and system control within the VR environment.

EVALUATION

Methodology

We conducted specific validation sessions addressing selected workflows of engineers being defined for the VR enhanced workplace. Here, we were setting up the hybrid desktop in which the engineer was capable of using legacy systems and VR tools for model preparation, model simplification, simulation and analysis. Several validation sessions were conducted in three iterations over a total period of four months within a first tier OEM (ZF[2]), Airbus and CSTB[3], covering the requirements from automotive, aerospace and urban planning departments[4]. As in-house COTS ("Commercial-Off-the-Shelf")- tools MSC.Patran/ Nastran, PERMAS, IDEAS-NX, ANSYS Fluent were used. Three questionnaires suited as basis to gather feedback from the test subjects. The intention was to collect qualitative as well as quantitative data during each validation session. The time required to conduct certain tasks within the new integrated workflow is considered as the quantitative parameter within the subsequent evaluation. In a first iteration focus was given to the completeness and availability of the required functionalities in order to conduct certain tasks. A subsequent validation session addressed the integration of tasks within one workflow and finally the system was benchmarked in a real life setting in which two installations were moved to end user sites and were evaluated under daily working conditions over a total period of two months. At the final stage, specific benchmarks were conducted in order to compare time savings of the new approach with the traditional workflow. Due to the limitations of this paper, only

[2] ZF – ZF Friedrichshafen, www.zf.de

[3] CSTB – Centre Scientific et Technologie du Batiment, France, www.cstb.fr

[4] It should be noted that the urban planning scenarios have been restricted to "simulate-'n-explore", rather local enhancements.

some findings can be presented. We focus on the most interesting ones.

Participants

The group of test subjects comprised experts from two engineering domains, 6 from engineering analysis and 3 from urban planning. In general, all involved experts from the CAE/CFD domain had little experience with VR technology. None of them had VR or user interface design experience, the majority is using traditional keyboard and mouse interactions, and several test subjects were using standard (or dual) LCD output screens.

Test Cases and Tasks

Three test cases from each domain had to be examined and analysed: A small and medium size model to verify the entire process chain within solid mechanics applications, and a large model to experience the current limits of the approach and to check the postprocessing capabilities within the hybrid desktop. The engineers had to accomplish the task list for each model:

- Import mesh with results from external solver
- Change visualisation options
- Display and evaluate principal stresses
- Display displacements and animate
- Conduct a submodel analysis
- Conceptual simulation (local enhancement)
- Define local mesh refinement (re-mesh)
- Solve local mesh (re-simulate)
- Postprocess sub-/local model (use freely definable cross section and element probes, based on a laser beam metaphor 'glued' to the virtual interaction device)
- Compare refinements and result sets

Analysis

In order to analyse the questionnaires, the experiments conducted by the test subjects were modelled as discrete random variables. The realizations of a random variable are called random variates. Thus, let Ω be a probability space and E a measurable space. The random variables $X_i : \Omega \to E, i = 1 ... l$, model a stochastic process on Ω, with $X_i^{-1}(A) \doteq \{\omega \mid X(\omega) \in A\}, A \subset E$. For l observations $X_i\ (\omega) = x_{i,n_i}, i = 1,...,l$ of samples with size $n = n_1 + ... + n_l$ the realisations (variates) are x_{i,n_l}.

For the analysis of returned feeback by the test subjects the software suite SPSS of IBM was used in order to derive the descriptive statistics of several modelled variables. The results have been analysed according to the descriptives: mean (M), variance (V) and standard deviation (SD). Where required, further information to the presentation is added, such as quartiles q_{25}, q_{50}, q_{75} of an observation. Quantitative evaluations are based on a one-way ANOVA (Analyis of Variance) which provides a statistical test whether three or more means are the same thus testing compared observations to be equal ("null hypothesis").

For the *quantitative* evaluation it uses a F-test based on the ratio of two scaled sums of squares reflecting different sources of variability ("degrees of freedom" /"df"). The construction of the test provides a possibility to evaluate the variational differences between obeservations in a way that its quantum tends to be greater if the null hypothesis is not true.

QUANTITATIVE ASSESSMENT

This section presents some quantitative benchmarks in order to compare the new developments with in-house COTS-tools. For this, the experts did use the following models

Table 1 Models used during the evaluation

	KB	Nb Elements	DOF
smaller models	12.300	23.000	120.000
larger models	154.000	332.000	1.800.000

The following table summarise the benchmarks for smaller sized and medium sized models in view of the postprocessing and coupled interactive refinement/re-simulation/postprocessing loop. A detailed presentation of the quantitative assessment on submodelling and local enhancement techniques for very large models, the reader should be referred to (Larson et al. 14).

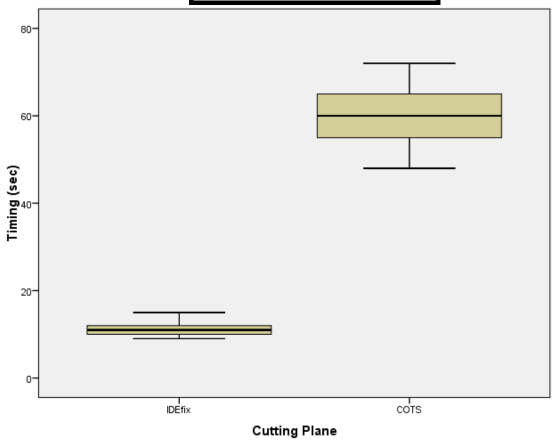
Postprocessing

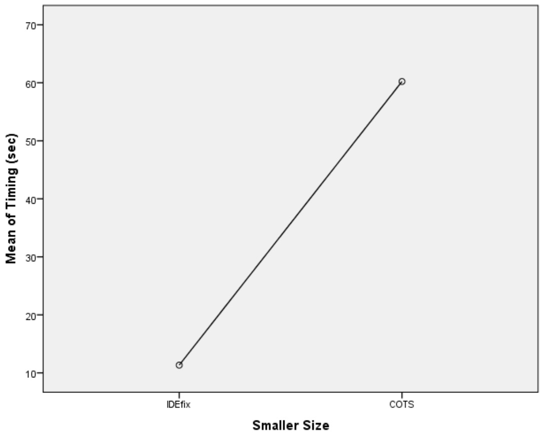
In a first step the experts had to identify area of high stress using typical postprocessing functionality. Here the time was captured from the start of the postprocessing session until he did come to a final assessment of the mock-up. Upload time was not taken into account. During a second step, the most important modalities data probing and cutting planes (i.e. cross sectioning incl. data probes incl. assessment) were evaluated separately (Table 2, Table 3).

Table 2 Quantitative Analysis 'Postprocessing' comparing IDEfix and COTS – Cutting Planes

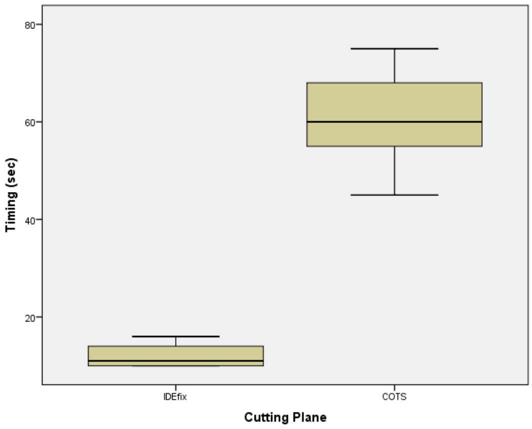
Smaller Models	Descriptives (sec)	Std. Error	Smaller Models	df	F	Sig.
IDEfix	M(9) = 11,333	,645	Between	1	333,793	,000
	V(9) = 3,750		Groups	16		
	SD(9) = 1,936		Within			

			Groups			
COTS	M(9) = 60,222 V(9) = 60,69 SD(9) = 7,79	2,596				





Medium Sized	Descriptives (sec)	Std. Error	Medium Sized	df	F	Sig.
IDEfix	M(9) = 12,11 V(9) = 5,361 SD(9) = 2,315	,771	Between Groups Within Groups	1 16	213,591	,000
COTS	M(9) = 60,556 V(9) = 93,528 SD(9) = 9,67	3,222				



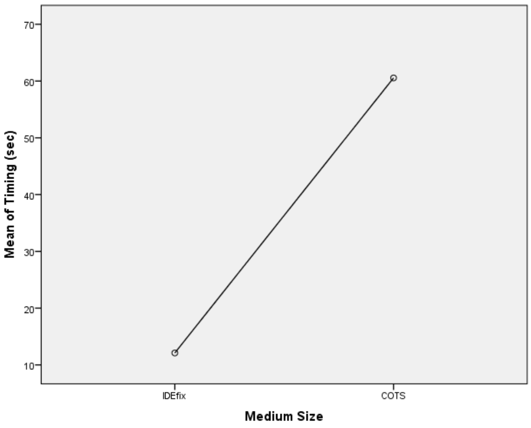
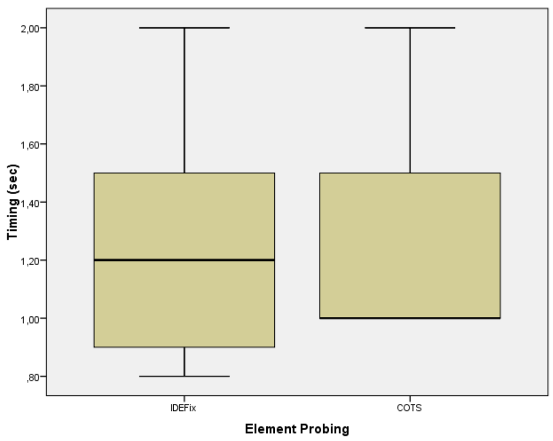
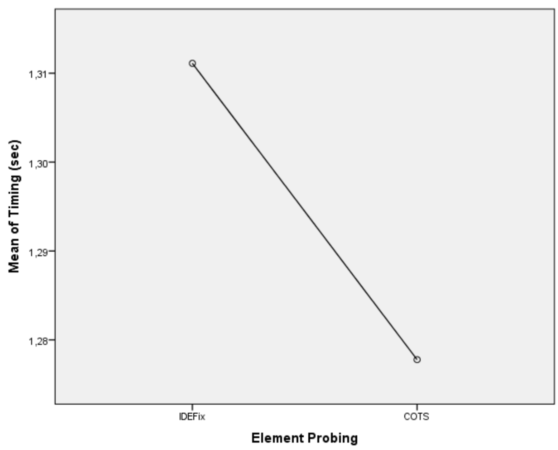


Table 3 Quantitative Analysis ‘Postprocessing’ comparing IDEfix and COTS – Element Probing

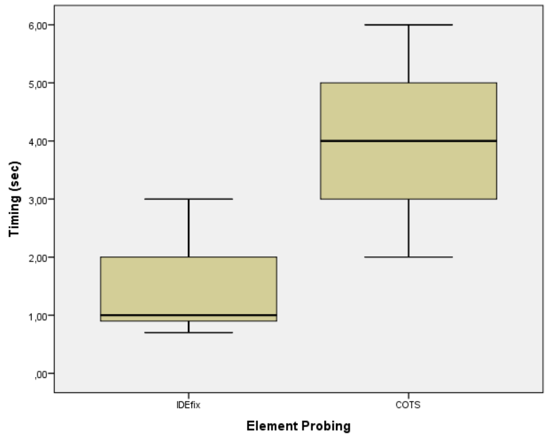
Smaller Models	Descriptives (sec)	Std. Error	Smaller Models	df	F	Sig.
IDEfix	M(9) = 1,3111 V(9) = 0,216 SD(9) = 0,464	,154	Between Groups Within Groups	1 16	,024	,878
COTS	M(9) = 0,9388 V(9) = 0,194 SD(9) = 0,441	,147				

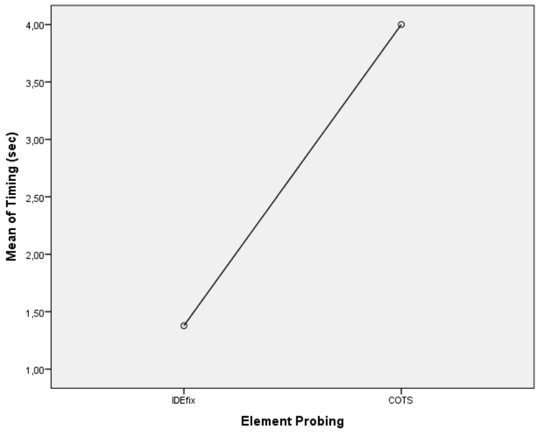




Medium Sized	Descriptives (sec)	Std. Error
IDEfix	M(9) = 1,377 V(9) = 0,607 SD(9) = 0,779	,259
COTS	M(9) = 4,00 V(9) = 2,00 SD(9) = 1,414	0,4714

Medium Sized	df	F	Sig.
Between Groups	1	23,738	,000
Within Groups	16		





According to the F-statistics, IDEfix offers a significant better performance in several used test cases (except the element probes on smaller models). This is in line with the qualitative assessment (below) and being observed by several engineers. Nevertheless, the postprocessing tasks in general are better supported by IDEfix as indicated by the descriptives and related figures.

QUALITATIVE ASSESSMENT OF IDEFIx

This section reflects on the results obtained for a qualitative assessment of IDEfix according to the task list and related questionnaire. As first task the user had to evaluate on the clarity, availability, efficiency and fulfilment of the requested functionality within the VR component. For both domains the implemented modes for visualisation and interaction have been similar. The variables $X_i : \Omega \rightarrow E, E = \{1, 2, ..., 5\}, \Omega =$ (“not available”, “poor”, “satisfactory”, “good”, “very good”) reflect on “clarity (c)”, “efficiency (e)”, “consistency (cs)” as well as “fulfilment of demands (f)” of the developed solutions. Here, the consistency reflects on the need of the functionality to be competitive to in-house COTS-tools. The mark “good” mirrors the same level of possibility provided by typical working tools, “very good” being superior to the traditional way. The efficiency reflects on the need of the functionality to allow the engineer to quickly get to an assessment of the problem areas or the overall mock-up. The higher the mark the faster he could accomplish his task.

Conceptual Simulation – Local Enhancement

The following table (Table 4) indicates the results for the implemented conceptual simulation based on substructuring and local enhancements of the resolution within structural mechanics. As it was not implemented for CFD and urban planning only feedback of the two involved structural mechanics departments has been gathered. It was expected that this functionality improves the workflow of current processes for substructuring and local refinements. During the first testing session the test subjects had to provide feedback on the available integrated simulation/visualisation/post-processing functionality and state their perception on the current realisation. A high ranking of the efficiency is expected.

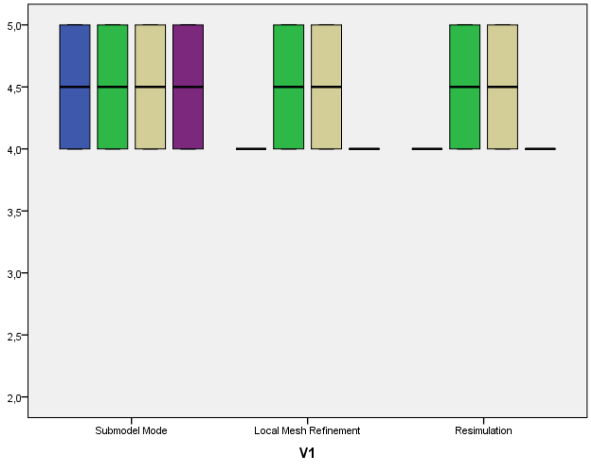
Table 4 Qualitative assessment of IDEfix – Conceptual simulation/substructuring/local enhancement

"Clarity" (c)	Descriptives	Std. Error
Submodel Mode	M(6) = 4,50 V(6) = 0,500 SD(6) = 0,707	,500
Local Mesh Refinement	M(6) = 4,0 V(6) = SD(6) = 0	,0
Re-simulation	M(6) = 4,0 V(6) = SD(6) = 0	,0

"Efficiency" (e)	Descriptives	Std. Error
Submodel Mode	M(6) = 4,60 V(6) = 0,300 SD(6) = 0,548	,245
Local Mesh Refinement	M(6) = 4,50 V(6) = 0,500 SD(6) = 0,707	,500
Re-simulation	M(6) = 4,60 V(6) = 0,300 SD(6) = 0,548	,245

"Consistency" (cs)	Descriptives	Std. Error
Submodel Mode	M(6) = 4,50 V(6) = 0,500 SD(6) = 0,707	,500
Local Mesh Refinement	M(6) = 4,50 V(6) = 0,500 SD(6) = 0,707	,500
Re-simulation	M(6) = 4,50 V(6) = 0,500 SD(6) = 0,707	,500

"Fulfilment" (f)	Descriptives	Std. Error
Submodel Mode	M(4) = 4,75 V(6) = 0,333 SD(6) = 0,577	,333
Local Mesh Refinement	M(4) = 4,0 V(6) = SD(6) = 0	,0
Re-simulation	M(4) = 4,0 V(6) = SD(6) = 0	,0



(As several observations achieved values above 3 ("satisfactory"), it was decided to start the y-scale at 2 in order to maximize scaling factors for the purpose of visualization)

X_i(not available, poor, satisfactory, good, very good) → $X_i \in \{1,2,3,4,5\}$, $i \in \{c,e,cs,f\}$

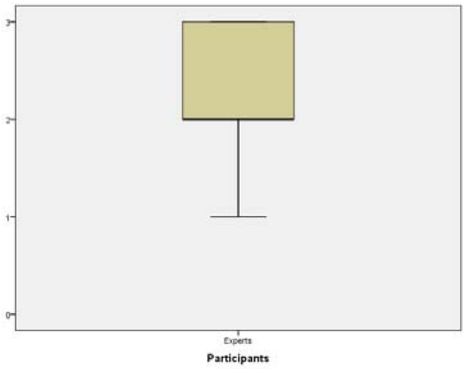
- Several functionalities realising the conceptual simulation based on substructuring and refinement process were perceived as "very good". The fulfillment of the demands was marked as "good" for local resolution enhancements.
- The engineers provided very positive feedback and expected to cut down individual substructuring and local remeshing tasks by 3-5 working hours.
- In view of the defined scenarios the realisation was sufficient, however, specialists need a larger number of elements being involved in an analyis.
- In general, the experts were missing the support of more elements for a dedicated analysis. As very positive statement, the experts were satisfied with the speed of simulation and re-simulation within an interactive environment.
- Some quotes of the free-text fields:
  - *"This function is extremely practical for large models" (ZF).*
  - *"The postprocessing tool with the mesh refinement and the nearly real time solver is the most valuable part" (ZF).*
  - *"In the VR environment Mesh refinements can easily be performed. The speed is outstanding compared to standard technologies. All tests of remeshing have been performed just within a few seconds only. A similar remeshing with MSC.Patran needs minutes." (Airbus)*
  - *"The advantages in standard packages are that they give the user more capabilities in influencing the meshes. For this reason the new approach is better for casual users (which needs quick and easy to use tools) whereas the standard tools are better for specialists" (Airbus)*

Assessment of the Hybrid Workspace

This section presents the results obtained for the validation of the overall Hybrid Workspace. The experts have been asked whether a hybrid approach embedding 2D and 3D functionality into one workspace concept and hybrid objects are suitable for daily work, whether they got exhausted, how long they might be able to work on it and if they would deploy VR as immersive CAE modelling tool. The variables have been modelled according to different variates. Thus the $X_i : \Omega_i \to E_i, E_i \subset \{1,2,...,5\}$. The following table (Table 5) shows the final results

Table 5 Hybrid Workspace – Suitability, exhaustiveness, duration of work, hybrid objects

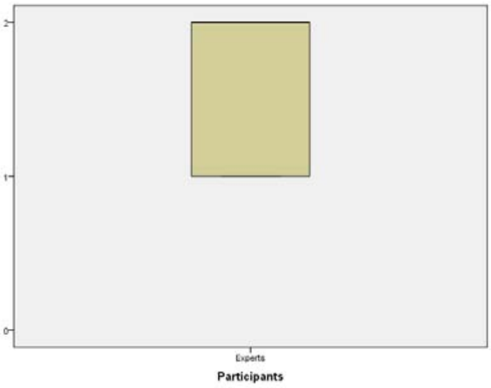
Variable	Descriptives	Std. Error
Suitability of Workspace	M(9) = 2,33 V(9) = 0,500 SD(9) = 0,707	,236



X_{Suit}(obsolete, not very useful, quite useful, useful) → $X_{Suit} \in \{0,1,2,3\}$,

- Up to 75% of the feedback indicated the usefulness of the concept. The mean of 2,33 indicates that more than half of the experts found the concept "quite useful" for daily work.

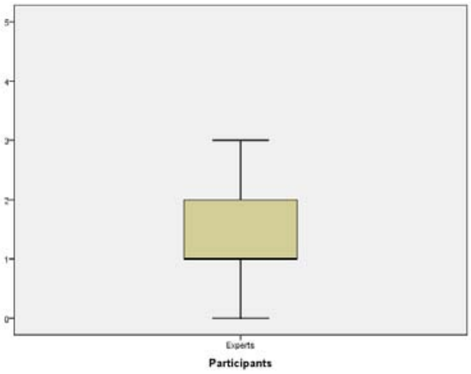
Variable	Descriptives	Std. Error
Exhaustiveness	M(9) = 1,67 V(9) = 0,250 SD(9) = 0,500	,167



X_{Exh}(no, not sure, yes) → $X_{Exh} \in \{0,1,2\}$,

- Several experts felt exhaustive working with the hybrid workspace. This has several reasons:
- The new interaction device has been too heavy for hourly work,
- the experts got tired working with the autostereoscopic display, indicating the change of depth perception from 2D space into 3D space is causing eye strain, and
- workspace ergonomics lacked due to aiming at very precise positioning tasks (back/neck stiffness were reported).

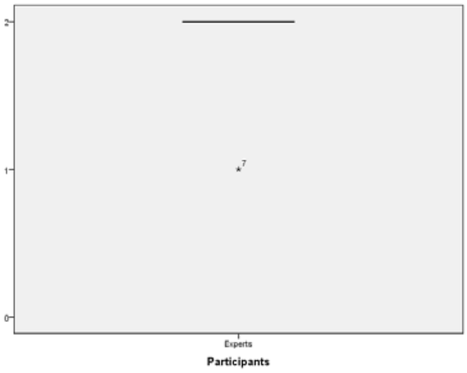
Variable	Descriptives	Std. Error
Duration	M(9) = 1,89 V(9) = 0,861 SD(9) = 0,928	,309



X_{Dur}("<1","1-2","2-3","3-4","4-5","5-6") → $X_{Dur} \in \{0,1,\ldots,5\}$

- The experts indicated to be able to work in average 1-2 hrs with the interactive desktop.
- In general up to 75% of the feedback ($q_{X_{Dur}75} = 2$) indicated to work longer, but no more than 2-3 hrs.

Variable	Descriptives	Std. Error
Hybrid Objects	M(9) = 1,89 V(9) = 0,111 SD(9) = 0,333	,111



X_{hUI}(not suitable at all, confusing, suitable) → $X_{hUI} \in \{0,1,2\}$

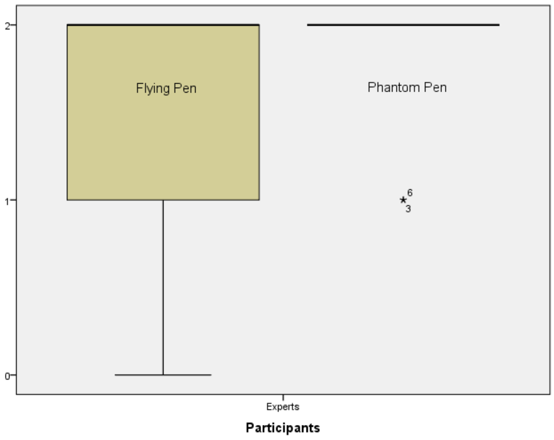
- Asked for the work with hybrid user interface elements, the engineers found it important to also steer the VR environment from the desktop. Thus only one outlier indicated that the user interface is confusing.

Interaction Paradigm

Within the different on-site set-ups several variants of interaction devices and paradigms were tested (Table 6): One, using the Cyberstilo© (FP) and the other using a phantom pen haptic device (PP). Furthermore, engineers were asked for the dislike or like of an "object in hand" metaphor at a virtual table (VT – a table like backprojection system used with active stereo glasses), indicating that they were able to keep a tracked artefact (assigned to the digital mock-up) in their free hands. This metaphor was tested subsequently at the hybrid desktop (HD).

Table 6 Interaction paradigm – Suitability, ergonomy, "object in hand"

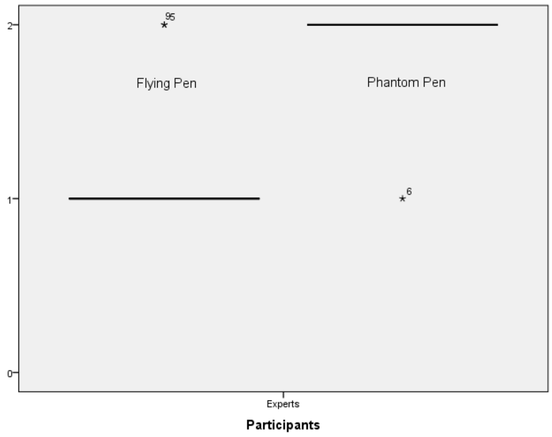
Variable	Descriptives	Std. Error
Interaction Paradigm (FP)	M(9) = 1,44 V(9) = 0,528 SD(9) = 0,726	,242
Interaction Paradigm (PP)	M(9) = 1,78 V(9) = 0,194 SD(9) = 0,441	,147



X_{IntP} (not useful at all, not sure, easy to use) → $X_{IntP} \in \{0,1,2\}$

- The evaluation of the preferred interaction paradigm based on the new interaction device and a phantom pen indicated a clear preference to the phantom pen.
- This is mainly due to the control provided by the haptic device. The degrees of freedom are limited to a small interaction volume in contrary to the flying pen.

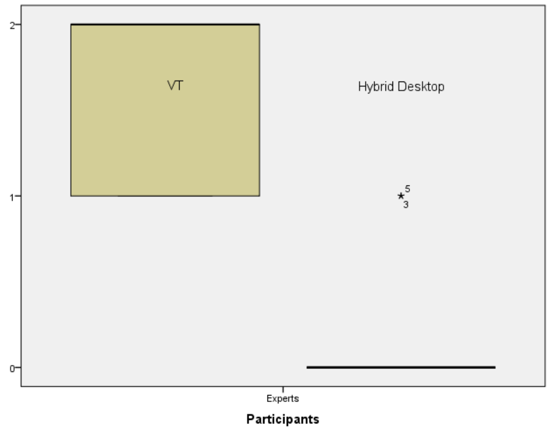
Variable	Descriptives	Std. Error
Ergonomy (FP)	M(9) = 1,22 V(9) = 0,194 SD(9) = 0,441	,147
Ergonomy (PP)	M(9) = 1,89 V(9) = 0,111 SD(9) = 0,333	,111



X_{Erg} (does not fit, needs to be improved, does fit) → $X_{Erg} \in \{0,1,2\}$

- The evaluation on ergonomical aspects reveiled that the flying pen requires improvements, whereas the phantom pen fits well and does not burden any heavy load.
- As main drawbacks, the experts named the weight, and the dimensions of the pen.

Variable	Descriptives	Std. Error
Object in Hand (VT)	M(9) = 1,44 V(9) = 0,528 SD(9) = 0,726	,242
Object in Hand (HD)	M(9) = 1,44 V(9) = 0,528 SD(9) = 0,726	,242



X_{IntP} (no, not sure, yes) → $X_{IntP} \in \{0,1,2\}$

- Asked for the like (yes) or dislike (no) of keeping an object in hand a clear voting for this metaphor was fed back by the experts.
- However, at the desktop they disliked this idea mainly due to the limited interaction space as well as imprecision introduced by additional degrees of freedom.

Remarks within the free-text fields indicated a high degree of satisfaction with the realisation of the Hybrid Desktop and integrated VR environment. Some excerpts of the validation protocols:

Hybrid 2D/3D Set-Up

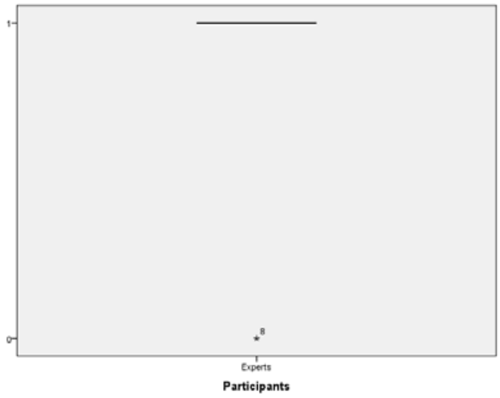
- *"The hybrid setup is from our testing experience mandatory, as it seems not to be suitable to perform all tasks in VR only. Performing the work in such a hybrid setup is suitable. Display quality is more or less ok as a starting point, but the area of finding positions for a* *good 3D view, needs to be expanded. Also all screens should be not too small" (Airbus).*

- *"The display seems to be in general suitable, but not for an 8 hour day. As an additional display for specific tasks, it is ok. Technological enhancements may lead soon to an acceptable 8 hour usage" (Airbus).*
- *"Estimation of time savings: At least 50% for all tasks needing 3D perception: Precise interaction with the model (picking, BC definition, data probing, area selection for submodelling). Visualisation / analyze of simulation results." (CSTB).*
- *"Benefits working with such a system during the process chain:*
  - *Increased automation to clear faults in CAD-models*
  - *Speed up of the meshing process*
  - *More precise results with the intuitive mesh refinement tool" (ZF)*
  - *"Enhanced understanding of non experts for calculation results" (ZF).*
  - *"Advantages: Very accurate display of a 3D structure. Low cost" (Airbus).*
  - *"This hybrid 2D/3D approach is very interesting, especially is the 2D GUI is used via a touchscreen" (CSTB).*

Finally, asked for the take-up and use of VR within their environments, several experts would make use of VR within their environment (Table 7)

Table 7 Hybrid Workspace – Use of VR for CAE tasks

Variable	Descriptives	Std. Error
Immersive Use	M(9) = 0,89 V(9) = 0,111 SD(9) = 0,333	,111



X_{ImU}("no", "yes") → $X_{ImU} \in \{0,1\}$

- The feedback did indicate that the engineers would be in favour of using VR within their workflow. However:
  - as additional tool not as a complement to current COTS, and
  - embedded within a hybrid workspace concept.

CONCLUSION

The open framework exposes an integrated CAD-to-CAE-to-VR process chain and an inclusion of typical preprocessing steps such as automatic geometry clean-up into a VR-based analysis tool. We have shown in earlier work that with our approach, the overall preparation, simulation and analysis time could be significantly shorten using VR. The presented evaluations here, show, that using advanced and integrated 3D interactive pre-/postprocessing facilities based on VR, leads to a faster assessment of the results (down to minutes and seconds compared to inhouse COTS). The elaborated new techniques for conceptual simulations proved to allow an engineer being concentrated on local problems without a need to re-calculate the overall global problem. Further on, it provides a basis for the engineer to find an answer to his question: "where do I have to spend my analysis time?". The involved engineers evaluated the VR-based hybrid desktop as an easy to use and intuitive access technology. Finally, they estimate to be able to shorten down their engineering workflow by several days. As a consequence, we have to object the hypothesis of Choi et al., having shown, how engineers could benefit from new simulation methodologies integrated into advanced front ends such as VR. Yet, still major challenges remain for making VR a widely accepted front end in the CAE domain as the way to more advanced simulations require new mathematical models, methodologies to control the creation of elements and assembly of matrices and thus, is stony and tough, but worth a try!

REFERENCES

Georgii, J., "Real-Time Simulation and Visualisation of Deformable Objects", *PhD Thesis, Institut für Informatik*, Technische Universität München (TUM), 2008

Graf, H., Stork, A, "Enabling Real-Time Immersive Conceptual CAE Simulations based on Finite Element Masks", In: *Proc. of the the ASME World Conference of Innovative VR, (WINVR'11),* Milan, Italy, 2011

Graf, H., Stork, A., "Virtual Reality based Interactive Conceptual Simulations Combining Post-Processing and Linear Static Simulations", *In: Proceedings of the International Conference on Human-Computer Interaction (HCII) – Virtual Augmented and Mixed Reality (VAMR)*: Part I (2013), Las Vegas, NV, USA, LNCS, pp. 13-22, 2013

Graf, H., "A 'Change n'Play' Software Architecture integrating CAD, CAE and Immersive Real-Time Environments for Conceptual Simulations", In: *Proc. of the IEEE International Conference on CAD/Graphics'11*, Jinan, China, 2011

Larson, M., Graf, H., Stork, A., "Interactive 3D Subdomaining using Adaptive FEM based on Solutions to the Dual Problem", In: *Proceedings of the ACM Virtual Reality International Conference,* Laval, France, ACM New York, NY, USA, 2014

Choi, S.S., Jung, K., and Do Noh, Sang, "Virtual reality applications in manufacturing industries: Past research, present findings, and future directions", *Concurrent Engineering: Research and Applications*, Vol23 (1), 2015

Tierney, C., Nolan, D., Robinson, T., Armstrong, C.G., "Using mesh-geometry relationships to transfer analysis models between CAE tools". *In: Journal of Engineering with Computers, Springer London*, July 2015

Weber, D., "Interactive Physically Based Simulation Efficient Higher-Order Elements, Multigrid Approaches and Massively Parallel Data Structures", *PhD Thesis, Technical University of Darmstadt*, 2016

SIMULATION STUDY OF 1DOF HYBRID ADAPTIVE CONTROL APPLIED ON ISOTHERMAL CONTINUOUS STIRRED-TANK REACTOR

Jiri Vojtesek, Lubos Spacek and Petr Dostal
Faculty of Applied Informatics
Tomas Bata University in Zlin
Nam. TGM 5555, 760 01 Zlin, Czech Republic
E-mail: {vojtesek,lspacek}@fai.utb.cz

KEYWORDS
Simulation, Mathematical Model, Adaptive Control, Continuous Stirred-Tank Reactor, 1DOF, Polynomial Approach.

ABSTRACT

A Continuous Stirred-Tank Reactor is typical system with nonlinear behavior and lumped parameters. The mathematical model of this type of reactor is described by the set of nonlinear ordinary differential equations that are easily solvable with the use of numerical methods. The big advantage of the computer simulation is that once we have reliable mathematical model of the system we can do thousands of simulation experiments that are quicker, cheaper and safer then examination on the real system. The control approach used in this work is a hybrid adaptive control where an adaptation process is satisfied by the on-line recursive identification of the External Linear Model as a linear representation of the originally nonlinear system. The polynomial approach together with the Pole-placement method and spectral factorization satisfies basic control requirements such as a stability, a reference signal tracking and a disturbance attenuation. Moreover, these methods produce also relations for computing of controller's parameters. As a bonus, the controlled output could be affected by the choice of the root position in the Pole-placement method. The goal of this contribution is to show that proposed controller could be used for various outputs as this system provides five possible options.

INTRODUCTION

The mathematical modelling and the computer simulation is a great tool for control engineering that helps with the understanding of the system's behavior without exanimation on a real equipment or a real model of the system (Honc *et al.* 2014), (Ingham *et al.* 2000). The benefits of the computer simulation are clear – it is quick, safe and of course much cheaper method than real experiments which, especially in chemical industry, could consume a lot of chemicals without a clear result. Even more, a lot of chemical experiments produce an exothermic reaction and wrong settings of the controller could end with the dangerous explosion.

This paper presents the simulation study from the initial steady-state and dynamic analyses to the hybrid adaptive control of the system. The steady-state and dynamic analyses observes the nonlinear behavior of the system and help us with the choice of the optimal control strategy. The system under the consideration is an isothermal Continuous Stirred-Tank Reactor (CSTR) the mathematical model of which is described by the set of five nonlinear ordinary differential equations (ODE) as there are five state variables – concentrations (Russell and Denn 1972). There were used Simple iteration method for the solving of the steady-state of this system which is, in fact, the numerical solution of the set of nonlinear algebraic equations that are transformed from the set of ODE with the condition that the derivative with respect to the time are equal to the zero in the steady-state. The dynamic analysis then employs the Standard Runge-Kutta's method for numerical solution of the set of ODE. Both methods are simple but accurate enough. Moreover, they are easily programmable and Runge-Kutta's methods are even build-in functions in the mathematical software Matlab (Vojtesek 2014) which was used as a simulation program in this work.

The control method here is based on the idea of the adaptive control (Åström and Wittenmark, 1989) where parameters of the controller are restored during the control according to the actual needs and state of the controlled system. The core function of this adaptive approach is the recursive identification of the External Linear Model (ELM) as a linear representation of the nonlinear system (Bobal *et al.*, 2005). Parameters of the controller than depends on the identified ELM and they are computed with the use the Polynomial method, the Pole-placement method and the Spectral factorization. As a result, this approach produces not only the controller that satisfies basic control requirements but also easily programmable relations for computing of controller's parameters which helps with the implementation inside the industrial controllers.

We call this approach the "hybrid" adaptive control because the polynomial approach used here is defined in the continuous-time which is more accurate but problematic for the on-line identification. Because of this, the special type of the discrete-time identification was used. This method is called the Delta-models (Middleton and Goodwin 2004) that belongs to the class of discrete-time models but its parameters approaches to the continuous-time ones for sufficiently small sampling

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

period (Stericker and Sinha 1993) as there are both input and output variables related to the sampling period.
The control strategy was applied on the control of two different outputs which shows that it could be also successfully applicable to similar types of processes.

ADAPTIVE CONTROL

The control approach used in this work is an adaptive control. The philosophy of this control method comes from the nature, where plants, animals and even human beings "adopt" their behavior to the actual conditions and an environment. This could be done, from the control point of the view, for example by the change of the controller's parameters, structure etc. (Bobal *et al.*, 2005).

External Linear Model

The approach used here starts with the dynamic analysis of the system that help us with the understanding of the system's behavior. Resulted step responses are then used for the choice of the External Linear Model (ELM) as a linear representation of usually nonlinear system. This ELM could be in the form of the polynomial transfer function and the adaptivity is then satisfied by on-line recursive identification that estimates parameters of the ELM in every moment. This procedure guarantees that this ELM describes the system accurately to the relative state of the system.
The general form of the ELM's transfer function is

$$G(s)=\frac{b(s)}{a(s)} \tag{1}$$

where parameters of polynomials *a*(*s*) and *b*(*s*) are computed from the recursive identification and both polynomials holds the feasibility condition for $\deg a(s)\geq \deg b(s)$.

Design of Controller

Now we know, that the controlled nonlinear system is described by the polynomial transfer function (1) and we can describe the controller also by the transfer function

$$Q(s)=\frac{q(s)}{p(s)} \tag{2}$$

where *q*(*s*) and *p*(*s*) are again commensurable polynomials with the properness condition $\deg p(s)\geq \deg q(s)$.

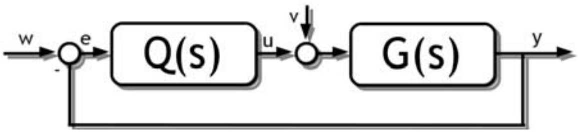


Figure 1: 1DOF control scheme

If we put this controller's transfer function in the feedback part of the closed-loop scheme displayed in Figure 1, the Laplace transform of the transfer function *G*(*s*) in (1) is then

$$G(s)=\frac{Y(s)}{U(s)}\Rightarrow Y(s)=G(s)\cdot U(s) \tag{3}$$

where Laplace transform of the input signal *u* is from Figure 1

$$U(s)=Q(s)\cdot E(s)+V(s)=Q(s)\cdot\left[W(s)-Y(s)\right]+V(s) \tag{4}$$

Then, if we put polynomials *a*(*s*), *b*(*s*), *p*(*s*) and *q*(*s*) from (1) and (2), the equation (3) has form

$$\begin{aligned} Y(s)&=\frac{b(s)q(s)}{a(s)p(s)+b(s)q(s)}\cdot W(s)+\ldots\\ &\ldots+\frac{a(s)p(s)}{a(s)p(s)+b(s)q(s)}\cdot V(s)\end{aligned} \tag{5}$$

The denominator here is a characteristic polynomial of the closed loop system we can write it generally

$$a(s)\cdot p(s)+b(s)\cdot q(s)=d(s) \tag{6}$$

where *d*(*s*) is a stable optional polynomial. The position of roots of this polynomial affects control results and the whole equation (6) is called Diophantine equation (Kucera 1993).
Every control system must fulfill basic control requirements such as a stability, an asymptotic tracking of the reference signal and the disturbance attenuation. The closed loop system is stable, if the polynomial *d*(*s*) on the left side of (6) is also stable. Asymptotic tracking of the reference signal and disturbance attenuation is gained if the polynomial *p*(*s*) includes the least common divisor *f*(*s*) of denominators of transfer functions of the reference signal *w* and disturbance signal *v*:

$$p(s)=f(s)\cdot\tilde{p}(s) \tag{7}$$

As both of these signals are expected as a step function, the least common divisor is *f*(*s*) = *s*.
The transfer function of the feedback controller is then

$$\tilde{Q}(s)=\frac{q(s)}{s\cdot\tilde{p}(s)} \tag{8}$$

and we can rewrite the Diophantine equation (6) to

$$a(s)\cdot s\cdot\tilde{p}(s)+b(s)\cdot q(s)=d(s) \tag{9}$$

Polynomials *a*(*s*) and *b*(*s*) are known from the recursive identification and polynomials $\tilde{p}(s)$ and *q*(*s*) are unknown parameters that are needed to be computed. The method used here for computation of those polynomials is the Method of uncertain coefficients.
The polynomial *d*(*s*) on the right side of the equation (9) is the stable optional polynomial. The simple method

used for the choice of the polynomial $d(s)$ is the Pole-placement method which divides this polynomial into one or more parts with double, triple, etc. roots, e.g.

$$d(s)=(s+\alpha)^m;d(s)=(s+\alpha_1)^{m/2}\cdot(s+\alpha_2)^{m/2},\ldots \quad (10)$$

where $\alpha > 0$. The disadvantage of this method can be found in the uncertainty. There is no general rule which can help us with the choice of roots which are, of course, different for different controlled processes. One way how we can overcome this unpleasant feature is to use spectral factorization. Big advantage of this method is that it can make stable roots from every polynomial, even if it is unstable. The polynomial $d(s)$ is in this case

$$d(s)=n(s)\cdot(s+\alpha)^{\deg d(s)-\deg n(s)} \quad (11)$$

where parameters of the polynomial $n(s)$ are computed from the spectral factorization of the polynomial $a(s)$ in the denominator of (1), i.e.

$$n^*(s)\cdot n(s)=a^*(s)\cdot a(s) \quad (12)$$

The use of the spectral factorization satisfies that the polynomial $n(s)$ is stable even if the identified polynomial $a(s)$ is unstable. This situation could occur in the adaptation part at the beginning of the control, where the estimator does not have enough information about the system. The second part is then regular Pole-placement method, but the number of unknown roots is reduced by the choice of the polynomial $n(s)$.

It is good to have some parameter that could affect the control results. This parameter is in this case the position of the root α in the Pole-placement method. The only condition which comes from the stability is that this root must be $\alpha > 0$.

Degrees of unknown polynomials $\tilde{p}(s)$, $q(s)$ and $d(s)$ are for the fulfilled properness condition generally:

$$\begin{aligned}&\deg \tilde{p}(s)\geq \deg a(s)-1\\&\deg q(s)=\deg a(s)\\&\deg d(s)=\deg a(s)+\deg \tilde{p}(s)+1\\&\deg n(s)=\deg a(s)\end{aligned} \quad (13)$$

Recursive Identification

The computation of the controller's parameters by the Method of uncertain coefficients needs parameters of the system, i.e. coefficients of polynomials $a(s)$ and $b(s)$ from the transfer function $G(s)$ in (1). It was mentioned before, that these coefficients are estimated recursively during the control. The Recursive Least-Squares (RLS) method (Rao and Unbehauen 2005) is ideal method for this task – it is easily programmable and together with forgetting factors accurate enough.

The RLS method uses two vectors: the first is data vector that comes from the measured input and output variables, generally in the discrete form

$$\begin{aligned}\boldsymbol{\varphi}(k-1)=[&-y(k-1),-y(k-2),\ldots,-y(k-n)\\&u(k-1),u(k-2),\ldots,u(k-m)]^T\end{aligned} \quad (14)$$

The second, unknown, vector is the vector of parameters

$$\boldsymbol{\theta}(k)=[a_n,\ldots,a_1,a_0,b_m,\ldots,b_1,b_0]^T \quad (15)$$

and this vector is in RLS computed from the set of equations:

$$\begin{aligned}&\varepsilon(k)=y(k)-\boldsymbol{\varphi}^T(k)\cdot\hat{\boldsymbol{\theta}}(k-1)\\&\gamma(k)=\left[1+\boldsymbol{\varphi}^T(k)\cdot\mathbf{P}(k-1)\cdot\boldsymbol{\varphi}(k)\right]^{-1}\\&\boldsymbol{L}(k)=\gamma(k)\cdot\mathbf{P}(k-1)\cdot\boldsymbol{\varphi}(k)\\&\mathbf{P}(k)=\frac{1}{\lambda_1(k-1)}\left[\mathbf{P}(k-1)-\frac{\mathbf{P}(k-1)\cdot\boldsymbol{\varphi}(k)\cdot\boldsymbol{\varphi}^T(k)\cdot\mathbf{P}(k-1)}{\dfrac{\lambda_1(k-1)}{\lambda_2(k-1)}+\boldsymbol{\varphi}^T(k)\cdot\mathbf{P}(k-1)\cdot\boldsymbol{\varphi}(k)}\right]\\&\hat{\boldsymbol{\theta}}(k)=\hat{\boldsymbol{\theta}}(k-1)+\boldsymbol{L}(k)\varepsilon(k)\end{aligned} \quad (16)$$

where $\boldsymbol{\varphi}$ is regression vector, ε denotes a prediction error, $\mathbf{P}$ is a covariance matrix and λ_1 and λ_2 are forgetting factors. There were defined different methods in (Fikar and Mikles 1999), for example constant exponential forgetting where $\lambda_2 = 1$ and λ_1 is computed from

$$\lambda_1(k)=1-K\cdot\gamma(k)\cdot\varepsilon^2(k) \quad (17)$$

where K is a very small value (e.g. $K = 0.001$).

SIMULATION MODEL

The proposed adaptive control was tested by the simulation on the mathematical model of the isothermal chemical reactor (Russell and Denn 1972), schematic representation of which is shown in Figure 2.

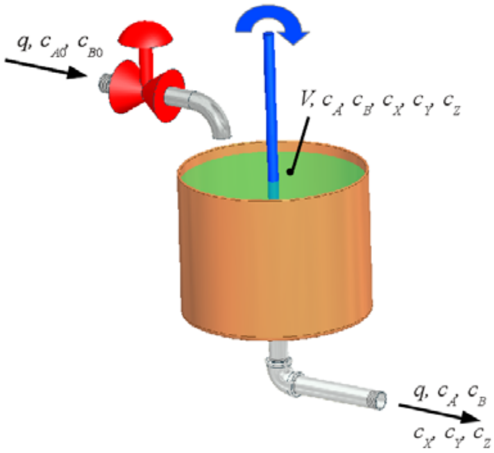


Figure 2: Isothermal Continuous Stirred-Tank Reactor

Reactions inside the reactor are

$$A+B\xrightarrow{k_1}X;A+X\xrightarrow{k_2}Y;A+Y\xrightarrow{k_3}Z \quad (18)$$

The mathematical model is constructed with the use of material balances, that are in the general word form

$$\left\{\begin{array}{c}\text{Mass flow of the}\\ \text{component into}\\ \text{the system}\end{array}\right\}=\left\{\begin{array}{c}\text{Mass flow of the}\\ \text{component out of}\\ \text{the system}\end{array}\right\}+\left\{\begin{array}{c}\text{Rate of}\\ \text{accumulation of}\\ \text{mass in the system}\end{array}\right\}$$

and as we have five state variables – concentrations c_A, c_B, c_X, c_Y and c_Z, the resulting mathematical model is described by five ordinary differential equations (Russell and Denn 1972)

$$\begin{aligned}\frac{dc_A}{dt}&=\frac{q}{V}\left(c_{A0}-c_A\right)-k_1\cdot c_A\cdot c_B\\ \frac{dc_B}{dt}&=\frac{q}{V}\left(c_{B0}-c_B\right)-k_1\cdot c_A\cdot c_B-\\ &\quad -k_2\cdot c_B\cdot c_X-k_3\cdot c_B\cdot c_Y\\ \frac{dc_X}{dt}&=\frac{q}{V}\left(c_{X0}-c_X\right)+k_1\cdot c_A\cdot c_B-k_2\cdot c_B\cdot c_X\\ \frac{dc_Y}{dt}&=\frac{q}{V}\left(c_{Y0}-c_Y\right)+k_2\cdot c_B\cdot c_X-k_3\cdot c_B\cdot c_Y\\ \frac{dc_Z}{dt}&=\frac{q}{V}\left(c_{Z0}-c_Z\right)+k_3\cdot c_B\cdot c_Y\end{aligned}\tag{19}$$

The mathematical model (19) includes besides state variables c_A, c_B, c_X, c_Y and c_Z also their initial values (with index $c_{\cdot 0}$), volumetric flow rate q, volume of the reactor V and rate constants of the reactions $k_1 - k_3$. Rate constants, the volume of the reactant and input concentrations are fixed parameters that are shown in Table 1(Russell and Denn 1972):

Table 1: Fixed parameters of the CSTR

$k_1 = 5\times10^{-4}\ m^3.kmol^{-1}.s^{-1}$	$c_{A0} = 0.4\ kmol.m^{-3}$
$k_2 = 5\times10^{-2}\ m^3.kmol^{-1}.s^{-1}$	$c_{B0} = 0.6\ kmol.m^{-3}$
$k_3 = 2\times10^{-2}\ m^3.kmol^{-1}.s^{-1}$	$c_{X0} = c_{Y0} = c_{Z0} = 0\ kmol.m^{-3}$
$V = 1\ m^3$	

The only quantity which could be used as an adjustable parameter is the volumetric flow rate of the reactant q which was later used as an action value for control.

The Steady-state and Dynamic Analyses

The goals of the steady-state and dynamic analyses are usually to observe the behavior of the system and its physical boundaries.

The steady-state analysis tries to find values of the state variables in the steady-state, i.e. for the time $t \rightarrow \infty$. It means that the set of nonlinear ordinary differential equations (ODE) (19) is transformed into the set of nonlinear algebraic equations that can be solved for example with the Simple iteration method. We have done various computations for different input volumetric flow rates q and the results are shown for example in (Zelinka *et al.* 2006). Our experiments have shown nonlinear behavior and the optimal working point is $q^s = 1\times10^{-4}\ m^3.s^{-1}$.

Steady-state values of the state variables for this working point are:

$$\begin{aligned}c_A^s&=0.2407\ kmol.m^{-3} & c_B^s&=0.1324\ kmol.m^{-3}\\ c_X^s&=0.0024\ kmol.m^{-3} & c_Y^s&=0.0057\ kmol.m^{-3}\\ c_Z^s&=0.1513\ kmol.m^{-3}\end{aligned}\tag{20}$$

It was already mentioned, that there are five theoretical state variables but we have chosen only two of them to be controlled – concentrations $c_B^s(t)$ and $c_Z^s(t)$.

As the steady-state values of quantities are also initial values for the dynamic analysis that examine the behavior after the step change of the input variable, the output variable then starts on its steady-state value. In this work, the initial value of the output variable is subtracted from the actual value which results in step responses that starts from zero. The both outputs in this work are then

$$y_1(t)=c_B(t)-c_B^s;\ y_2(t)=c_Z(t)-c_Z^s\quad\left[kmol.m^{-3}\right]\tag{21}$$

where c_B^s and c_Z^s are those values in (20).

There were done eight step changes of the input variable

$$u(t)=\frac{q(t)-q^s}{q^s}\cdot100\,[\%]\tag{22}$$

and results are shown in Figure 3 and Figure 4.

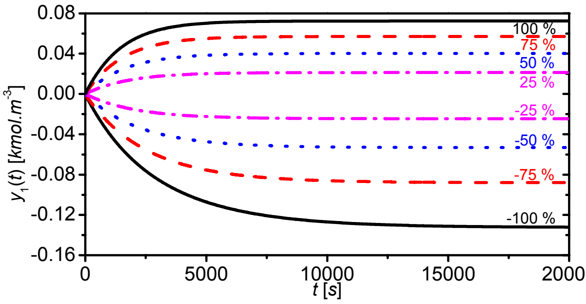


Figure 3: The course of the output variable $y_1(t)$ after step changes of the input variable $u(t)$

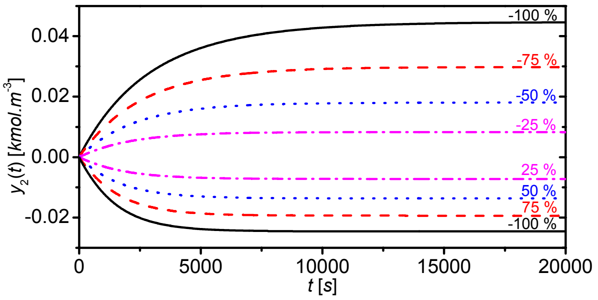


Figure 4: The course of the output variable $y_2(t)$ after step changes of the input variable $u(t)$

Step responses in Figure 3 and Figure 4 have shown nonlinear behavior of the system and also limitations of output variables. The first output differs in the boundaries from -0.1322 (-100 %) to 0.0725 (+100 %) $kmol.m^{-3}$. The second output $y_2(t)$ has boundaries from -0.0245 (+100 %) to 0.0446 (-100 %).

Results of the dynamic analysis can help us with the choice of the ELM's transfer function (1). According to courses of the output the transfer function has a form:

$$G(s)=\frac{Y(s)}{U(s)}=\frac{b(s)}{a(s)}=\frac{b_1 s+b_0}{s^2+a_1 s+a_0} \qquad (23)$$

SIMULATION OF ADAPTIVE CONTROL

The ELM (23) is in the continuous-time *s*-plain and on-line identification of those models is more accurate but also problematic that those in discrete-time, where we can read input and output variables are read only in the defined time intervals and the time before this interval can be used for the recomputation of the systems or controller parameters.

One compromise could be use of the so called *δ*-models which are special types of discrete-time identification models, where input and output variables are related to the sampling period T_v. A new complex variable *γ* is computed from (Mukhopadhyay et al. 1992)

$$\gamma=\frac{z-1}{\alpha\cdot T_v\cdot z+(1-\alpha)\cdot T_v} \qquad (24)$$

where *z* is complex variable, T_v denotes a sampling period and *α* is an optional parameter. It is clear, that we can obtain infinitely many models for optional parameter *α* from the interval $0 \le \alpha \le 1$ and a sampling period T_v, however a *forward δ-model* was used in this work which has *γ* operator computed via

$$\alpha=0\Rightarrow\gamma=\frac{z-1}{T_v} \qquad (25)$$

The general form of the ELM (23) is then rewritten to the general differential equation

$$a'(\delta)y(t')=b'(\delta)u(t') \qquad (26)$$

where *t'* denotes discrete time and *δ* is the operator defined according to (25). Some previous experiments (Stericker and Sinha 1993) have shown, that parameters of polynomials *a'*(*δ*) and *b'*(*δ*) approach the parameters of the continuous-time model with decreasing value of the sampling period T_v.

The relation for the actual output is derived from the (26) as

$$\begin{aligned}y_\delta(k)=&-a_1 y_\delta(k-1)-a_0 y_\delta(k-2)+\\&+b_1 u_\delta(k-1)+b_0 u_\delta(k-2)\end{aligned} \qquad (27)$$

where $y_\delta$is the recomputed output to the *δ*-model:

$$\begin{aligned}&y_\delta(k)=\frac{y(k)-2y(k-1)+y(k-2)}{T_v^2}\\&y_\delta(k-1)=\frac{y(k-1)-y(k-2)}{T_v};\ y_\delta(k-2)=y(k-2)\\&u_\delta(k-1)=\frac{u(k-1)-u(k-2)}{T_v};\ u_\delta(k-2)=u(k-2)\end{aligned} \qquad (28)$$

and the data vector is then

$$\begin{aligned}\boldsymbol{\phi}(k-1)=&\left[-y_\delta(k-1),-y_\delta(k-2),\ldots\right.\\&\left.\ldots,u_\delta(k-1),u_\delta(k-2)\right]^T\end{aligned} \qquad (29)$$

and the vector of estimated parameters

$$\hat{\boldsymbol{\theta}}(k)=\left[a_1^\delta,a_0^\delta,b_1^\delta,b_0^\delta\right]^T \qquad (30)$$

can be computed from the ARX (Auto-Regressive eXtrogenous) model

$$y_\delta(k)=\boldsymbol{\theta}_\delta^T(k)\cdot\boldsymbol{\varphi}_\delta(k-1) \qquad (31)$$

by the recursive least squares methods described in the theoretical part.

As the ELM in Equation (23) is of the second order with relative order one, degrees of polynomials $\tilde{p}(s)$, *q*(*s*) and *d*(*s*) in (13) are

$$\begin{aligned}&\deg\tilde{p}(s)\ge\deg a(s)-1=2-1=1\\&\deg q(s)=\deg a(s)=2\\&\deg d(s)=\deg a(s)+\deg\tilde{p}(s)+1=2+1+1=4\\&\deg n(s)=\deg a(s)=2\end{aligned} \qquad (32)$$

and the transfer function of the controller (8) is

$$\tilde{Q}(s)=\frac{q(s)}{s\cdot\tilde{p}(s)}=\frac{q_2 s^2+q_1 s+q_0}{s\cdot(p_1 s+p_0)} \qquad (33)$$

The stable polynomial *d*(*s*) on the right side of the Diophantine equation (9) is

$$d(s)=n(s)\cdot(s+\alpha)^2=(s^2+n_1 s+n_0)\cdot(s+\alpha)^2 \qquad (34)$$

where parameters of the polynomial *n*(*s*) are computed from the spectral factorization (12) as

$$n_0=\sqrt{a_0^2},\ n_1=\sqrt{a_1^2+2n_0-2a_0} \qquad (35)$$

and we have one tuning parameter – the position of the double root *α* and parameters of polynomials $\tilde{p}(s)$, *q*(*s*) in the transfer function of the controller (33) are computed from the Diophantine equation (9) by the Method of uncertain coefficients.

Simulation Results

There were done different simulation experiments for *u*(*t*) as a change of the volumetric flow rate *q* (22) and changes of the output concentrations c_B and c_Z respectively – see (21).

The simulation time was 30 000 *s* and five step changes of the reference signal were done during this time. The sampling period was T_v = 10 *s*, the initial covariance matrix **P**(0) has on the diagonal $1\cdot10^6$ and starting vectors of parameters for the identification was chosen according to some previous measurements.

It is good to qualify the control results also by some quantitative criterion. In this case, we used control quality criteria S_u and S_y that reflects the changes of the input variable u and the control error $e = w - y$. These criteria are then computed in the whole control from

$$S_u = \sum_{i=2}^{N} \left(u(i) - u(i-1)\right)^2 \qquad , \text{ for } N = \frac{T_f}{T_v} \qquad (36)$$
$$S_y = \sum_{i=1}^{N} \left(w(i) - y(i)\right)^2$$

where T_f is the final time – in this case $T_f = 30\ 000\ s$.

The first simulation analysis observes the effect of the tuning parameter α = 0.002, 0.004 and 0.2 on the control response of the output $y_1(t)$.

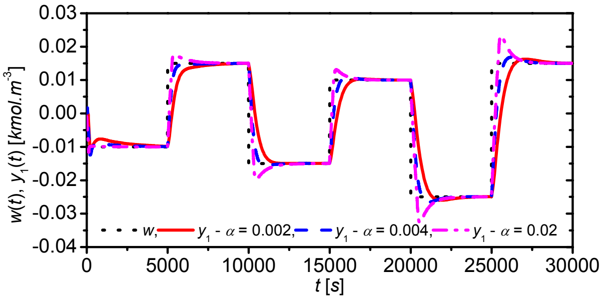


Figure 5: The course of the output variable $y_1(t)$ and the reference signal $w(t)$ for various α

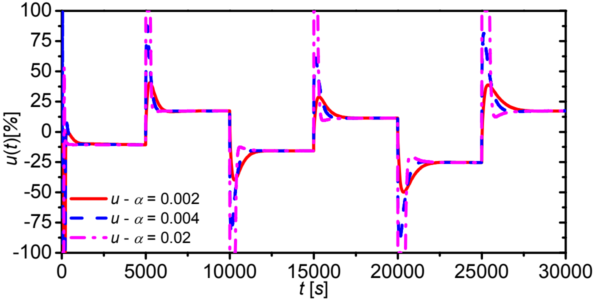


Figure 6: The course of the input variable $u(t)$ in the control of the output $y_1(t)$ for various α

Control results of the first simulation are shown in Figure 5 and Figure 6. It can be seen, that the proposed 1DOF hybrid adaptive controller does not have problem with the control of this output concentration cB(t). The course of the controlled output can be affected by the choice of the parameter α as a position of the root and it is clear, that bigger value of this parameter results in quicker output response but overshoot of the output variable. The question is if we want to have quicker response or overshoots are unwanted feature of the control system? The choice of the parameter α then depends on the answer for the previous question.

Values of criteria S_u and S_y for the first control simulation study are shown in Table 2. Lower value of α results in smoother course of the input variable which could be also important feature of the controller – quicker changes of $u(t)$ could affect the cost of the control and moreover it could harm the hardware of the controller that is also important. This smoother course reflects in the value of S_u which is minimal for the lowest value of α – see Table 2.

Table 2: Control quality criteria for the first simulation study – control of the output y_1

	S_u [-]	S_y [$kmol^2.m^{-6}$]
$\alpha = 0.002$	11 340	0.1554
$\alpha = 0.004$	68 736	0.0777
$\alpha = 0.02$	86 403	0.0636

The same simulation of the control was done for the second output y_2. As the second output has different course and especially final values which is shown in the graph Figure 4 that display results of the dynamic analysis, a little bit different reference was chosen in this case. Also simulations have shown better control results for the parameter α = 0.005, 0.01 and 0.075. Other parameters were equal to those in the previous study.

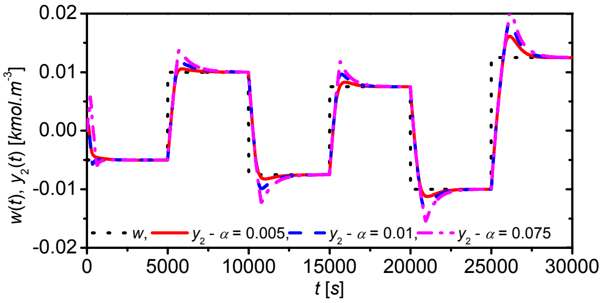


Figure 7: The course of the output variable $y_2(t)$ and the reference signal $w(t)$ for various α

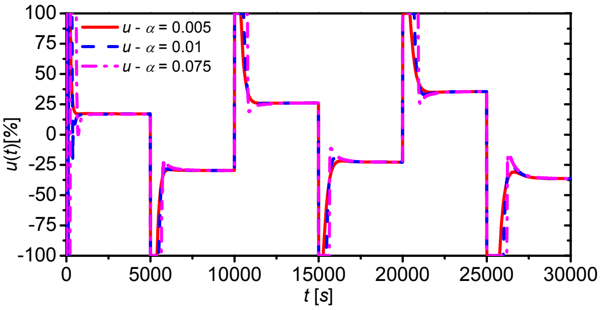


Figure 8: The course of the input variable $u(t)$ in the control of the output $y_2(t)$ for various α

Presented results in Figure 7 and Figure 8 show that the same controller can be used also for the control of the second output y_2 that represents the change of the output concentration c_Z in time t.

The effect of the tuning parameter α is the same as in previous case – an increasing value of this criterion produces quicker output response which is more evident for the bigger changes of the reference signal $w(t)$. This claim is also supported by values of the criteria $\underline{S_u}$ and $\underline{S_y}$ in Table 3 – the best tracking, i.e. the lowest value of $\underline{S_y}$, is for $\alpha = 0.005$.

Table 3: Control quality criteria for the second simulation study – control of the output y_2

	S_u [-]	S_y [$kmol^2.m^{-6}$]
$\alpha = 0.005$	255 704	0.0306
$\alpha = 0.01$	370 423	0.0328
$\alpha = 0.075$	193 016	0.0386

CONCLUSIONS

The paper shows the main benefits of the computer simulation – once we have a reliable mathematical model of the controlled system, we can do various simulations which can help us with the understanding of the behavior of the system (the steady-state and dynamic analyses) or with the choice and setting of the possible controller. The controller in this work was chosen as a hybrid adaptive controller where is the adaptivity satisfied by the recursive identification of the External Linear Model (ELM) as a linear representation of originally nonlinear system. This simplification is being supported by the on-line estimation of the ELM's parameters. Moreover, the controller could be tuned by the choice of the parameter α and it was proofed that increasing value of parameter results in quicker output response but bigger overshoots that are usually inappropriate.

Proposed results of simulations have shown that this controller can be used for this type of systems and does not matter which output you control – both controlled outputs representing changes of output concentrations of compounds B and Z indicates good control results. The only thing that differs is the choice of the reference signal which depends of the physical properties of the controlled output and the different value of α. We are then back in the main advantage of the computer simulation – while we have mathematical model and the relations that computes the parameters of the controller both in the form of simulation programs, we can do thousands of simulations that can help us with the choice of the optimal setting.

REFERENCES

Åström, K.J.; Wittenmark, B. 1989. *Adaptive Control*. Addison Wesley. Reading. MA, 1989, ISBN 0-201-09720-6.

Bobal, V.; Böhm, J.; Fessl, J.; Machacek, J. 2005 *Digital Self-tuning Controllers: Algorithms. Implementation and Applications*. Advanced Textbooks in Control and Signal Processing. Springer-Verlag London Limited. 2005, ISBN 1-85233-980-2.

Fikar, M.; J. Mikles 1999. *System Identification*. STU Bratislava

Honc, D.; Dusek, F.; Sharma, R. 2014 "GUNT RT 010 Experimental Unit Modelling and Predictive Control Application". In *Nostradamus 2014: Prediction, Modeling and Analysis of Complex Systems*. New York : Springer, 2014, s. 175-184. ISBN 978-3-319-07400-9.

Ingham, J.; Dunn, I. J.; Heinzle, E.; Prenosil, J. E. 2000 *Chemical Engineering Dynamics. An Introduction to Modeling and Computer Simulation*. Second. Completely Revised Edition. VCH Verlagsgesellshaft. Weinheim, 2000. ISBN 3-527-29776-6

Kucera, V. 1993. "Diophantine equations in control – A survey". *Automatica*. 29, 1993, p. 1361-1375.

Middleton, H.; Goodwin, G. C. 2004. *Digital Control and Estimation - A Unified Approach*. Prentice Hall. Englewood Cliffs, 2004, ISBN 0-13-211798-3

Mukhopadhyay, S.; Patra, A. G.; Rao, G. P. 1992 "New class of discrete-time models for continuos-time systems". *International Journal of Control*, vol.55, 1992, 1161-1187

Rao, G. P.; Unbehauen, H. 2005 "Identification of continuous-time systems". *IEEE Process-Control Theory Application*, 152, 2005, p.185-220.

Russell, T.; Denn, M. M. 1972 "Introduction to chemical engineering analysis". New York: Wiley, 1972, xviii, 502 p. ISBN 04-717-4545-6.

Stericker, D. L.; Sinha, N. K. 1993 "Identification of continuous-time systems from samples of input-output data using the δ-operator". *Control-Theory and Advanced Technology*. vol. 9, 1993, 113-125.

Vojtesek, J. 2014 "Numerical Solution of Ordinary Differential Equations Using Mathematical Software". In A*dvances in Intelligent Systems and Computing*. Heidelberg: Springer-Verlag Berlin, p. 213-226. ISSN 2194-5357. ISBN 978-3-319-06739-1.

Zelinka, I.; Vojtesek, J.; Oplatkova, Z. 2006. "Simulation Study of the CSTR Reactor for Control Purposes". In: *Proc. of 20th European Conference on Modelling and Simulation ESCM 2006*. Bonn, Germany, p. 479-482

AUTHOR BIOGRAPHIES

JIRI VOJTESEK was born in Zlin, Czech. He studied at Tomas Bata University in Zlin, Czech Republic, where he received his M.Sc. degree in Automation and control in 2002. In 2007 he obtained Ph.D. degree in Technical cybernetics at Tomas Bata University in Zlin. In the year 2015 he became associate professor. His research interests are modeling and simulation of continuous-time chemical processes, polynomial methods, optimal, adaptive and nonlinear control. You can contact him on e-mail address `vojtesek@fai.utb.cz`.

LUBOS SPACEK studied at the Tomas Bata University in Zlín, Czech Republic, where he obtained his master's degree in Automatic Control and Informatics in 2016. He currently attends PhD study at the Department of Process Control. His e-mail address is `lspacek@fai.utb.cz`.

PETR DOSTAL studied at the Technical University of Pardubice. He obtained his PhD. degree in Technical Cybernetics in 1979 and he became professor in Process Control in 2000. His research interest are modelling and simulation of continuous-time chemical processes. polynomial methods. optimal. adaptive and robust control. Unfortunatelly, prof. Dostal has died in January 2017.

TEACHING PROCESS MODELLING AND SIMULATION AT TOMAS BATA UNIVERSITY IN ZLIN USING MATLAB AND SIMULINK

Frantisek Gazdos
Faculty of Applied Informatics
Tomas Bata University in Zlin
Nam. T. G. Masaryka 5555, 760 01 Zlin, Czech Republic
E-mail: gazdos@fai.utb.cz

KEYWORDS
Modelling, Simulation, Education, MATLAB, Simulink.

ABSTRACT

This paper summarizes author's experiences of teaching a course on process modelling and simulation at Faculty of Applied Informatics, Tomas Bata University in Zlin, Czech Republic. It briefly presents contents of the course in both lectures and tutorials together with adopted methodology and used software tools. Requirements for the students to pass the course are also given as well as some statistics concerning their results. At the end of the contribution one of the final students' projects is also briefly presented.

INTRODUCTION

Modelling and simulation plays an important role in the process of education nowadays, e.g. (Kincaid et al. 2003; Stoffa 2004; Lean et al. 2006; Andaloro et al. 2007; Zavalani and Kacani 2012). It saves time, money and even prevents from injuries that could happen e.g. during some hazardous real-time experiments, e.g. (Jenvald and Morin 2004; Skarka et al. 2013). Thanks to the rapid developments in the field of computer hardware and software it is now possible in a safe place of simulation labs, offices or even at home perform experiments that could not be realized in the past decades without a proper hardware models. This places high demands on the process of education for experts in this field, in order to produce reliable (simulation) models and reasonable results (Kincaid et al. 2003).

The skills related to process modelling and simulation are useful in most engineering disciplines and applications, including also control engineering, e.g. (Ljung and Torkel 1994; Thomas 1999; Severance 2001; Egeland and Gravdahl 2002, Bequette 2003). Most of the control methods is based on some knowledge of a process model, therefore a control engineer must be able to obtain a proper model of the process to be controlled. In addition, it is advisable to test the designed control system properly using simulation means before real-time implementation in order to prevent from possible problems.

This contribution summarizes experiences related to teaching process modelling and simulation at Faculty of Applied Informatics, Tomas Bata University in Zlin, CZ, during studies of Master's degree programme "Automatic Control & Informatics" (FAI TBU in Zlin 2017). Here, in the first year of follow-up Master's studies during the winter semester students have to complete the course "Analysis and Simulation of Technological Processes" which is focused on the deepening of the knowledge in the field of modelling, computer simulation and analysis of common technological processes. All the things taught here are oriented so that they can be subsequently used easily for further control system design.

The paper is structured as follows: after this introductory part the contribution presents detailed structure of the presented course, including contents of both – lectures and tutorials (labs). Further, methodology of teaching and used software tools are discussed, next part introduces requirements for the students to pass the course and presents also some statistics for recent 10 years. Final section enables to see briefly the results of one of the simpler students' final projects. Some concluding remarks give insight into possible future directions of the course.

STRUCTURE OF THE COURSE

This part starts with some information on prerequisites of the students starting the course "Analysis and Simulation of Technological Processes" and follows by detailed description of contents of both – lectures and tutorials (labs). The course has 2 hours of lectures and 2 hours of tutorials (labs) per week and is donated by 5 credits after its successful completion. In our institution, there are 14 weeks of lectures per semester, followed by 5 weeks of examinations.

Students' Prerequisites

Students starting the course "Analysis and Simulation of Technological Processes" in the 1st year of their follow-up Master's studies (lasting 2 years) should already have some basic knowledge of university mathematics, physics, programming and computing software from their Bachelor's degree studies (lasting 3 years), e.g. they should complete the following courses that can be further useful for modelling and simulation in control engineering:

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

- Seminar of Mathematics; Mathematic Analysis; Differential Equations;
- Physical Seminary; Electricity, Magnetism and Wave Motion; Electrotechnics and Industrial Electronics; Microelectronics;
- Programming; Object-oriented Programming; Programs Theory; Algorithms and Data Structures; Matlab and Simulink; Programmable Logic Computers; Microcomputer Programming; JAVA Technology;
- Automation; Optimisation; System Theory.

These courses above are just a part of their Bachelor's studies and were selected as ones giving some basics that can be further exploitable in the field of process modelling and simulation for control engineers.

In their follow-up Master's studies, besides having the course "Analysis and Simulation of Technological Processes" students have to study simultaneously e.g. Mathematical Statistics, Process Engineering, Discrete Control System, Sensors, and others. Those students coming from different study programmes or different universities can also choose the course "Matlab and Simulink" besides some obligatory courses that equalize students' entry level.

Contents of Lectures

This course is primarily focused on modelling common continuous-time technological processes and their simulation/solution using the apparatus of numerical mathematics. The 14 weeks of 2-hours lectures/per week are divided into 2 main blocks – while in the first half students learn to derive analytically (simplified) first-principles mathematical models of common industrial processes, in the second block they study how to solve these models using the methods of numerical mathematics. After some introductory information where students gain motivation for studying this course, learn basic approaches to process modelling, become familiar with basic terminology and classification of models, these typical process models are derived step-by-step using the first-principles analytical modelling:

- liquid tanks with constant and non-constant cross-sections;
- processes with heat transfer, mixed and tubular heat exchangers;
- processes with mass transfer, distillation and staged processes;
- processes with chemical reactions, batch, semi-batch and continuous stirred tank reactors.

In the presented list of processes there are both linear and non-linear representative models as well as lumped and distributed parameters systems. If they are non-linear, a subsequent linearization and transformation into deviation models is also given. From the derived dynamical models, also their steady-states models are obtained and analysed, all with respect for subsequent control system design.

The second part of the lectures is focused on numerical solution of such models as obtained in the first part of the course. It begins with introduction into general approximation of functions, followed by polynomial approximations and then common numerical methods of solving introduced models are presented, from the simplest problems to more complex ones.

First, simulation/solution of steady-state behaviour of lumped-parameters processes is studied, resulting in the solution of sets of linear and nonlinear equations. For this purposes the principles of following common iteration methods are presented: simple iteration method, Jacobi, Gauss-Seidel and Relaxation methods, Newton method, and others, with obvious discussion on the conditions of convergence of all these algorithms.

Further, the problem of simulating/solving dynamic behaviour of lumped-parameters systems is explained, resulting in the solution of ordinary differential equations. Principles of both, simple one-step and more complex multi-steps methods are presented, including e.g. the simple Euler method, popular Runge-Kutta methods, and others, with further analysis on their numerical stability.

Finally at the end of the course, also the most complex problem in this field – simulation of steady-state and dynamics of distributed parameters systems is briefly presented, resulting in numerical solution of partial differential equations. Here, boundary value problems are discussed, together with the practical usage of the finite difference methods.

Contents of Tutorials

Tutorials (or laboratory exercises/practices/labs) are oriented more practically while following the course of more theoretically oriented lectures. In the first part of the semester students, with the help of a teacher, derive mathematical models of common industrial processes, followed by their practical simulation in a popular simulation software. They derive, e.g.:

- liquid-storage tanks with cylindrical, spherical and funnel-like shape, tanks in series;
- mixed and tubular heat exchangers;
- continuous (flow) stirred-tank reactors,

while learning the typical procedure of modelling and simulation:

- schematic picture,
- definition of variables (inputs, outputs, states),
- simplifying assumptions,
- energy/material balances,
- steady-states analysis,
- (classification of the model),
- choice/estimate/determination of model parameters,

- process variables limits, singular states, model validity,
- choice of initial/boundary conditions and operating point(s) for simulation,
- implementation of the model,
- simulation experiments,
- experiments evaluation,
- model verification / corrections…

For practical solution of the models the MATLAB computing software is fruitfully exploited together with its popular graphical multi-domain simulation library Simulink. In this environment student try to solve both steady-state and dynamical models using various approaches. For example, when solving models described by ordinary differential equations (ODEs) they learn how to solve it using the standard function *ode45* (based on an explicit Runge-Kutta (4,5) formula), how to build the model in the Simulink (including building their own blocks) or are advised to use e.g. the *state-space block* in the case of linear systems.

In the second part of the semester students are more practically familiarized with the numerical methods of solving the models. They start with recalling basics of solving sets of linear equations with the focus on the iterative methods and their practical implementation, i.e. programming in the MATLAB or other software. Then they go on to solve sets of nonlinear equations and finally (sets) of ODEs, with examples from the modelling part of the course or from practice. Discussion on the numerical aspects of the methods, i.e. convergence, accuracy, initial estimate, stability, etc., is a natural part of the explanations.

Completion of the Course

After successful completion of the course, students should be able to derive mathematical models of basic technological processes using the first-principles analytical modelling. Further, they should be able to analyse the models in order to obtain important information (e.g. linearity, stability, gain and time-constants…) and prepare them for subsequent control system design. Finally students should be able to solve/simulate and investigate these models using numerical methods, independent of the used simulation language.

While lectures attendance is voluntary, laboratory practices require min. 80% attendance and active students can gain "extra" points which can improve overall classification of the course. The classification is based on the unified credit system (compatible with the ECTS student mobility within European education programmes, e.g. European Union 2015) and therefore it is expressed on a common six-point scale: "A" (Excellent), "B" (Very good), "C" (Good), D (Satisfactory), E (Sufficient) and F (Fail/Unsatisfactory). The course is evaluated by 5 credits, where one credit represents 1/60 of the average annual student workload within the standard length of study. In order to obtain the credits students have to:

- have 80% attendance at tutorials/labs,
- have to elaborate and defence a "final project" on a given topic, obtaining min. 50% of points from it.

In the "final project" students show that they are able to derive simple mathematical models further usable for control system design and that they are able to analyse and solve/simulate these models effectively. So basically they try to follow the procedure they have learnt in the tutorials/labs. The final projects are assigned as soon as the students have basics knowledge and skills to elaborate it, typically after first 3 weeks. Students can come with their "own process", if not, they are assigned randomly from a regularly updated list of projects. The list of project includes, e.g.:

- cylindrical/spherical/funnel-like tanks in series,
- mixed and tubular heat exchangers,
- room heating process in various set-ups,
- continuous flow hot water systems and boilers,
- concentration and temperature mixers,
- swimming pool heating systems,
- continuous (flow) stirred-tank reactors,
- landfill site systems,
- various current/voltage controlled motors,
- conveyor systems
- and others…

while students follow the procedure they have learnt during the course (see "Contents of Tutorials" above), deriving the process model and analysing its steady-state and dynamic behaviour using the simulation means. One such typical final project is briefly presented at the end of this contribution.

STATISTICS OF THE RESULTS

This section summarizes briefly some statistical information concerning the number of students enrolling the course and their successfulness. The presented course is a part of "Automatic Control"– oriented study programme taught at our institution for several decades. Number of students enrolling studies in this field is not big – usually 1-2 study groups, as a result, the courses can be taught more individually and tailored to the actual needs of students and practice. This is also the case of the course "Analysis and Simulation of Technological Processes", referred in this contribution. General table with number of students enrolling this course in the last decade together with their successfulness according to the ECTS grading scale is presented in Table 1. From the table it can seen that overall number of students in the last decade was 126 and that the number of students in the last few years decreases, unfortunately, as also seen in Fig. 1. In the last several years, this is a trend in our country attributable to the drop in the population curve and also decreasing interest in technical studies, unfortunately.

Table 1: Number of Students and Their Successfulness in ECTS Grading Scale

Year	A	B	C	D	E	F	Sum
15/16	2	1	1	0	0	0	**4**
14/15	5	2	2	0	0	1	**10**
13/14	6	1	0	0	0	3	**10**
12/13	7	3	4	1	0	2	**17**
11/12	8	2	4	1	0	6	**21**
10/11	3	3	1	0	0	0	**7**
09/10	6	2	2	3	0	2	**15**
08/09	2	2	1	1	0	0	**6**
07/08	3	1	2	0	0	0	**6**
06/07	10	6	7	0	3	4	**30**
Sum	**52**	**23**	**24**	**6**	**3**	**18**	**126**
Sum [%]	**41%**	**18%**	**19%**	**5%**	**2%**	**14%**	

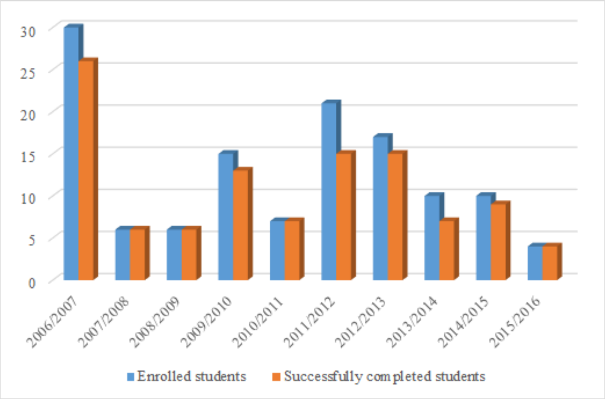


Figure 1: Number of Students and Their Successfulness

The table and graph show also successfulness of the students enrolling this course in each year, which is 86% on the whole, i.e. 86% of the students obtain the grade from “A” (Excellent) to “E” (Sufficient) according to the ECTS grading scale, and 14% of them do not complete the course successfully. Percentage in each category is presented in the table above or more clearly, in the next graph, Fig. 2.

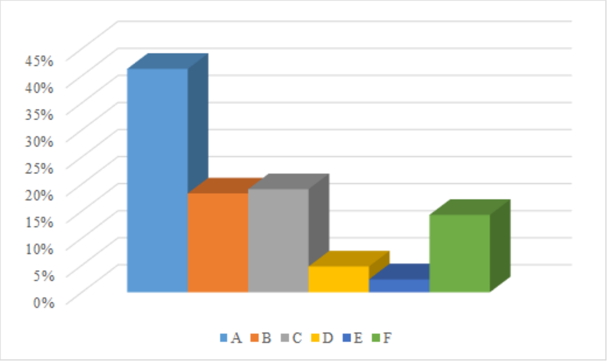


Figure 2: Students’ Successfulness in the ECTS Grading Scale

Generally speaking, unsuccessful students are usually those who enroll studies and this course and for some reasons decide to withdraw from their studies, after some time during the semester.

CASE STUDENT’S FINAL PROJECT

This section presents one of the simpler final students’ projects needed for the successful completion of the course. It starts with the problem assignment, followed by the elaboration including also main results and final summary.

Problem Formulation

Assume a room heated using an electric heater. Choose all the physical parameters so that they approx. correspond to real conditions.

- derive a simplified mathematical model of this system describing the room temperature $T(t)$ as a function of outdoor temperature $T_C(t)$ and heating power $P(t)$;
- derive and discuss also the steady-states model;
- determine the minimum necessary heating power to heat the room up to 20°C in case of outside temperature -10°C;
- display static characteristics $T^S = f(P^S, T_C^S)$;
- simulate a response of the room temperature to the step change in outside temperature and heating power ± 20%, compared to the chosen operating point; discuss the results;
- classify the derived model.

Simplified Mathematical Model

The modelled system can be sketched simply as presented in Fig. 3 below, where V stands for the volume and α is the average heat transfer coefficient.

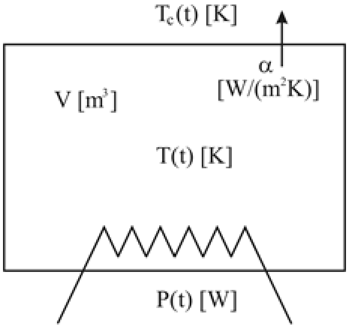


Figure 3: Schematic Picture of the Process

Definition of variables can be as follows: *input* variables are the heating power $P(t)$ in [W] and outdoor temperature $T_C(t)$ in [°C] (the latter one can be alternatively considered as a disturbance); *state* variable is the room temperature $T(t)$ in [°C], which is also the *output* variable, from the systems theory point of view, as displayed in Fig. 4.

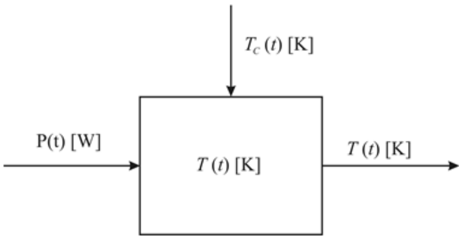


Figure 4: Process from the Systems Theory View

For the derivation of a mathematical model, the following common simplified assumptions are adopted:

- ideal air mixing,
- constant process parameters (air volume V, density ρ, heat capacity c_P, overall (average) heat transfer coefficient α, heat transfer surface area A, …),
- heat accumulation in the walls neglected.

Based on the heat balance:

heat input = heat output + heat accumulation,

the following simple mathematical model holds:

$$P(t) = \alpha A\left[T(t) - T_C(t)\right] + V\rho c_P \frac{dT(t)}{dt}, \qquad (1)$$

for some initial room temperature $T(0)$. For simulation purposes, the derivative is expressed as:

$$\frac{dT(t)}{dt} = \frac{1}{V\rho c_P}P(t) - \frac{\alpha A}{V\rho c_P}\left[T(t) - T_C(t)\right]. \qquad (2)$$

The steady-states model is obtained from (1) simply for the derivative equal to zero, i.e.

$$P^S = \alpha A\left(T^S - T_C^S\right), \qquad (3)$$

where the steady variables are denoted with s-superscript, as usual. Therefore, the steady room temperature reads simply as:

$$T^S = T_C^S + \frac{P^S}{\alpha A}, \qquad (4)$$

which is further used to generate the static characteristics.

Model parameters were chosen as follows: A = 55 m^2, V = 70 m^3, α = 1.82 W/m^{2}K, ρ = 1.205 kg/m^3, c_P = 1005 J/kgK. Initial conditions and operating point for simulation were defined as $T(0)$ = 20 °C, P = 2000 [W], T_C = 5 °C.

The dynamical and statical models (2), (4) have no singular states and are valid in common (reasonably chosen) conditions; the heating power can vary in the interval: $P(t) \in < 0; 4000 >$ W.

From the steady-states model (3) it is straightforward to compute the necessary heating power to heat the room to the temperature 20°C in case of outside temperature -10°C:

$$P^S = \alpha A\left(T^S - T_C^S\right) = 1.82 \times 55 \times 30 = 3003\ W, \qquad (5)$$

therefore, under the defined conditions, we need more than 3 kW to keep the temperature above 20°C when outside is freezing -10°C.

Simulation Results

The steady-states model (3) displayed graphically generates the static characteristics of Fig. 5, which shows linearity of the derived model.

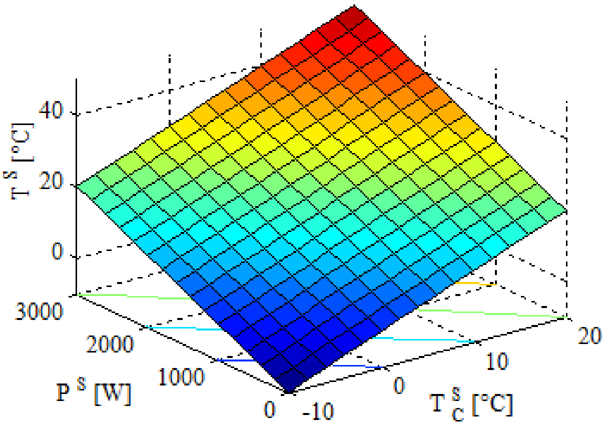


Figure 5: Static Characteristics

Dynamic step-responses obtained using the standard MATLAB ODE solver *ode45* are presented on the next graphs, Fig. 6 and Fig. 7.

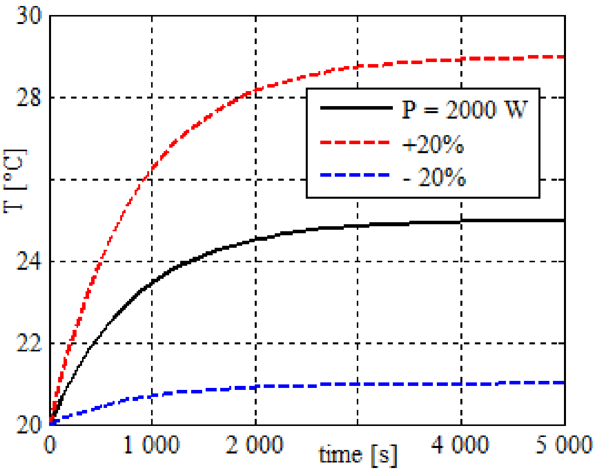


Figure 6: Step-response of Room Temperature with Heating Power

The first one shows the case of different heating power, starting with the nominal (P = 2000 W) and then small variations ± 20% from this value. As can be seen, when the outside temperature is around 5 °C, the electric heater enables to heat up the room to 25 °C approximately, when the power decreases to 1600 W, the room temperature will be around 21 °C and with 20% more (2400 W) the temperature settles around 29 °C, for the same initial conditions.

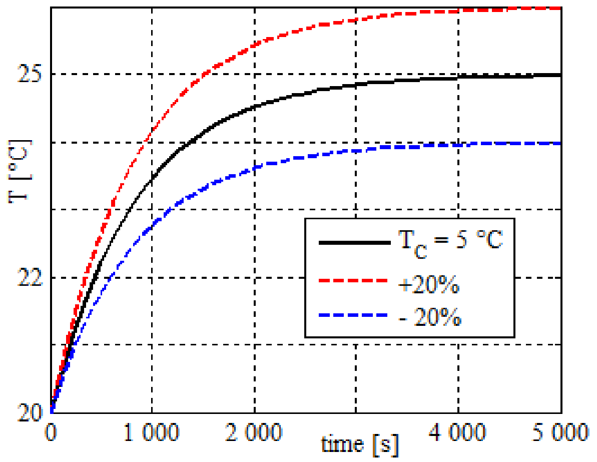


Figure 7: Step-response of Room Temperature with Outside Temperature

The second graph shows the case of different outside temperature and constant (nominal) power. As expected, lower outdoor temperature results in lower indoor temperature and vice versa.

Presented behaviour of the simplified mathematical model corresponds to the general expectation, therefore the model can be further used for e.g. subsequent control system design and analysis. More computations using the MATLAB software generated 3D plots of Fig. 8-9 with continuous intervals of heating power and outside temperature.

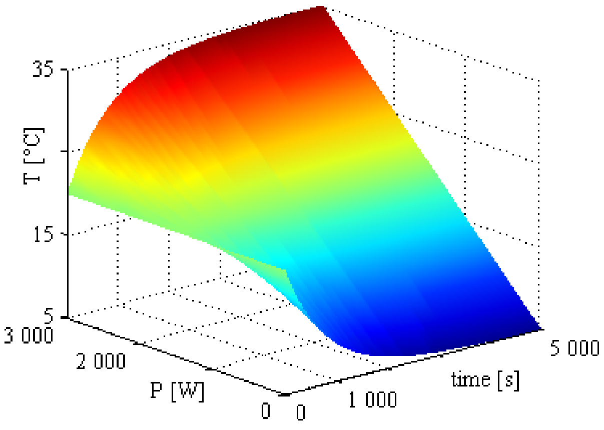


Figure 8: Step-responses of Room Temperature with Heating Power – 3D

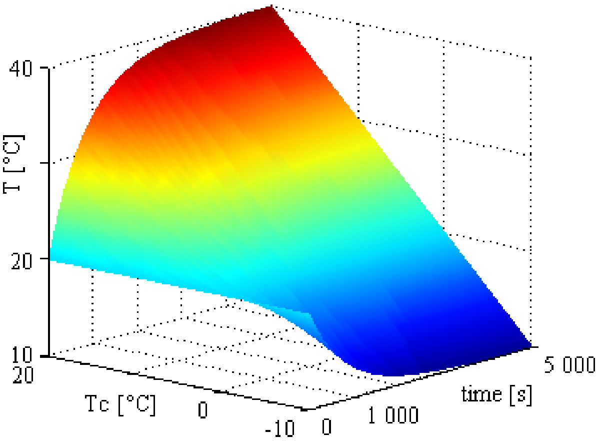


Figure 9: Step-responses of Room Temperature with Outside Temperature – 3D

Classification of the Model

Based on the adopted mathematical model and presented simulation results it is possible to classify it as:

- linear 1st order stable aperiodic system,
- with lumped parameters,
- continuous-time,
- deterministic,
- two-input and single-output,
- time-invariant,

which can further help to design a convenient control system for the adopted mathematical model, and real process as well.

Presented information and results outlined a possible form of students' final projects in the mentioned course focused on process modelling and simulation.

CONCLUSIONS

This paper has presented the structure and contents of the course "Analysis and Simulation of Technological Processes" taught in the first year of Master's degree study programme "Automatic Control & Informatics" at Faculty of Applied Informatics, Tomas Bata University in Zlin, Czech Republic. Requirements for the students to complete the course were also given together with some statistics concerning their successfulness. The presented case study has shown one of the simpler final students' projects for which the MATLAB computing system and its toolboxes for simulation and optimization are fruitfully utilized during the course. Future direction of the course aims to more practically-oriented modelling and simulation, connected to practical real-life examples and actual projects with industrial companies. There is also an obvious effort to teach the students not only to build reasonable process models but also to be able to utilize them for the next step - control system design. Therefore, in the next semester, after successful completion of the course "Analysis and

Simulation of Technological Processes" students use their built models in the next course – "State-space and Algebraic Control Theory" where they are taught how to design a convenient control systems. This course is also completed by a "final project" where students try to design and implement suitable control algorithms for their models/systems, again, with the strong help of MATLAB and Simulink.

REFERENCES

Andaloro, G.; V. Donzelli and R.M. Sperandeo-Mineo. 2007. "Modelling in physics teaching: the role of computer simulation." International Journal of Science Education, Vol.13, No.3, 243-254.

Bequette, B.W. 2003. *Process Control: Modeling, Design and Simulation*. Prentice Hall, New Jersey.

Egeland O. and J.T. Gravdahl. 2002. *Modeling and Simulation for Automatic Control*. Marine Cybernetics, Trondheim.

European Union. 2015. *ECTS User's Guide*. Publications Office of the European Union. Luxembourg.

Faculty of Applied Informatics, Tomas Bata University in Zlin. 2017. [online]. Available at: http://www.utb.cz/fai-en

Jenvald J. and M. Morin. 2004. "Simulation-Supported Live Training for Emergency Response in Hazardous Environments." *Simulation and Gaming*, Vol.35, No.3, 363-377.

Kincaid, J.P.; R. Hamilton, R.W. Tarr and H. Sangani. 2003. "Simulation in Education and Training". In *Applied System Simulation: Methodologies and Applications*, M.S. Obaidat and G.I. Papadimitriou (Eds.). Springer Science, New York, 437-456.

Lean, J.; J. Moizer; M. Towler and C. Abbey. 2006. "Simulation and games: Use and barriers in higher education." *Active learning in higher education*, Vol.7, No.3, 227-242.

Ljung L. and G. Torkel. 1994. *Modeling of Dynamic Systems*. Prentice Hall, New Jersey.

Severance, F.L. 2001. *System Modeling and Simulation: An Introduction*. Wiley, Chichester.

Skarka W; M. Otrebska and P. Zamorski. 2013. "Simulation of Dangerous Operation Incidents in Designing Advanced Driver Assistance Systems" *Proceedings of the Institute of Vehicles*, Vol.96, No.5, 131-139.

Stoffa, V. 2004. "Modelling and Simulation as a Recognizing Method in Education" *Educational Media International*, Vol.41, No.1, 51-58.

Thomas, P. 1999. *Simulation of Industrial Processes for Control Engineers*. Butterworth-Heinemann, Oxford.

Zavalani, O. and J. Kacani. 2012. "Mathematical Modelling and Simulation in Engineering Education". In *Proceedings of the 15th International Conference on Interactive Collaborative Learning (ICL)* (Villach, Austria, Sep. 26-28). IEEE, Picataway, N.J., 1-5.

AUTHOR BIOGRAPHIES

FRANTIŠEK GAZDOŠ was born in Zlín, Czech Republic in 1976, and graduated from the Brno University of Technology in 1999 with MSc. degree in Automation. He then followed studies of Technical Cybernetics at Tomas Bata University in Zlín, obtaining Ph.D. degree in 2004. He became Associate Professor for Machine and Process Control in 2012 and now works as the Head of the Department of Process Control, Faculty of Applied Informatics of Tomas Bata University in Zlín.
He is author or co-author of more than 80 journal contributions and conference papers giving lectures at foreign universities, such as University of Strathclyde Glasgow, Instituto Politécnico do Porto, Università di Cagliari and others. His research covers the area of process modelling, simulation and control. His e-mail address is: gazdos@fai.utb.cz.

BIOMETRIC IDENTIFICATION OF PERSONS

Milan Adámek, Petr Neumann, Dora Lapková, Martin Pospíšilík and Miroslav Matýsek
Tomas Bata University in Zlín
Faculty of Applied Informatics
Nad Stráněmi 4511, 760 05, Zlín, Czech Republic
E-mail: adamek@fai.utb.cz

KEYWORDS

Biometric systems, reliability, attack, fake fingerprint, dactyloscopy.

ABSTRACT

Algorithms - used for the identification and verification of individuals through fingerprint recognition technology have long been extensively used in Forensic Science and in the private sector. This work is concerned with the verification of the reliability of biometric systems that use fingerprints for their activities. Further, the eFinger programme is used to study similarities between men, women´s and family members´ fingerprints.

INTRODUCTION

Biometrics has been used since ancient times to recognise/distinguish people. People mutually recognised each other by voice, face or the way they walked. Some characteristics do not change during human life; while others, on the contrary continue to be shaped with increasing age [1].

Differentiating people by their fingerprints is one of the oldest Biometric recognition methods. From the earliest times, this method was used by a lot of civilisations that had some form of knowledge of papillary lines, which are included on human skin. The first provable evidence of the use of modern Biometrics however, dates back to somewhere around the mid-19th Century. This was when fingerprints began to be used in Criminology. William James Herschel was one of the first people to take advantage of Biometrics then. He used railway employees´ fingerprints to confirm their identity. Using fingerprints was the only possible way to prove the identity of individual workers, because the majority of them could neither read nor write - and therefore, one could not expect a signature from them. This fingerprint confirmed their identity when being paid their salary.

In 1865, Francis Galton came out with a “Study of the Inheritance of Physical Characteristics.” The study dealt with the issue that newly-born babies take over and inherit some characteristics/properties from their parents. These characteristics can include both physical characteristics, as well as some properties - such as, behaviour or conduct. In 1869, Galton became co-founder of the science called Eugenics, which is the Science of Hereditary Diseases and Defects in the Foetus. A year later, Galton became the founder of research into twins. In 1880, he came up with a branch of science called Anthropometry, which deals with the measurement of human body dimensions. In 1892, Galton published his work entitled "Fingerprints", which led to the introduction of fingerprinting into practice in 1900. In the same year, Galton advocated the use of fingerprinting for identification and verification purposes. He demonstrated the permanence - and uniqueness, of papillary lines on the fingers. After this, fingerprinting/Dactyloscopy was introduced into police work [1].

BIOMETRICS

In Biometrics, several terms exist that are (also) used in Security Technologies. These include identity, identification, authentication, authorisation, verification, and recognition/recognisance. The term Biometrics, is a combination of two words - the word “bio” = life, and “metric” = measurement. Overall then, Biometrics can be seen as a science that deals with the measurement and examination of “live” human characteristics [1] [2].

The notion of identity is derived from the word “idem” - the same. This term is used when - for instance, comparing an object, situation, concept, and such like. One can divide “Identity” into two types; namely, “electronic identity” and “physical identity”. One can have several “Electronic Identities” at the same time – e.g. an identity registered on a Web-site. Conversely, (with regard to) “Physical Identity”, we each have only one, which is unique. Two people, who should have/share the same physical identity, do not exist. It composed of physiological, anatomical and behavioural traits [2][3].

Identification represents the process of discovering and identifying the validity of individuals. To begin with, the person must register such that it passes-on one´s biometric data into the system, which is stored in a database. In the course of determining the identity of a person, a comparison of the information stored in the database (template) and the currently-scanned information (sample) is carried out. This comparison process for as long as it needs to find compliance with data in the database. The output is either - finding the identity … and authorisation to enter; or to refuse entry because there was no consensus in the data.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

BIOMETRIC IDENTIFICATION METHODS

For Biometric identification needs, the human body can be divided into several basic components – or fields. These include the head, the arm/hand(s), the leg/foot/feet, and others. For personal identification purposes, one can make use of the methods set out in the table below.

Table 1 Comparison of Biometric Methods [1][4][5]

Method	**Field of Use**		**Interfaces with Users**	**Characteristics**		**Accuracy**
	P-S	B-K		A-F	B	
Scanning Faces	+	-	The face is scanned from a distance up to 2m	+	-	• •
Irises	-	+	Looking into the camera from a distance of cca. 30 cm	+	-	• ••
Retina	-	+	The eye is focused on the centre of a sensor at a distance of about 2 cm	+	-	• ••
Outer Ear	+	+	Users sets their ear close to a sensor	+	-	• •
Voice and Speech	+	+	Users pronounce words or phrases into a sensor	-	+	•
Fingerprints	+	+	Fingers are pressed on the surface of a sensor	+	-	• ••
Palm-prints	+	+	Palms are pressed on the surface of a sensor	+	-	• ••
Scalloping of Nails	-	+	Fingers are inserted into a special sensor	+	-	• ••
Veins on the Back of the Hands	-	+	Hands are inserted into a sensor	+	-	• •
Veins in the Palm of a Hand	-	+	A palm is placed into a sensor	+	-	• •
Veins in the Fingers	-	+	Fingers are inserted into a sensor	+	-	• •
Signature Dynamics	+	+	The signature is made with a special pen on a special surface	-	+	•
Computer Key-stroke Dynamics	-	+	Users write a sample text on a special keyboard	-	+	• •

P –	**For Policing - Forensic Identification**
B – K	**For Safety - Commercial Identification**
A – F	**For Anatomical - Judicial Characteristics**
B –	**For Behavioral Characteristics**
P – S	**For Police – Court Identification**

The human body undergoes many changes, whereby some properties/characteristics are more – or less, dependent on stability in time. The two most consistent properties over time - include the (human) iris, and DNA – in which almost no change occurs. Conversely, the characteristics of the human voice change a lot throughout life - especially during puberty. The time constancies of these biometric characteristics are depicted in Figure 1; the degree of temporal stability is expressed in percentages [4], [7].

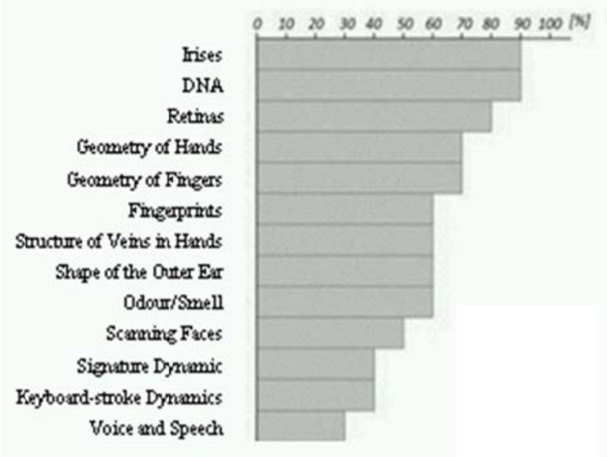


Figure 1: The Degree of Human Biometric Temporal Stability - expressed in percentages

BIOMETRIC SYSTEMS

The basic component of a Biometric Identification System (BIS), is a sensing module that ensures the scanning of biometric characteristics. The core part is the decision-making module, which compares the biometric features defined in the database. The output of the biometric identification system is the communication interface, or “lock” - allowing access to the space provided.

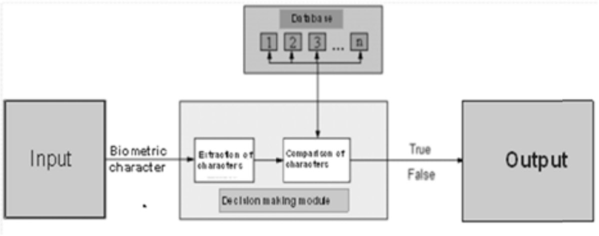


Figure 2: Structure of a Biometric Identification System, [5]

BIOMETRIC SYSTEMS´ RELIABILITY

One of the important characteristics of biometric systems includes the ability to clearly and faultlessly identify the identity of the rightful user - who is officially stored in a system database - and also, to differentiate/identify any unknown persons. Two parameters are used to express the degree of reliability of the system; these are:

- FRR – **False Rejection Rate** – (probability of erroneous rejection) – sometimes, the term also used for this is: Type I Error Rate
- FAR – **False Acceptance Rate** (probability of erroneous rejection) – sometimes, the term also used for this is: Type II Error Rate [4][6].

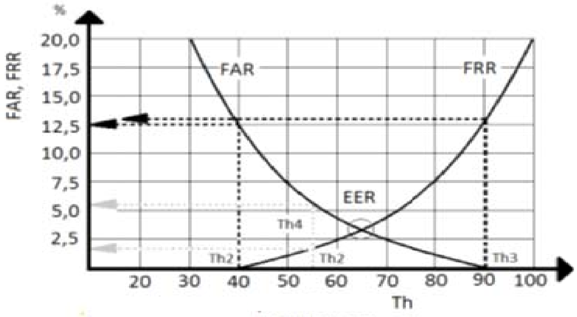


Figure 3: Dependence of FAR and FRR on Sensitivity Threshold, (Th) [4]

The False Acceptance Rate (FAR), and False Rejection Rate (FRR), and express the probability of the occurrence of a given error in percentages. From these errors, it follows that the higher the FRR - the lower the FAR; and vice versa. FRR and FAR are both dependent (Figure 3.), at the “Threshold Value”. The setting of the “Threshold Value” depends on the use of the system in practice; that is to say, if the bigger problem is if someone erroneously accepts or rejects it. When FAR and FRR the values are equal, this equality is referred to as the EER - Equal Error Rate. The EER allows one to determine the “approximate value of a security system.”

FINGERPRINT SENSOR PRINCIPLES

A. Optical Fingerprint Sensors

Optical fingerprint sensors are based on the reflection, or transmission of light. These sensors exploit the use of different reflections of light from papillary lines - and the space between these lines. The reflected light is then evaluated through a CCD or CMOS sensor.

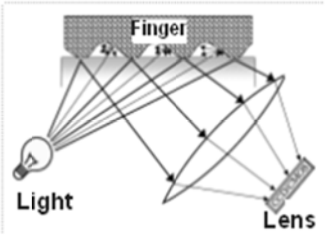


Figure 4: An Optical Sensor based on the principle of Reflections [3]

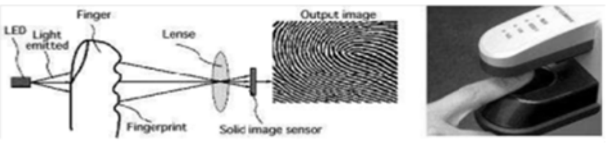


Figure 5: An Optical Sensor based on a Scanning Transmission, [3]

The optical sensor - using light transmission, is based on the backlighting of a finger from the upper side (from the nail), and on recording the sensor´s image on the opposite side.

B. Capacitive Fingerprint Sensors

The principle of this sensor is based on measuring the differences in capacity between the sensor-plate and the finger. The sensing area is equipped with a large number of sensor micro-electrodes in order to evaluate the capacity difference between the peaks and recesses in the papillary ridges in a finger.

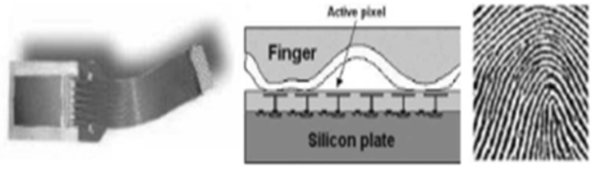


Figure 6: The Principle of a Capacitor Sensor [3]

C. Thermal Fingerprint Readers

Thermal fingerprint scanners use a small “pyro-detector” as a heat-sensitive element. The principle of this technology is based on measuring the temperature difference between peaks and valleys in finger papillary lines.

Figure 7: The Principle of a Thermal Sensor [4]

D. Ultrasonic Fingerprint Readers

This sensor transmits an ultrasonic signal from the transmitter to the fingerprint. The signal captures the reflected and deformed waves by rotating the transmitter or receiver. These are then evaluated farther and captured.

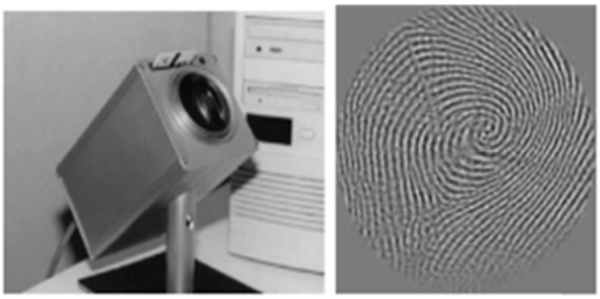


Figure 8: An Ultrasonic Reader + sample [5], [8]

FALSE FINGERPRINT-MAKING METHODS

Two approaches can be used for false fingerprint production:

1. Fake fingerprints can be created directly using appropriate materials. Several materials can be used to create a fake fingerprint - with regard to the preservation of papillary lines including their characteristics.

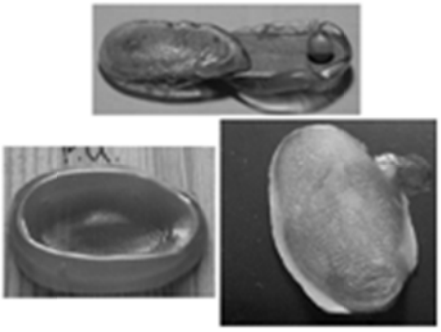


Figure 9: Plastic Finger-prints (Gelatine, Silicone, Plastic Moulds).

Granulated plastic can be used for the production of this material, which is malleable after being warmed up. Plastic materials have similar properties. Original fingerprints are pressed into plastic materials - thereby creating fake fingerprint templates. The fake fingerprint template is filled with materials like gelatine, silicone or plastic.

2. False Impressions Created by Secured Latent Traces

In order to produce a fingerprint, a “latent print” needs to be highlighted - and pictures taken – i.e. a scan; the resulting image is inverted and trimmed and then rendered in black and white shades. Enhanced image transfer of the material is designed to create a form that can be used for to “screen print” a plastic material - or create rubber stamps, etc.

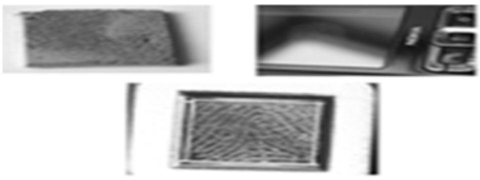


Figure 10: A Fake Fingerprint; Latent Fingerprint on a Mobile Phone.

Both procedures can create relatively high-quality fingerprints - but do not have a long shelf-life - they cannot be used with Vibrancy Control scanners. For example, gelatine or silicone cannot be applied to all touch sensors, because some methods do not meet the properties of materials that remain close to human skin properties, etc.

TESTING FINGERPRINT SENSOR RELIABILITY WITH THE USE OF FAKE FINGERPRINTS

The false fingerprints were measured against the immunity of fingerprint sensors. A Capacitive Fingerprint Sensor was used for testing the false fingerprints. The “fake fingerprints” were made from a plastic material rubber stamp.

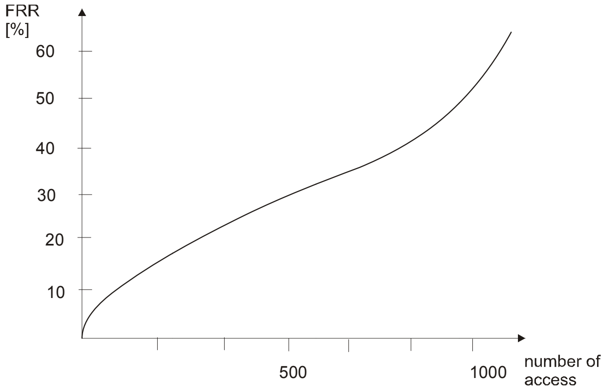


Figure 11: FRR of Capacitive Fingerprint Sensor: A fake fingerprint made from a rubber stamp

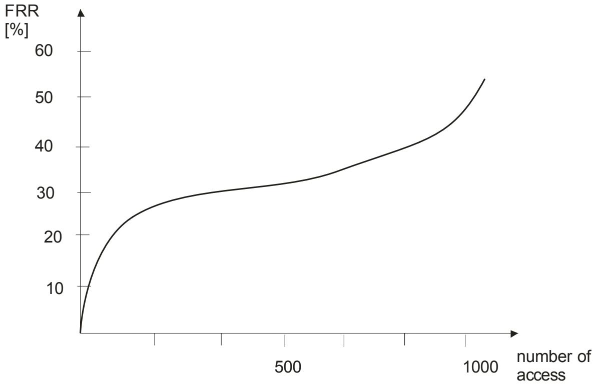
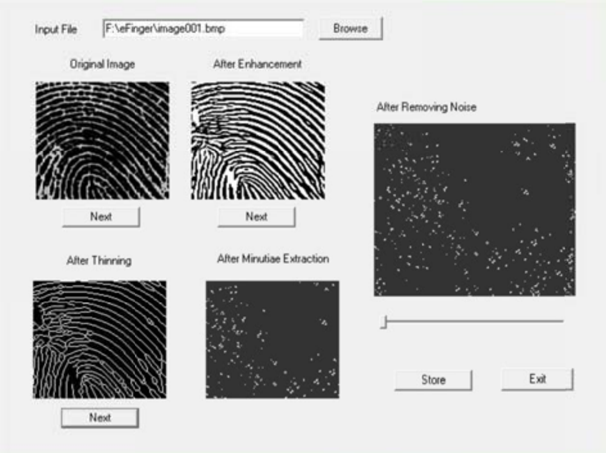


Figure 12: FRR of a Capacitive Fingerprint Sensor: The fake fingerprint is made from plastic

COMPARING THE MATCHING OF FINGERPRINTS

The eFinger programme was used for the comparison of fingerprints. The papillary ridge lines and points were extracted from a set of fingerprints stored in the database; the results are shown in Figure 13.



Input File F:\eFinger\image001.bmp

Figure 13: Extracted Fingerprint Papillary Ridge Lines and Points

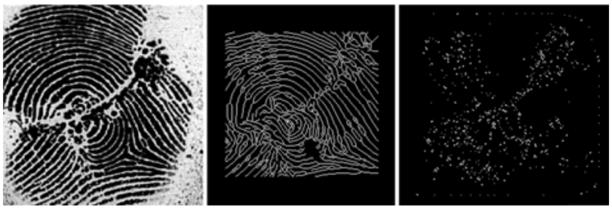
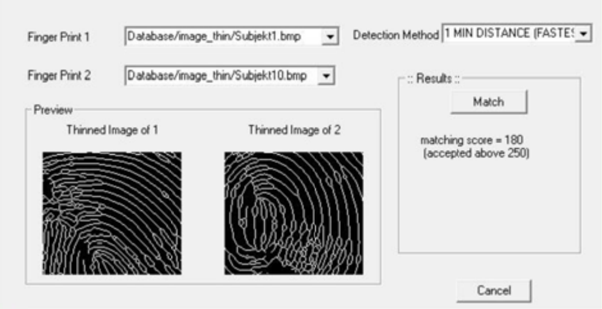


Figure 14: Demonstrations of the course of fingerprint extraction

Euclidean Metrics were used for the comparison of the match in identity of fingerprints – in which, the Euclidean Distance between Two Points, A and B is given by:

$$\rho\left(A,B\right)=\sqrt{\left(a_1-b_1\right)^2-\left(a_2-b_2\right)^2} \qquad (1)$$



Finger Print 1: Database/image_thin/Subjekt1.bmp

Detection Method: 1 MIN DISTANCE (FASTE!

Finger Print 2: Database/image_thin/Subjekt10.bmp

:: Results ::

matching score = 180
(accepted above 250)

Figure 15: Mutual Comparisons of Two Fingerprints

For the MIN DISTANCE methods used in the eFINGER programme, consensus is expressed numerically in intervals ranging from 0 to 1000. The maximum match is expressed by 1000; that is to say, there is 100% concordance of the two fingerprints. In this programme, values below 250 are considered unsatisfactory. Upon reaching the minimum number of matches - expressed by a number greater than 250, the comparison of the fingerprints reaches the minimum number of matching markers.

7 women´s, (Subjects: S1, S3, S4, S5, S6, S8 and S15); and 9 males´, (Subjects: S2, S7, S9, S10, S11, S12, S13, S14 and S16), fingerprints were matched and compared.

Table 2. Comparisons and Matches in Fingerprints in Women

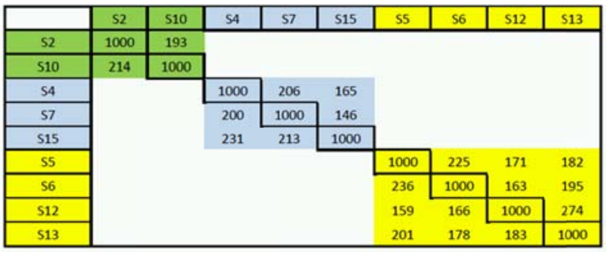
	S1	S3	S4	S5	S6	S8	S15
S1	1000	217	217	202	218	188	140
S3	241	1000	225	256	209	246	224
S4	250	249	1000	194	234	193	165
S5	247	247	250	1000	225	235	166
S6	226	233	238	236	1000	212	154
S8	229	220	213	192	235	1000	220
S15	234	208	231	207	200	193	1000

Table 3. Comparisons and Matches in Fingerprints in Men

	S2	S7	S9	S10	S11	S12	S13	S14	S16
S2	1000	205	218	193	182	181	196	210	171
S7	229	1000	214	181	187	195	161	193	189
S9	247	224	1000	192	175	172	203	233	163
S10	214	291	229	1000	190	190	186	211	174
S11	144	137	138	137	1000	233	244	128	127
S12	178	163	164	170	189	1000	274	172	182
S13	222	191	240	222	141	183	1000	251	157
S14	164	131	201	146	161	185	182	1000	153
S16	167	161	173	173	190	201	168	146	1000

From the tables above, it shows that, when comparing fingerprints, women show a higher degree of matching fingerprints; unlike the men´s fingerprints. Despite this, the rate of matches between individual subjects does not exceed the value of 300; that is to say, the fingerprint comparison of subjects matched the minimum number of markers T.

Table 4. Comparisons and Matches in Fingerprints between Family Members



	S2	S10	S4	S7	S15	S5	S6	S12	S13
S2	1000	193							
S10	214	1000							
S4			1000	206	165				
S7			200	1000	146				
S15			231	213	1000				
S5						1000	225	171	182
S6						236	1000	163	195
S12						159	166	1000	274
S13						201	178	183	1000

Furthermore, an assessment was made for matches and compliance between the fingerprints of family members. Subjects S2 and S10 are siblings - brothers. Subjects S4, S7 and S15 represent another group of family members - a brother, sister and cousin. The last, yellow coloured group in Table 4 is made up of Subjects S5, S6, S12 and S13. This quartet is composed of a brother, sister, mother - and their cousin). From the table above, when comparing the cardinal fingerprint elements of family members, there is seemingly no strong match between family members. Even in this case, the match compliance rate is less than 250, so – a sufficient number of fingerprint comparison markers cannot be matched, or made.

CONCLUSION

Biometric systems are closely-linked to reliability, which is given by the values: FAR and FRR. The aim of this paper was to suggest ways that can significantly impair the reliability of Biometric Systems. One such example (presented), is the production of false fingerprints – e.g., by using plastic and rubber models. Some types of fingerprint sensors are unable to recognise fingerprint copies, thus significantly impairing the reliability of Biometric Systems. Furthermore, the study also resolves the question of consensus (matching) – or respectively, the similarity of fingerprints for women and men, and between family members. The eFINGER programme was used to tackle this issue. It is based on Euclidean Distance Metrics when comparing individual points. Even a very small set of fingerprint comparisons shows that fingerprint-matches between family members are very low. Fingerprints matches only on a minimum number of minutia. Thus, they can be distinguished from one other.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the National Sustainability Programme project No. LO1303 (MSMT-7778/2014) and also by the European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089

REFERENCES

[1] RAK, R. *Biometrics and identity of people: the forensic and commercial applications*, BEN, Prague, 2008. ISBN 978-80-247-2365-5.

[2] COUFAL, T. *What is FingerChip* [online]. 2007. <http://hw.cz/teorie-praxe/art2020-co-je-fingerchip.html>.

[3] BITTO, O. *Encryption and biometrics: or arcane bits and touches*.: Computer Media, 2005. ISBN 80-86686-48-5.

[4] LI, Haizhou, Liyuan LI a Kar-Ann TOH. *Advanced topics in biometrics*. New Jersey: World Scientific, c2012, xv, 500 s. ISBN 978-981-4287-84-5.

[5] JORGENSEN, Z a T. YU. *On Mouse Dynamics as a Behavioral Biometric for Authentication*. Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security. 2011; 476-482

[6] AHMED, Awad E. Ahmed a Issa TRAORÉ. *A New Biometrics Technology based on Mouse Dynamics*. IEEE Transactions on Dependable and Secure Computing. 2007; 4: 165-179.

[7] NAZAR, Akif, Issa TRAORÉ a AHMED Awad E. Ahmed Inverse *Biometrics for Mouse Dynamics*. International Journal of Pattern Recognition and Artificial Intelligence. 2008; 22: 461-495.

[8] Fujitsu Palmsecure. Fujitsu [online]. 2014. <http://www.fujitsu.com/cz/solutions/high-tech/palmsecure/>.

AUTHOR BIOGRAPHIES

MILAN ADÁMEK graduated in 1990 from the Olomouc Palacky University, Czech Republic. He received his Ph.D. degree in Technical Cybernetics at Tomas Bata University in Zlin in 2002. From 1997 to 2008 he worked as senior lecturer at the Faculty of Technology, Brno University of Technology. From 2008 he has been working as an associate professor at the Department of Electronic and Measurement, Faculty of Applied Informatics of the Tomas Bata University in Zlín, Czech Republic. Current work covers following areas: power lines, camera system, sensors. His e-mail address is: adamek@fai.utb.cz.

Petr NEUMANN has been graduated from the Brno Technical University in Electronic Technology in 1974. He has acquired the industrial experience in the field of medical electronics and quality management as R&D engineer. He received his Ph.D. degree in Technical Cybernetics at Tomas Bata University in Zlin in 2001. He has been lecturing and working in the university research area since 1994. He was engaged in the SMT technology training, equipment installation and servicing more than 10 years between 1997 and 2009. He is currently working as a senior lecturer at Tomas Bata University in Zlin. His research work is aimed at the electronic component authenticity analysis and failure diagnosis. His e-mail address is: neumann@fai.utb.cz

DORA LAPKOVÁ has been graduated from the Tomas Bata University in Zlin in 2012. Her scientific research is oriented into professional defence and self-defence, physical security and physical security technical equipment. Her e-mail address is: dlapkova@fai.utb.cz

MARTIN POSPÍŠILÍK graduated in 2008 from Czech Technical University in Prague, Czech Republic, in Microelectronics. Having received his Ph.D. degree in Engineering Informatics at Tomas Bata University in 2013, he became an assistant and researcher at the Department of Computer and Communication Systems of Faculty of Applied Informatics of the Tomas Bata University in Zlín, Czech Republic. His current research covers the following topics: electromagnetic compatibility, shielding effectiveness of materials for avionics, design of construction of electrical circuits and testing of electrical devices considering the security of communication. The security issues are investigated in cooperation with Escola Superior de Tecnologia e Gestão, Beja, Portugal. His e-mail address is: pospisilik@fai.utb.cz.

Simulation and Optimization

APPLICATION OF TWO PHASE MULTI-OBJECTIVE OPTIMIZATION TO DESIGN OF BIOSENSORS UTILIZING CYCLIC SUBSTRATE CONVERSION

Linas Litvinas and Romas Baronas
Faculty of Mathematics and Informatics
Vilnius University
Naugarduko 24, LT-03225, Vilnius, Lithuania
Email: linas.litvinas@mif.vu.lt

Antanas Žilinskas
Institute of Mathematics and Informatics
Vilnius University
Akademijos 4, LT-08663, Vilnius, Lithuania
Email: antanas.zilinskas@mii.vu.lt

KEYWORDS

Computational Modelling, Multi-Objective Optimization, Biosensor.

ABSTRACT

A method for the optimal design of amperometric biosensors with cyclic substrate conversion is proposed. The design is multi-objective since biosensors must meet numerous, frequently conflicting, requirements of users and manufacturers. Moreover, they should be technologically and economically competitive. To apply a multi-optimization technique, a mathematical model should be developed where the most important characteristics of the biosensor are defined as objectives, and the other characteristics and requirements are defined as constrains. For the considered biosensors the following characteristics are taken as objectives: the output current, the enzyme amount, and the biosensor sensitivity. The proposed method consists of two phases. At the first phase an approximated Pareto front is constructed, and a preliminary solution is selected. The second phase is aimed at specification of the Pareto front around the preliminary solution, and at making the final decision. A numerical example is presented using a computational model of an industrially relevant biosensor.

INTRODUCTION

A biosensor is a device capable to measure the concentration of an analyte (Scheller and Schubert 1992; Turner et al. 1987). A catalytic biosensor is based on the enzymatic reaction where the analyte is turned to a measurable product (Banica 2012). Amperometric biosensors measure the changes in the output current on the working electrode due to the direct oxidation or reduction of the products of biochemical reactions (Grieshaber et al. 2008). In this work a biosensor utilizing the cyclic substrate conversion was considered. Biosensors with cyclic substrate conversion are of particular interest due to high sensitivity possible by utilizing a cyclic substrate conversion in a single enzyme membrane (Kulys and Vidziunaite 2003).

The efforts to design biocatalytical systems can be mightily reduced by the aid of computer tools (Dagan and Bercovici 2014). The multi-objective optimization has been successfully applied to design biochemical systems (Vera et al. 2010; Taras and Woinaroschy 2012), to increase the productivity of multi-enzyme systems (Ardao and Zeng 2013) and for optimal design of synergistic amperometric biosensors (Baronas et al. 2016).

The multi-objective optimization is often used as a tool to get a representative set of trade-off solutions for a product to be designed (Žilinskas et al. 2006). The trade-off solutions, also known as Pareto optimal solutions, together with trade-off curves as their visualization are widely used for learning and making decisions when designing products (Maksimovic et al. 2012). Amperometric biosensors can be rather efficiently designed and optimized by combining the multi-objective optimization and multi-dimensional visualization with computational modelling (Baronas et al. 2016).

One and two compartment mathematical and corresponding numerical models for particular amperometric biosensors utilizing cyclic substrate conversion are already known (Sorochinskii and Kurganov 1996; Baronas et al. 2004a,b). In this work, a more complex biosensor involving a dialysis membrane is modelled and optimized. The modelling biosensor comprises three compartments, an enzyme layer, a dialysis membrane and an outer diffusion layer. The model is based on non-stationary reaction-diffusion equations containing a non-linear term related to Michaelis-Menten kinetic of the enzymatic reaction (Baronas et al. 2010). The numerical simulation of the biosensor action was carried out using the finite difference technique (Britz and Strutwolf 2016).

The objective functions in such applications are often non-linear and therefore the multi-objective optimization consumes much computational time, especially when the objectives are considered as expensive black-box functions, whose values can be obtained by solving nonlinear partial differential equations numerically. Because of this a two phase multi-objective optimization was applied. Firstly, a particular trade-off solution is obtained and a Pareto front is analyzed. Then the preliminary solution is adjusted in a specific part of the Pareto front, the decision variables are calculated, and finally pilot plant tests can be performed. For the amperometric biosensors with cyclic substrate conversion the output current, the enzyme amount and the gain of the current were consid-

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

ered in the optimization.

MATHEMATICAL MODELLING

Modelling biosensor

An amperometric biosensor may be considered as enzyme membrane attached to an electrode. In the case of the biosensor utilizing cyclic substrate conversion, a measured substrate (S) is electrochemically converted to a product (P) which in an enzyme (E) reaction in turn is converted to the substrate (S) (Baronas et al. 2004b),

$$\mathrm{S} \longrightarrow \mathrm{P} \xrightarrow{\mathrm{E}} \mathrm{S}. \tag{1}$$

The modelling biosensor involves four regions: the enzyme layer where the enzymatic reaction as well as the mass transport by diffusion take place, a dialysis membrane and a diffusion limiting region where only the mass transport by diffusion takes place, and a convective region where the analyte concentration remains constant. The schematic view of the biosensor is presented in Fig. 1, where d_1, d_2 and d_3 are the thicknesses of the enzyme, dialysis and diffusion layers, respectively, $x = a_0 = 0$ corresponds to the electrode surface, and a_1, a_2, a_3 denote boundaries between the adjacent regions.

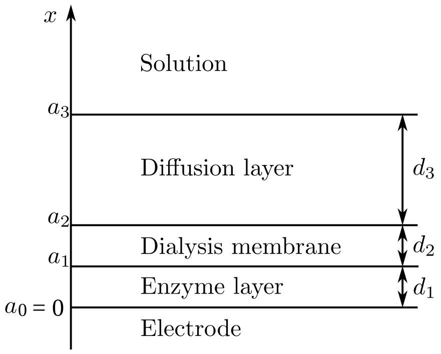


Figure 1: The Schematic View of the Biosensor.

Mathematical model of biosensor

Assuming symmetric geometry of the enzyme electrode, homogeneous distribution of the enzyme in the enzyme membrane and the uniform thickness of the dialysis membrane, the dynamics of the biosensor action can be described by the following reaction-diffusion system ($t > 0$):

$$\frac{\partial S_1}{\partial t} = D_{S_1}\frac{\partial^2 S_1}{\partial x^2} + \frac{V_{max}P_1}{K_M + P_1}, \tag{2a}$$

$$\frac{\partial P_1}{\partial t} = D_{P_1}\frac{\partial^2 P_1}{\partial x^2} - \frac{V_{max}P_1}{K_M + P_1}, x \in (0, a_1), \tag{2b}$$

$$\frac{\partial S_i}{\partial t} = D_{S_i}\frac{\partial^2 S_i}{\partial x^2}, \tag{2c}$$

$$\frac{\partial P_i}{\partial t} = D_{P_i}\frac{\partial^2 P_i}{\partial x^2}, x \in (a_{i-1}, a_i), i = 2, 3, \tag{2d}$$

where x and t stand for space and time, respectively, $S_i(x,t), P_i(x,t), i = 1, ..., 3$ are the substrate (S) and reaction product (P) concentrations in the enzyme layer ($i = 1$), dialysis membrane ($i = 2$) and diffusion layer ($i = 3$), D_{S_i}, D_{P_i}, $i = 1, ..., 3$ are the diffusion coefficients, V_{max} is the maximal enzymatic rate, and K_M is the Michaelis constant (Baronas et al. 2004a,b).

The biosensor operation starts when the substrate appears in the bulk solution. This is used in the initial conditions ($t = 0$),

$$P_i(x, 0) = 0, x \in [a_{i-1}, a_i], i = 1, ..., 3 \tag{3a}$$

$$S_i(x, 0) = 0, x \in [a_{i-1}, a_i], i = 1, 2, \tag{3b}$$

$$S_3(x, 0) = 0, x \in [a_2, a_3), S_3(a_3, 0) = S_0, \tag{3c}$$

where S_0 is the concentration of the substrate to be analyzed.

The substrate is assumed to be an electro-active substance. Due to the electrode polarization the concentration of the substrate at the electrode surface is permanently reduced to zero. The substrate at the electrode surface is electrochemically converted to the product. The product is generated at the same rate as the substrate is reduced. On the boundary between the adjacent regions having different diffusivities the matching conditions are defined ($t > 0$),

$$S_1(0, t) = 0,\ S_3(a_3, t) = S_0,\ P_3(a_3, t) = 0 \tag{4a}$$

$$D_{P_1}\left.\frac{\partial P_1}{\partial x}\right|_{x=0} = -D_{S_1}\left.\frac{\partial S_1}{\partial x}\right|_{x=0}, \tag{4b}$$

$$D_{S_i}\left.\frac{\partial S_i}{\partial x}\right|_{x=a_i} = D_{S_{i+1}}\left.\frac{\partial S_{i+1}}{\partial x}\right|_{x=a_i}, \tag{4c}$$

$$S_i(a_i, t) = S_{i+1}(a_i, t), \tag{4d}$$

$$D_{P_i}\left.\frac{\partial P_i}{\partial x}\right|_{x=a_i} = D_{P_{i+1}}\left.\frac{\partial P_{i+1}}{\partial x}\right|_{x=a_i}, \tag{4e}$$

$$P_i(a_i, t) = P_{i+1}(a_i, t),\ i = 1, 2. \tag{4f}$$

For simplicity, the functions S_i and P_i applicable to particular intervals $[a_{i-1}, a_i]$ ($i = 1, ..., 3$) are replaced with $S(x,t)$ and $P(x,t)$ applicable to whole domain, $x \in [0, a_3]$, $t \geq 0$. Both concentration functions are continuous in the entire domain.

During a physical experiment the anodic or cathodic current is measured as the biosensor response. The current depends on the flux of electro-active substrate (S) at electrode surface ($x = 0$). The current density $I(t)$ at time t can be explicitly calculated from Faraday's and Fick's laws,

$$I(t) = n_e F D_{S_1}\left.\frac{\partial S}{\partial x}\right|_{x=0} = -\ n_e F D_{P_1}\left.\frac{\partial P}{\partial x}\right|_{x=0}, \tag{5}$$

where n_e is the number electrons involved in charge transfer, F is Faraday's constant, $F \approx 9.65 \times 10^4$ C/mol.

During the biosensor operation the system (2)-(4) as well as the biosensor current $I(t)$ approaches the steady state,

$$I_\infty = \lim_{t\to\infty} I(t). \tag{6}$$

Computational simulation

The initial boundary value problem (2)-(4) is a nonlinear. Because of this the problem was solved numerically by applying the finite difference technique (Baronas et al. 2010). An explicit finite difference scheme was build as a result of the model discretization (Baronas et al. 2004a,b).

Some parameters of the model (2)-(4) are application-specific and cannot be changed or optimized by a biosensor designer (Banica 2012; Grieshaber et al. 2008). Meanwhile, values of some other parameters, e.g., the concentration of the enzyme as well as the biosensor geometry, can be selected by the designer quite freely. The following values specific to phenol sensitive biosensors and commercially available dialysis membranes were assumed to be constant (Kulys and Vidziunaite 2003; Banica 2012):

$$D_{S_1} = D_{P_1} = 3 \times 10^{-6}\text{cm}^2\text{/s}, \tag{7a}$$

$$D_{S_2} = D_{S_1}/10, \quad D_{P_2} = D_{P_1}/10, \tag{7b}$$

$$D_{S_3} = 2D_{S_1}, \quad D_{P_3} = 2D_{P_1}, \tag{7c}$$

$$K_M = 10^{-7}\text{mol/cm}^3, \quad n_e = 2. \tag{7d}$$

The chemical signal amplification is one the main features of amperometric biosensors utilizing cyclic substrate conversion (Kulys and Vidziunaite 2003). The rate of the steady state current of enzyme active electrode ($V_{max} > 0$) to the steady state current of the corresponding enzyme inactive electrode ($V_{max} = 0$) is considered as the gain G of the biosensor sensitivity (Baronas et al. 2004b),

$$G(V_{max}) = \frac{I_\infty(V_{max})}{I_\infty(0)}. \tag{8}$$

Due to the substrate cyclic conversion, the gain G of the sensitivity significantly depends on the geometry as well as catalytic activity of the biosensor and can be increased in some tens of times (Baronas et al. 2004a,b).

OPTIMAL DESIGN OF THE BIOSENSOR

The design of a biosensor can be mathematically reduced to a multi-objective optimization problem (Baronas et al. 2016). The complexity of biosensors involves consideration and simultaneous optimization of several often conflicting objectives, which means that if one of them is improved, the others get worse (Sadana and Sadana 2011). The solutions of multi-objective optimization is called Pareto optimal front (Deb 2009).

The goal of the optimal design is to find Pareto optimal solutions and by use of expert evaluation to find a particular trade-off solution which would satisfy needs of user and manufacturer (Žilinskas et al. 2006; Maksimovic et al. 2012; Žilinskas 2013; Žilinskas et al. 2015).

Most enzymes are expensive products and some of them are produced in very limited quantity (Sadana and Sadana 2011; Banica 2012). In such cases the optimization of the enzyme amount is important though the greater amount of the enzyme in some cases increases the range of calibration curve (Baronas et al. 2010).

The biosensor response is often perturbed by noise, e.g. white noise, sinusoidal power electrical noise, or if the biosensor response is biased, e.g. by temperature change (Hassibi et al. 2007). Miniaturized biosensors with small sensing area has low signal-to-noise ratio and it may result problems in measurement (Sadana and Sadana 2011). To reduce the negative influence of the signal noise to the biosensor sensitivity the biosensor current should be as high as possible.

The current of the biosensor utilizing cyclic substrate conversion monotonously increases with increasing the substrate concentration (Baronas et al. 2004b). Because of this, the biosensor current I_M calculated at a moderate concentration $S_0 = K_M$ of the substrate was assumed as the characteristics of the magnitude of the current of a particular biosensor,

$$I_M = I_\infty(K_M), \tag{9}$$

where $I_\infty(K_M)$ is density of the steady state current calculated assuming $S_0 = K_M$.

The gain of the biosensor sensitivity G shows the increase of the steady state current due to the enzyme catalized reaction. The high G indicates that the biosensor with a particular configuration effectively uses enzyme to amplify the current.

The maximal enzymatic rate V_{max} is proportional to enzyme amount ($V_{max} = kE$, k - reaction rate constant, E - enzyme concentration) and is attainable with that amount of enzyme, when the enzyme is fully saturated with the substrate. So, the maximal enzymatic rate can be changed by changing the enzyme concentration. The relative enzyme amount can be calculated as the product $V_{max}d_1$ of the maximal enzymatic rate and the thickness of enzyme layer.

The enzyme amount d_1V_{max}, the density I_M of the steady state current and the gain G of the sensitivity were optimized for the optimization of the biosensor utilizing cyclic substrate conversion.

Multi-objective optimization problem

Design of the biosensor with the cyclic substrate conversion can be stated as a three-objective optimization problem with the objective function $\Phi(x) = (\varphi_1(x), \varphi_2(x), \varphi_3(x))^T$, where $\varphi_1(x)$ is G, $\varphi_2(x)$ is I_M and $\varphi_3(x)$ is d_1V_{max}. Decision variables of the optimal design are given in Table 1.

Table 1: Decision Variables $x = (d_1, d_2, d_3, V_{max})^T$ for the Cyclic Biosensor Design Problem

Variable	Description	Range
d_1	Enzyme layer thickness, cm	$[2 \times 10^{-4}, 5 \times 10^{-2}]$
d_2	Dialysis membrane thickness, cm	$[10^{-4}, 10^{-2}]$
d_3	Diffusion layer thickness, cm	$[10^{-4}, 10^{-1}]$
V_{max}	Maximal enzymatic rate, mol/(cm^{3}s)	$[0, 10^{-6}]$

Range values of the decision parameters should be expertly evaluated. It depends on technological possibilities, e.g. the thicknesses of the commercially available dialysis membranes or the thicknesses of nylon nets used for the enzyme layer (Scheller and Schubert 1992).

The defined multi-objective minimization problem is difficult since the objective function Φ was defined by expensive black-box functions calculated from the numerical solution of the system (2)-(4) of non-linear partial differential equations. This feature of the objective function is a crucial factor for selecting an appropriate algorithm to find a Pareto front representation. The classical methods (Miettinen 1999) are efficient for smooth convex problems and not suitable here because of the of non-smoothness of the objective functions implied by the numerical errors of the solution of equations (2)-(4).

The application of metaheuristic methods is limited due to expensiveness of the function to be optimized, i.e. calculation of the objective function $\Phi(x)$ takes about 5 minutes using Intel Core i7-4770 3.5 GHz based personal computer. For optimization problems with given characteristics the most suitable is statistical model based algorithm (Žilinskas 2014), however software currently available only for bi-objective problems. Among other alternatives Chebyshev scalarization based methods seem most promising (Miettinen 1999).

First two objectives functions $(\varphi_1(x), \varphi_2(x))$ are maximized while the last one $(\varphi_3(x))$ is minimized. The optimization task to minimize objectives normalised to unit interval $[0, 1]$ can be formulated as follows:

$$\mathbf{F_P} = \min_{x \in \mathbf{A}} F(x), \tag{10a}$$

$$F(x) = (f_1(x), f_2(x), f_3(x))^T, \tag{10b}$$

$$f_i(x) = \frac{\varphi_i^+ - \varphi_i(x)}{\varphi_i^+ - \varphi_i^-}, \quad i = 1, 2, \tag{10c}$$

$$f_3(x) = \frac{\varphi_3(x) - \varphi_3^-}{\varphi_3^+ - \varphi_3^-}, \tag{10d}$$

$$\varphi_i^+ = \max_{x \in \mathbf{A}} \varphi_i(x), \quad i = 1, 2, 3, \tag{10e}$$

$$\varphi_i^- = \min_{x \in \mathbf{A}} \varphi_i(x), \quad i = 1, 2, 3, \tag{10f}$$

$$x = (x_1, \ldots, x_4)^T, \tag{10g}$$

$$\mathbf{A} = \{x : 0 \le x_j \le 1,\ j = 1, ..., 4\}, \tag{10h}$$

where $x_1, ..., x_4$ are the decision variables d_1, d_2, d_3 and V_{max} re-scaled to the unit interval.

The ranges for the objective functions $(\varphi_i^-, \varphi_i^+)$ were found by using the single criteria optimization of the corresponding function $\varphi_i(x)$, $i = 1, 2, 3$. The multi-start of Hooke-Jeeves algorithm was used for this (Kelly 1999).

The multi-criteria optimization besides finding Pareto front approximation $\mathbf{F_P}$ also finds optimal decision variables,

$$\mathbf{X_P} = \{x : F(x) \in \mathbf{F_P}\}. \tag{11}$$

Results of optimization

The Chebyshev scalarization was used to the transform multi-objective problem (10a) to a single objective problem,

$$f(x) = \max_{1 \le i \le 3} w_i f_i(x), \quad x(w) = \arg\min_{x \in \mathbf{A}} f(x), \tag{12a}$$

$$w = (w_1, w_2, w_3)^T,\ 0 \le w_i \le 1,\ \sum_{i=1}^{3} w_i = 1, \tag{12b}$$

where the minimizer $x(w)$ is the Pareto optimal solution of the original problem (10a).

All the Pareto optimal solutions can be found by solving (12a) with an appropriate weight vector w. To find an approximation of the Pareto front $\mathbf{F_P}$ and the decision vectors $\mathbf{X_P}$ the solution of (10a) should be found with a set of different weight vectors w. The optimization of the scalarized function $f(x)$ was performed by using the multi-start Hooke Jeeves algorithm (Kelly 1999), since it was successfully applied to similar problems (Žilinskas et al. 2006; Baronas et al. 2016). Some solutions of (10a) are weak Pareto optimal but they may be easily filtered.

The selection of weights to get an uniform distribution of $\mathbf{F_P}$ is rather complicated task. Search of Pareto front solutions was performed by a two step procedure. In the first step, the uniformly distributed weights were used as shown in Fig. 2a to solve the task (12a). Found Pareto optimal solutions are shown in Fig. 2b. One can see in Fig. 2b a gap in the Pareto front near square points. The corresponding weight vectors are shown as squares in Fig. 2a. To abolish the gap a more detailed representation of the Pareto front is needed in the neighbourhood of square points. In the second step, additional weight vectors (black points) are added to find solutions in neighbourhood of the square points in Fig. 3a. The supplemented representation of the Pareto front is presented in Fig. 3b. The gap is now completed by new solutions (black points). In figures the Pareto front solutions were given in the original dimensions $\Phi(x) = (\varphi_1(x), \varphi_2(x), \varphi_3(x))^T$ to be able expertly evaluate solutions in further analysis.

The described implementation of the two phase optimization procedure is still appropriate to run on a personal computer, even the multi-objective optimization was done with expensive black-box function. The computation of whole Pareto front took about a week (154 hours). The algorithm was parallelized by a master-slave approach using the Open MPI protocol (Gabriel et al. 2004). Eight parallel threads were used since the personal computer is based on Intel Core i7-4770 3.5 GHz processor. Each thread optimized function (12a) with different weight vectors w.

The analysis of the Pareto front $\mathbf{F_P}$ was performed to find an acceptable trade-off solution. The solution with the lowest enzyme amount $d_1 V_{max}$ = 8.4 pmol/(cm^{2}s) corresponds to the lowest steady state current I_M = 1.7 μA/(cm^2) and the lowest gain of the sensitivity $G = 1.8$. The solution with the highest enzyme amount

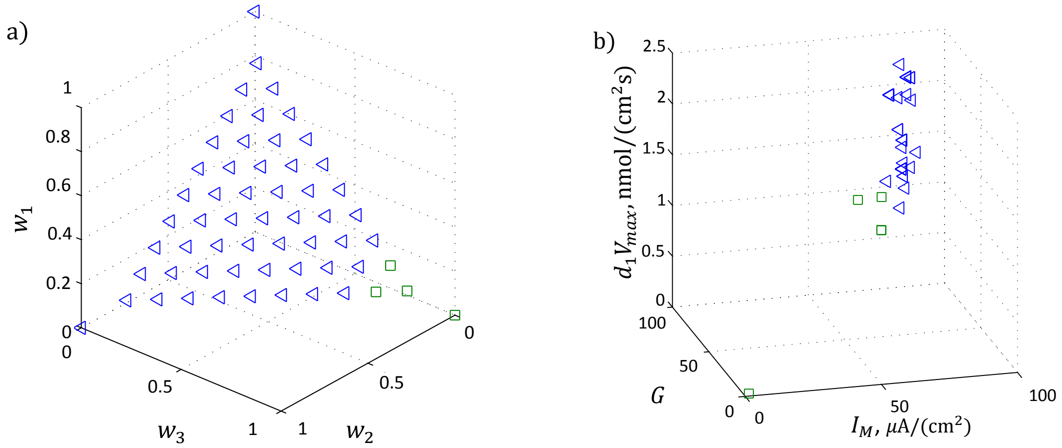


Figure 2: Weights (**a**) Used at the First Step of the Optimization Procedure and the Pareto Optimal Solutions (**b**)

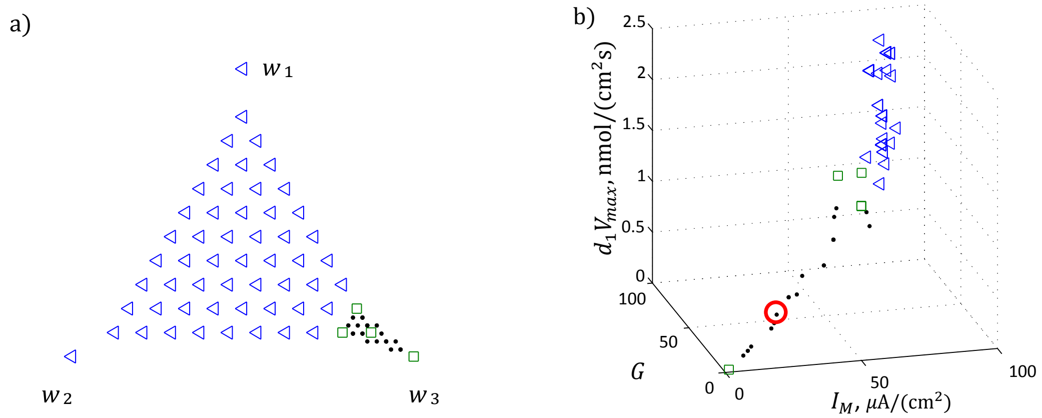


Figure 3: Additional Weights Indicated by Black Points in the Triangle of Weights (**a**) and the Complementary Representation of the Pareto Front (**b**)

$d_1V_{max} = 2.5$ nmol/(cm^{2}s) has the highest steady state current $I_M = 79.1$ μA/(cm^2) and the highest gain of the sensitivity $G = 80.1$. So, the steady state current and the gain of the sensitivity are proportional to the enzyme amount.

The steady state current and the sensitivity gain are not conflicting parameters, i.e. while one parameter increases also the other increases. An expert analysis of the Pareto front revealed that the solution marked with red circle (G, I_M, d_1V_{max}) = (25.3, 25.2 μA/(cm^2), 0.3 nmol/(cm^{2}s)) is best trade-off solution for the practical use $(d_1, d_2, d_3, V_{max}) = (1.45 \times 10^{-3}$ cm, 5.56×10^{-3} cm, 1.14×10^{-3} cm, 2.03×10^{-7} mol/(cm^{3}s)). It uses a relatively small amount of enzyme and gives an acceptably high the steady state current as well as the sensitivity gain.

The analysis of Pareto front decision variables $\mathbf{X_P}$ revealed that a very thin enzyme layer is used in Pareto optimal solutions, i.e. the range of enzyme layer thickness is near the low boundary of selected d_1 range (see Table 1): $d_1 \in (2.84 \times 10^{-4}\ \text{cm}, 2.52 \times 10^{-3}\ \text{cm})$. So, a thin enzyme layer should be used in the biosensor with the cyclic substrate conversion.

When comparing obtained the best trade-off solution with known configurations of the modelling biosensor particularly used for continuous flow-through measurements of phenol compounds in a alarm systems (Kulys and Vidziunaite 2003; Baronas et al. 2004a,b), one can see that the optimized biosensor provides about ten times greater signal gain $G = 25.3$ than the others at approximately the same enzyme amount ($d_1V_{max} =$ 0.3 nmol/(cm^2s)).

CONCLUSIONS

Optimal design of a biosensor is reducible to a problem of black-box multi-objective optimization. The objectives are computationally intensive since they are defined numerically via solution of a system of non linear partial differential equations. The Chebyshev scalarization based two phases method is appropriate to construct the approximation of the Pareto front the visualization of which greatly aids the design of the biosensor in question. The performed testing has shown that the computing resources of a personal computer are sufficient to design an industrially relevant biosensor by means of the proposed method.

ACKNOWLEDGEMENTS

The authors are grateful to Prof. Juozas Kulys for valuable discussions and his contribution to modelling of biosensors.

REFERENCES

Ardao, I. and A. P. Zeng. 2013. ”In silico evaluation of a complex multi-enzymatic system using one-potand modular approaches: Application to the high-yield production of hydrogen from a synthetic metabolic pathway.” *Chemical Engineering Science* 87, 183–193.

Banica, F. 2012. *Chemical Sensors and Biosensors: Fundamentals and Applications*. John Wiley & Sons, Chichester, UK.

Baronas, R.; F. Ivanauskas; and J. Kulys. 2004a. ”The effect of diffusion limitations on the response of amperometric biosensors with substrate cyclic conversion.” *Journal of Mathematical Chemistry* 35, No.3, 199–213.

Baronas, R.; F. Ivanauskas; and J. Kulys. 2010. *Mathematical Modeling of Biosensors*. Springer, Dordrecht.

Baronas, R.; J. Kulys; and F. Ivanauskas. 2004b. ”Modelling amperometric enzyme electrode with substrate cyclic conversion.” *Biosensors and Bioelectronics* 19, 915–922.

Baronas, R.; A. Žilinskas; and L. Litvinas. 2016. ”Optimal design of amperometric biosensors applying multi-objective optimization and decision visualization.” *Electrochimica Acta* 211, 586–594.

Britz, D. and J. Strutwolf. 2016. *Digital Simulation in Electrochemistry*. Monographs in Electrochemistry. Springer, Cham, 4 edition.

Dagan, O. and M. Bercovici. 2014. ”Simulation tool coupling nonlinear electrophoresis and reaction kinetics for design and optimization of biosensors.” *Analytical Chemistry* 86, 7835–7842.

Deb, K. 2009. *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons, Chichester, UK.

Gabriel, E.; G. Fagg; G. Bosilca; T. Angskun; J. Dongarra; J. Squyres; V. Sahay; P. Kambadur; B. Barrett; A. Lumsdaine; R. Castain; D. Daniel; R. Graham; and T. Woodall. 2004. ”Open mpi: Goals, concept, and design of a next generation mpi implementation.” In *Proceedings, 11th European PVM/MPI Users' Group Meeting*, 97–106.

Grieshaber, D.; R. MacKenzie; J. Vörös; and E. Reimhult. 2008. ”Electrochemical biosensors - sensor principles and architectures.” *Sensors* 8, 1400–1458.

Hassibi, A.; H. Vikalo; and A. Hajimiri. 2007. ”On noise processes and limits of performance in biosensors.” *Journal of Applied Physics* 102, 014909.

Kelly, C. T. 1999. *Iterative methods for optimisation*. SIAM, Philadelphia.

Kulys, J. and R. Vidziunaite. 2003. ”Amperometric biosensors based on recombinant laccases for phenols determination.” *Biosensors and Bioelectronics* 18, 319–325.

Maksimovic, M.; A. Al-Ashaab; R. Sulowski; and E. Shehab. 2012. ”Knowledge visualization in product development using trade-off curves.” In *IEEE International Conference on Industrial Engineering & Engineering Management*, 708–711.

Miettinen, K. 1999. *Nonlinear multiobjective optimization*. Kluwer, Dordrecht.

Sadana, A. and N. Sadana. 2011. *Handbook of Biosensors and Biosensor Kinetics*. Elsevier, Amsterdam.

Scheller, F. and F. Schubert. 1992. *Biosensors*. Elsevier Science, Amsterdam.

Sorochinskii, V. and B. Kurganov. 1996. ”Steady-state kinetics of cyclic conversions of substrate in amperometric bienzyme sensors.” *Biosens. Bioelectron.* 11, 225–238.

Taras, S. and A. Woinaroschy. 2012. ”An interactive multiobjective optimization framework for sustainable design of bioprocesses.” *Computers and Chemical Engineering* 43, 10–22.

Turner, A. P.; I. Karube; and G. S. Wilson. 1987. *Biosensors: fundamentals and applications*. Oxford University Press, Oxford.

Vera, J.; C. González-Alcón; A. Marín-Sanguino; and N. Torres. 2010. ”Optimization of biochemical systems through mathematical programming: Methods and applications.” *Computers & Operations Research* 37, 1427–1438.

Žilinskas, A.; E. S. Fraga; J. Beck; and A. Varoneckas. 2015. ”Visualization of multi-objective decisions for the optimal design of a pressure swing adsorption system.” *Chemom. Intell. Lab. Syst.* 142, 151–158.

Žilinskas, A.; E. S. Fraga; and A. Mackutė. 2006. ”Data analysis and visualisation for robust multi-criteria process optimisation.” *Computers & Chemical Engineering* 30, No.6, 1061–1071.

Žilinskas, A. 2013. ”On the worst-case optimal multi-objective global optimization.” *Optimization Letters* 7, 1921–1928.

Žilinskas, A. 2014. ”A statistical model-based algorithm for black-box multi-objective optimization.” *International Journal of System Science* 45, No.1, 82–93.

AUTHOR BIOGRAPHIES

LINAS LITVINAS was born in 1987 in Vilnius, Lithuania. He is a PhD student of computer science in Vilnius University. His research interests lie computational modelling, optimization and artificial intelligence.

ROMAS BARONAS was born in 1959 in Kybartai, Lithuania. He is a professor and serves as chair of the Department of Software Engineering at Vilnius University. Prof. Baronas received his MSc degree in Applied Mathematics in 1982 and then obtained his PhD degree in Computer Science in 2000 from the Vilnius University. His teaching and research interests lie in the areas of database systems and computational modelling of biochemical processes.

ANTANAS ŽILINSKAS was born in 1946 in Naujamiestis, Lithuania. Studied engineering cybernetics in Kaunas University of Technology, and got there the PhD degree. Habilitation in St. Petersburg University. Principal Researcher and Head of Optimization Section of Institute of Mathematics and Informatics, Professor of Department of Computer Science. Interested in theory and application of global and multi-objective optimization, multi-dimensional data visualization, optimal design.

EVIDENCE OF THE RELEVANCE OF MASTER PRODUCTION SCHEDULING FOR HIERARCHICAL PRODUCTION PLANNING

Thorsten Vitzthum
Innovation and Competence Centre for Production
Logistics and Factory Planning (IPF)
Ostbayerische Technische Hochschule Regensburg
Prüfeninger Str. 58, 93049 Regensburg, Germany
E-mail: thorsten.vitzthum@oth-regensburg.de
Faculty of Business and Economics
Technische Universität Dresden
Markt 23, 02763 Zittau, Germany
E-mail: thorsten.vitzthum@tu-dresden.de

Frank Herrmann
Innovation and Competence Centre for Production
Logistics and Factory Planning (IPF)
Ostbayerische Technische Hochschule Regensburg
Prüfeninger Str. 58, 93049 Regensburg, Germany
E-mail: frank.herrmann@oth-regensburg.de

KEYWORDS

Hierarchical production planning, aggregate production planning, master production scheduling, material requirement planning, planned independent requirements.

ABSTRACT

This paper deals with the significance of master production scheduling for hierarchical production planning.

Production planning in a typical manufacturing organization is a sequence of complex decisions which depends on a number of factors, such as number of products, complexity of products, number of production sites, and number of work centres in each production site.

The main idea in hierarchical production planning is to break down larger problems into smaller, more manageable sub problems.

Starting with aggregate production planning, the benefits of using master production scheduling for material requirements planning will be conveyed.

The main benefit of master production scheduling is more detailed planning. The production groups are thereby disaggregated into final products and the production site disaggregated into work centres. In order to plan capacities, resource profiles are used. By working with resource profiles production lead time data are taken into account to provide time-phased projections of the capacity requirements for each work centre (Vollmann et al. 2005).

The case study will show that more accurate planning and consideration of production lead time through master production scheduling results in a demand program that can be realized without shortages or delays.

INTRODUCTION

In the simultaneous planning approach all relevant decision parameters for the production program planning are taken into account at the same time. A high number of parameters and their mutual dependencies lead to very long computation times. Another disadvantage of this approach is that not all data are available at the same time and at the same level of detail.

In contrast, in the hierarchical production planning approach (Herrmann 2011), the complex singular problem of production planning is replaced by several manageable problems. These problems are solved successively. The individual solutions are then combined into one overall solution. This successive planning concept is implemented in most commercial production planning and control systems.

In this planning concept aggregation plays an important role. There are three different types of aggregation. The aggregation of time especially the period size, decision variables like grouping of products or constraints like grouping of machines to work centers or production sites (Stadtler 1988 and Gebhard 2009).

The hierarchical production planning approach is attributed to Hax and Meal 1975. This paper is based on a hierarchical planning concept that is expanded by limited capacities of resources, as suggested in Drexl et al. 1994. Typical hierarchical production planning is shown in figure 1.

The hierarchical production planning approach consists of the following steps: Aggregate production planning, master production scheduling, material requirements planning and scheduling.

The period size is freely scalable for each step and depends primarily on the planned product and the associated product structure tree.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

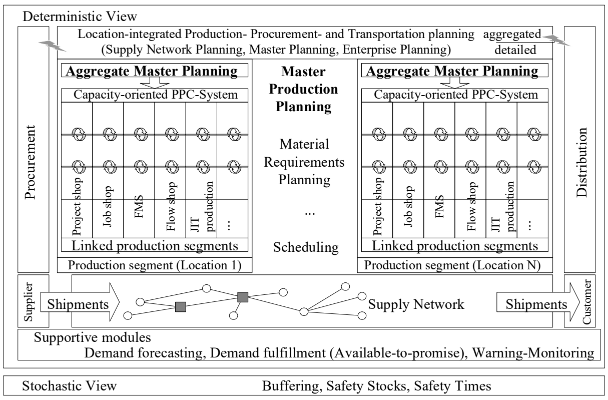

Deterministic View
Location-integrated Production- Procurement- and Transportation planning
(Supply Network Planning, Master Planning, Enterprise Planning)
aggregated
detailed
Procurement
Distribution
Aggregate Master Planning
Capacity-oriented PPC-System
Project shop
Job shop
FMS
Flow shop
JIT production
...
Linked production segments
Production segment (Location 1)
Master Production Planning
Material Requirements Planning
...
Scheduling
Aggregate Master Planning
Capacity-oriented PPC-System
Linked production segments
Production segment (Location N)
Supplier
Shipments
Supply Network
Shipments
Customer
Supportive modules
Demand forecasting, Demand fulfillment (Available-to-promise), Warning-Monitoring
Stochastic View
Buffering, Safety Stocks, Safety Times

Figure 1. Overview Production Planning and Control (Günther and Tempelmeier 2014)

AGGREGATE PRODUCTION PLANNING

Aggregate production planning predicts the planned independent requirements for one production site on a product group level based on demand forecasts. It is used by companies which offer a large spectrum of end products. It is difficult to predict the requirements for each individual product. Any forecast for a product has a forecast error. If the forecasts for individual products are summarized by prognosticating a product group, the forecast error is reduced due to variance reduction. The main benefit is the more accurate demand for planning process.

The planning horizon at this step covers at least one year, typically covering one seasonal demand pattern of the product group. So the planning horizon will typically cover between one and two years. The planning horizon will typically have a granularity of quarters, months or weeks.

MASTER PRODUCTION SCHEDULING

Master production scheduling determines the quantities of final products for a medium term period (several weeks to one quarter of a year). Therefore, aggregate production planning provides guidelines for master production scheduling in regard to the minimum production quantity and maximum overload capacity. The planning horizon will typically cover a medium term period (several weeks, to one quarter of a year). Master production scheduling is based on demand forecast for final products. The planning horizon will typically have a granularity of weeks.

At this stage the periods of aggregate production planning are disaggregated into smaller periods which are used in master production scheduling. Also the production groups are disaggregated into final products and the production site is disaggregated into work centres.

To reduce the complexity of disaggregation, in aggregate production planning and master production scheduling the same period size is used.

MATERIAL REQUIREMENTS PLANNING

The main task of material requirements planning is to determine the secondary requirements. It is a process of translating primary requirements into component part requirements which considers existing inventories and scheduled receipts (Kurbel 2013).

The planning horizon at this step covers usually one week and as period size is there will be used size days or shifts.

At these step typically lot size planning is also done. The main task of lot sizing is to minimize set up and storage costs (Herrmann 2009).

Material requirements planning determines the quantities and the completion dates for a short-term periode (several days) based on the results of the master production scheduling.

The result of the material requirements are planned orders.

SCHEDULING

At these step a production order is allocated to a concrete machine. Scheduling determines the exact processing times and the order of the operations on the machines.

Also important for scheduling are the capacities which can be used. These depends on capacity and maintenance data of machines, but also of factory calendars and shift models (Kurbel 2013).

CASE STUDY

In this paper it will be discussed, whether aggregate production planning is sufficient to get a realizable demand programm or whether master production scheduling is needed.

The point of this paper is best illustrated by following a clear example from a case study.

In order to clearly highlight the benefiting effects, it is assumed that the predicted planned independent requirements are identical to the real customer requirements.

To be able to compare results, the period size of aggregate production planning and of master production scheduling is the same.

The planning horizon starts in period 1 and ends with the last customer demand in period 10.

In this case study one production site with two work centres is considered. The capacity of each work centre is 500 hours per period. So the production site has a capacity of 1000 hours per period. In this case study the human capacity is equal the technical capacity and there is no additional capacity.

The production site produces one production group (P). The production group (P) consists of one final product

(E). The final product (E) consists of one component (V). The product structure tree is shown in figure 2.

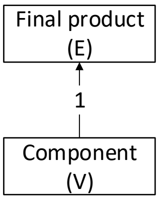


Figure 2. Product structure tree of the final product

The final product (E) is produced in work centre A and needs 4 hours of capacity per unit, the component (V) is produced in work centre B and needs 6 hours of capacity per unit. So the production group (P) needs in the production site a capacity of 10 hours per unit.

The estimated lead time for the final product (E) and the component (V) is 1 period. No set-up time is needed and there are no set-up costs. The inventory holding costs (h_k) are 2 € per unit.

The derived requirements (dependent requirements) to all components are produced just in time.

In this case study the model of a closed production is used. Closed production is characterized by the fact that each planned order has to be finished and stored, before it can be further processed or shipped (Herrmann 2011).

The demand for this case study is shown in table 1.

Table 1. Demand

Period (t)	1	2	3	4	5	6	7	8	9	10	$\sum$
Demand (d_t^A) [units]	0	0	0	0	80	100	120	100	80	120	600

The production program will be created once with the LP-Model for aggregate production planning and once with the LP-Model for the master production scheduling, which are located in the same named sections.

LP MODEL FOR AGGREGATE PRODUCTION PLANNING

The following shows the LP model for aggregate production planning (Günther and Tempelmeier 2014).

Parameters:

K Number of product groups $(1 \leq k \leq K)$.

T Length of the planning interval $(1 \leq t \leq T)$.

b_t Maximum technical production capacity in period t $\forall\, 1 \leq t \leq T$.

co_t Costs for one unit overtime in period t $\forall\, 1 \leq t \leq T$.

$d_{k,t}^A$ Demand for product group k in period t $\forall\, 1 \leq k \leq K$ and $1 \leq t \leq T$.

h_k Inventory holding costs for one unit of product group k $\forall\, 1 \leq k \leq K$.

n_t^{max} Maximum normal human capacity in period t $\forall\, 1 \leq t \leq T$.

o_t^{max} Maximum overtime in period t $\forall\, 1 \leq t \leq T$.

tb_k Human capacity absorption factor for one unit of product group k $\forall\, 1 \leq k \leq K$.

tc_k Technical capacity absorption factor for one unit of product group k $\forall\, 1 \leq k \leq K$.

Variables:

o_t Used overtime in period t $\forall\, 1 \leq t \leq T$.

$x_{k,t}^A$ Production quantity of product group k in period t $\forall\, 1 \leq k \leq K$ and $1 \leq t \leq T$.

$y_{k,t}^E$ Inventory of product group k at the end of period t $\forall\, 1 \leq k \leq K$ and $1 \leq t \leq T$.

Objective function:

The inventory costs for product groups (k) and the cost of additonal capacity are to be minimized.

$$Minimize\ Z = \sum_{k=1}^{K}\sum_{t=1}^{T} h_k \cdot y_{k,t}^E + \sum_{t=1}^{T} co_t \cdot o_t \qquad (1)$$

Constraints:

Inventory balance equation: (2)

$$y_{k,t-1}^E + x_{k,t}^A - y_{k,t}^E = d_{k,t}^A\ \forall\, 1 \leq k \leq K \text{ and } 1 \leq t \leq T$$

Human capacity constraints: (3)

$$\sum_{k=1}^{K} tb_k \cdot x_{k,t}^A - o_t \leq n_t^{max}\ \forall\, 1 \leq t \leq T$$

Technical capacity constraints: (4)

$$\sum_{k=1}^{K} tc_k \cdot x_{k,t}^A - o_t \leq b_t\ \forall\, 1 \leq t \leq T$$

Additional capacity constraints: (5)

$$o_t \leq o_t^{max}\ \forall\, 1 \leq t \leq T$$

Non-negativity constraints for all variables: (6)

$$x_{k,t}^A, y_{k,t}^E, o_t \geq 0\ \forall\, 1 \leq k \leq K \text{ and } 1 \leq t \leq T$$

Initialization of the starting inventory: (7)

$y^{E}_{k,0}$ given $\forall\, 1 \leq k \leq K$

Minimization problem:

Minimize Z.

AGGREGATE PRODUCTION PLANNING

Solution of Aggregate Production Planning

The optimal solution of the LP-Model of aggregate production planning is shown in table 2 and illustrated in figure 3. The objective is 120€. Because of the usage of a LP-Model there is no shortage.

Table 2. Solution of aggregate production planning

Period (t)	1	2	3	4	5	6	7	8	9	10	$\sum$
$\mathrm{d}^{\mathrm{A}}_{\mathrm{E,t}}$ [units]	0	0	0	0	80	100	120	100	80	120	600
$\mathrm{x}^{\mathrm{A}}_{\mathrm{E,t}}$ [units]	0	0	0	100	100	100	100	100	100	0	600
$\mathrm{y}^{\mathrm{E}}_{\mathrm{E,t}}$ [units]	0	0	0	0	20	20	0	0	20	0	60

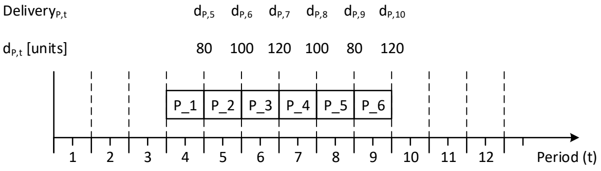


Figure 3. Solution of aggregate production planning

Realization of Aggregate Production Planning

The situation changes completely, when we look at the realization of the solution. Because of a delay no costumer requirements can be satisfied in time. The total shortage in the planning period is 1320 units and the total delay in periods is 13. The inventory holding costs will increase from 120 € to 160 €.

The realization of the solution is shown in table 3 and illustrated in figure 4. The first three periods do not contain demand or planned orders. For this reason they are not mentioned in the table.

Table 3. Realization of aggregate production planning

Period (t)	4	5	6	7	8	9	10	11	12	13	$\sum$
Demand ($\mathrm{d}^{\mathrm{A}}_{\mathrm{E,t}}$)	0	80	100	120	100	80	120	0	0	0	600
Planned order$^{\mathrm{A}}_{\mathrm{E,t}}$	100	100	100	100	100	100	0	0	0	0	600
Receipt$^{\mathrm{A}}_{\mathrm{E,t}}$	0	0	0	100	100	100	100	100	0	100	600
Delivery$^{\mathrm{A}}_{\mathrm{E,t}}$	0	0	0	80	100	120	100	80	0	120	600
Delay in periods	0	0	0	2	2	2	2	2	0	3	13
Inventory$^{\mathrm{E}}_{\mathrm{E,t}}$	0	0	0	20	20	0	0	20	20	0	80
Shortage$^{\mathrm{E}}_{\mathrm{E,t}}$	0	80	180	220	220	180	200	120	120	0	1320

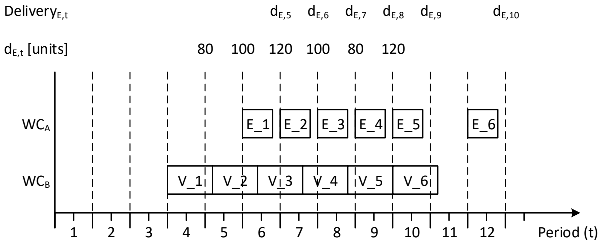


Figure 4. Realization of aggregate production planning

LP-MODEL FOR MASTER PRODUCTION SCHEDULING

The following shows the LP model for master production scheduling. (Günther and Tempelmeier 2014)

Parameters:

J Number of production segments $(1 \leq j \leq J)$.

K Number of products $(1 \leq k \leq K)$.

T Length of the planning interval $(1 \leq t \leq T)$.

$b_{j,t}$ Production capacity of production segment j in period t $\forall\, 1 \leq j \leq J$ and $1 \leq t \leq T$.

$co_{j,t}$ Costs for one unit overtime in production segment j in period t $\forall\, 1 \leq j \leq J$ and $1 \leq t \leq T$.

$d^{A}_{k,t}$ Demand for product k in period t $\forall\, 1 \leq k \leq K$ and $1 \leq t \leq T$.

$f_{j,k,z}$ Capacity absorption factor for product k with respect to production segment j in offset period z $\forall\, 1 \leq j \leq J$, $1 \leq k \leq K$ and $1 \leq z \leq Z_k$.

h_k Inventory holding costs for one unit of product k $\forall\, 1 \leq k \leq K$.

$o^{max}_{j,t}$ Maximum additional capacity in production segment j in period t $\forall\, 1 \leq j \leq J$ and $1 \leq t \leq T$.

Z_k Maximum lead time for product k $(1 \leq k \leq K)$.

Variables:

$o_{j,t}$ Used overtime in production segment j in period t $\forall\, 1 \leq j \leq J$ and $1 \leq t \leq T$.

$x^{A}_{k,t}$ Production quantity of product k in period t $\forall\, 1 \leq k \leq K$ and $1 \leq t \leq T$.

$y^{E}_{k,t}$ Inventory of product k at the end of period t $\forall\, 1 \leq k \leq K$ and $1 \leq t \leq T$.

Objective function:

The inventory costs for product (k) and the cost of additional capacity are to be minimized.

$$Minimize\ Z = \sum_{k=1}^{K} \sum_{t=1}^{T} h_k \cdot y^{E}_{k,t} + \sum_{t=1}^{T} \sum_{j=1}^{J} co_{j,t} \cdot o_{j,t} \tag{8}$$

Constraints:

Conditional statement inventory balance equation: (9)

$$y^E_{k,t-1} - y^E_{k,t} = d^A_{k,t} \;\forall\; 1 \le k \le K \text{ and } 1 \le t \le Z_k$$

$$y^E_{k,t-1} + x^A_{k,t} - y^E_{k,t} = d^A_{k,t} \;\forall\; 1 \le k \le K$$
and $Z_k + 1 \le t \le T$

Capacity constraints: (10)

$$\sum_{k=1}^{K} \sum_{z=0}^{Z_k} f_{j,k,z} \cdot x_{k,t+z} - o_{j,t} \le b_{j,t}$$

$\forall\; 1 \le j \le J$ and $1 \le t \le T$

Additional capacity constraint: (11)

$o_{j,t} \le o^{max}_{j,t} \;\forall\; 1 \le j \le J$ and $1 \le t \le T$

Non-negativity constraints for all variables: (12)

$x^A_{k,t}, y^E_{k,t}, o_{j,t} \ge 0 \;\forall\; 1 \le k \le K$ and $1 \le t \le T$

Initialization of the starting inventory: (13)

$y^E_{k,0}$ given $\forall\; 1 \le k \le K$

Minimization problem:

Minimize Z.

MASTER PRODUCTION SCHEDULING

Solution of Master Production Scheduling

The optimal solution of the LP-Model of master production scheduling is shown in table 4 and illustrated in figure 5. The objective is 872€.

Table 4. Solution of master production scheduling

Period (t)	1	2	3	4	5	6	7	8	9	10	$\sum$
$\mathrm{d}^{\mathrm{A}}_{\mathrm{E,t}}$ [units]	0	0	0	0	80	100	120	100	80	120	600
$\mathrm{x}^{\mathrm{A}}_{\mathrm{E,t}}$ [units]	0	19	83	83	83	83	83	83	83	0	600
$\mathrm{y}^{\mathrm{E}}_{\mathrm{E,t}}$ [units]	0	0	19	102	105	88	51	34	37	0	436

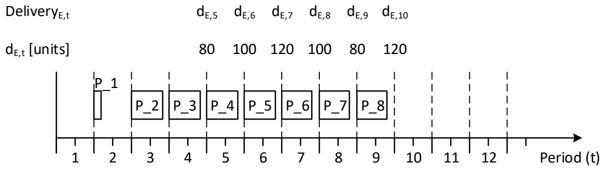


Figure 5. Solution of master production scheduling

Realization of Master Production Scheduling

Because of the usage of resource profiles in master production scheduling, in each period the lots can be processed in any order without exceeding the period capacity. Consequently, no shortage occurs in the individual periods, no matter which processing sequence is applied.

The complete data is shown in table 5 and illustrated in figure 6. The objective is 872€.

Table 5. Realization of master production scheduling

Period (t)	1	2	3	4	5	6	7	8	9	10	$\sum$
Demand ($\mathrm{d}^{\mathrm{A}}_{\mathrm{E,t}}$)	0	0	0	0	80	100	120	100	80	120	600
Planned order$^{\mathrm{A}}_{\mathrm{E,t}}$	0	19	83	83	83	83	83	83	83	0	600
Receipt$^{\mathrm{A}}_{\mathrm{E,t}}$	0	0	19	83	83	83	83	83	83	83	600
Delivery$^{\mathrm{A}}_{\mathrm{E,t}}$	0	0	0	0	80	100	120	100	80	120	600
Delay in periods	0	0	0	0	0	0	0	0	0	0	0
Inventory$^{\mathrm{E}}_{\mathrm{E,t}}$	0	0	19	102	105	88	51	34	37	0	436
Shortage$^{\mathrm{E}}_{\mathrm{E,t}}$	0	0	0	0	0	0	0	0	0	0	0

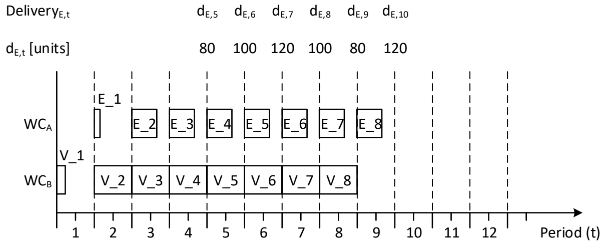

Delivery$_{E,t}$
d$_{E,5}$ d$_{E,6}$ d$_{E,7}$ d$_{E,8}$ d$_{E,9}$ d$_{E,10}$
d$_{E,t}$ [units]
80 100 120 100 80 120
E_1
WC$_A$
E_2 E_3 E_4 E_5 E_6 E_7 E_8
V_1
WC$_B$
V_2 V_3 V_4 V_5 V_6 V_7 V_8
1 2 3 4 5 6 7 8 9 10 11 12
Period (t)

Figure 6. Realization of master production scheduling

RESULTS

While in aggregate production planning capacity planning is done for product groups and production sites master production scheduling is done for final products and work centres. Furthermore master production scheduling takes production lead time into account through the usage of resource profiles (Vollmann et al. 2005).

A resource profile is a time-phased projection of capacity requirements for individual work centres. A resource profile gives information about the capacity requirements for final products in terms of work centres and offset periods.

The time-phased projection takes lead time into account, and so the planned orders start early enough to be produced in time.

This can be seen in this case study. Aggregate production planning starts with the production at the beginning of period 4 to cover the demand of period 5. As we see in the master production scheduling the final product needs two offset periods to be produced in time. Let's take a look what will happen in the case when we give the realization of aggregate production planning 2 offset periods. It should be mentioned that this is not part of the model and is only used for demonstration of the advantage of master production scheduling.

In this case the inventory costs are still 160€. The shortage decreases from 1320 units to 720 units and the delay in periods decreases from 10 to 5. This is a better result than without the consideration of the lead time, but master production scheduling is due to the non-occurrence of a shortage still significantly better.

The complete result is shown in table 6 and in figure 7.

Table 6. Realization of aggregate production planning with consideration of the lead time

Period (t)	4	5	6	7	8	9	10	11	12	13	$\sum$
Demand ($\mathrm{d}^{\mathrm{A}}_{\mathrm{E,t}}$)	0	80	100	120	100	80	120	0	0		600
Planned order$^{\mathrm{A}}_{\mathrm{E,t}}$	100	100	100	100	100	100	0	0	0		600
Receipt$^{\mathrm{A}}_{\mathrm{E,t}}$	0	0	100	100	100	100	100	0	100		600
Delivery$^{\mathrm{A}}_{\mathrm{E,t}}$	0	0	80	100	120	100	80	0	120		600
Delay in periods	0	0	1	1	1	1	1	0	2		7
Inventory$^{\mathrm{E}}_{\mathrm{E,t}}$	0	0	20	20	0	0	20	20	0		80
Shortage$^{\mathrm{E}}_{\mathrm{E,t}}$	0	80	100	120	100	80	120	120	0		720

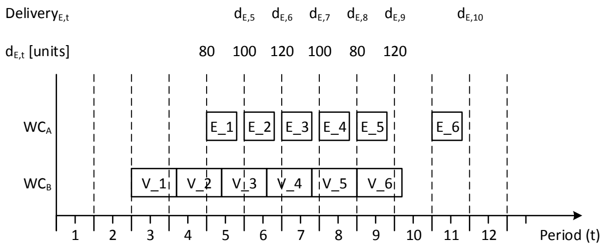


Figure 7. Realization of aggregate production planning with consideration of the lead time

Aggregated production planning, does not create a demand program which can be realized without shortages, despite the consideration of the lead time. To avoid shortages we have to increase the amount of offset periods by two to four. The result is shown in table 7 and in figure 8.

Table 7. Realization of aggregate production planning without delay

Period (t)	4	5	6	7	8	9	10	11	12	13	$\sum$
Demand ($\mathrm{d}^{\mathrm{A}}_{\mathrm{E,t}}$)	0	80	100	120	100	80	120				600
Planned order$^{\mathrm{A}}_{\mathrm{E,t}}$	100	100	100	100	100	100	0				600
Receipt$^{\mathrm{A}}_{\mathrm{E,t}}$	100	100	100	100	100	0	100				600
Delivery$^{\mathrm{A}}_{\mathrm{E,t}}$	0	80	100	120	100	80	120				600
Delay in periods	0	0	0	0	0	0	0				0
Inventory$^{\mathrm{E}}_{\mathrm{E,t}}$	100	120	120	100	100	20	0				560
Shortage$^{\mathrm{E}}_{\mathrm{E,t}}$	0	0	0	0	0	0	0				0

In this case the realization has no shortage as expected. The in inventory holding costs increase to 1020€ and the production has to start in period 1. For comparison the master production scheduling has inventory holding costs of 872€ and starts with the production in period 2.

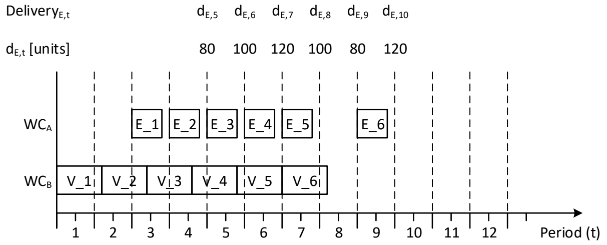


Figure 8. Realization of aggregate production planning without delay

That aggregated production planning, does not create a demand program which can be realized without shortages, without changing production program manually, is based on the following points:

The first reason for this is that in capacity planning the aggregate production planning considers the production site as a whole, while the master production scheduling considers each work centre separately.

The second and more important reason is the link between the final products, and the required capacity for work centres, which is taken into account by master production scheduling.

In this case study both work centres have the same capacity. The final product (E) needs 4 hours of capacity per unit of work centre A and the component (V) needs 6 hours of capacity per unit of work centre B.

Due to the asymmetric distribution of the capacity needed at the work centre aggregate production planning estimates the possible production quantity ($x^A_{E,t}$) false, aggregate production planning allows a production quantity ($x^A_{E,t}$) of 100 units per period, but there are only 83 units possible per period. Master production scheduling takes the right number of production quantity in account.

The production quantity ($x^A_{E,t}$) of aggregate production planning and master production scheduling can be compared in table 2 and table 4.

CONCLUSION

As shown in this case study, it is to be expected that aggregate production planning leads to shortages. Through a closer examination of the capacity of the master production scheduling, shortages can be avoided.

For this reason, the production planning in the context of hierarchical production planning should be additionally carried out by master production scheduling through upstream aggregated production planning approach.

Through a capacity restriction and the consideration of the exact offset periods on aggregate production planning level shortages are likely to be avoided. This and the relevance of industrial practice have to be further explored.

REFERENCES

Drexl, Andreas et al. (1994). “Konzeptionelle Grundlagen kapazitätsorientierter PPS-Systeme [Conceptual Foundations of Capacity-Oriented Systems for Production Planning and Control]”. In: *Zeitschrift für betriebswirtschaftliche Forschung* 46, pp. 1022–1045.

Gebhard, Marina (2009). *Hierarchische Produktionsplanung bei Unsicherheit [Hierarchical Production Planning under Uncertainty]*. 1. Aufl. Produktion und Logistik. Wiesbaden: Gabler.

Günther, Hans-Otto and Horst Tempelmeier (2014). *Produktion und Logistik: Supply Chain und Operations Management [Production and Logistics: Supply Chain and Operations Management]*. 11., verb. Aufl. Norderstedt: Books on Demand.

Hax, Arnoldo C. and Harlan C. Meal (1975). “Hierarchical Integration of Production Planning and Scheduling”. In: *Logistics*. Ed. by Murray A. Geisler. Vol. 1. TIMS studies in the management sciences. Amsterdam u.a.: North-Holland, pp. 53–69.

Herrmann, Frank (2009). *Logik der Produktionslogistik [Logic of production logistics]*. München: Oldenbourg.

Herrmann, Frank (2011). *Operative Planung in IT-Systemen für die Produktionsplanung und -steuerung: Wirkung, Auswahl und Einstellhinweise von Verfahren und Parametern [Operative Planning in IT-Systems for Production Planning and Control: Effects, Selection and Adjustment Notes]*. 1. Aufl. Studium : IT-Management und -Anwendung. Wiesbaden: Vieweg + Teubner.

Kurbel, Karl (2013). *Enterprise resource planning and supply chain management: Functions, business processes and software for manufacturing companies*. Progress in IS. Heidelberg: Springer.

Stadtler, Hartmut (1988). *Hierarchische Produktionsplanung bei losweiser Fertigung*. Vol. 23. Physica-Schriften zur Betriebswirtschaft. Heidelberg: Physica-Verl.

Vollmann, Thomas E. et al. (2005). *Manufacturing planning and control for supply chain management*. 5. ed., internat. ed. McGraw-Hill international editions. Boston, Mass.: McGraw-Hill.

AUTHOR BIOGRAPHIES

THORSTEN VITZTHUM was born in Augsburg, Germany in 1976. He received the B.Sc. in Business Information Technology and the M.Sc. in Applied Research in Engineering Sciences from Ostbayerische Technische Hochschule Regensburg (OTH Regensburg), Germany in 2015 and 2016. Since 2016 he is a doctoral student at the Faculty of Business and Economics at the Technische Universität Dresden (TU Dresden) and works on his PhD thesis Consideration of sustainability in production planning, integration of ecological and social boundary conditions and objectives into the hierarchical production planning and control. His e-mail address is: thorsten.vitzthum@tu-dresden.de

FRANK HERRMANN was born in Münster, Germany and went to the RWTH Aachen, where he studied computer science and obtained his degree in 1989. During his time with the Fraunhofer Institute IITB in Karlsruhe he obtained his PhD in 1996 about scheduling problems. From 1996 until 2003 he worked for SAP AG on various jobs, at the last as director. In 2003 he became Professor for Production Logistics at the University of Applied Sciences in Regensburg. His research topics are planning algorithms and simulation for operative production planning and control. His e-mail address is Frank.Herrmann@OTH-Regensburg.de and his Web-page can be found at http://www.oth-regensburg.de/frank.herrmann.

INFLUENCE OF RANDOM ORDERS ON THE BULLWHIP EFFECT

Hans-Peter Barbey
University of Applied Sciences Bielefeld
Interaktion 1, 33619 Bielefeld, Germany
Email: hans-peter.barbey@fh-bielefeld.de

KEYWORDS
Supply chain, bullwhip effect, simulation, random orders, seasonal trend, order strategies.

ABSTRACT

Supply chains in industry have a very complex structure. The influence of many parameters is not known. Therefore the control of the orders, material flow and stock is rather difficult. In order to recognize the basic relationships between the parameters, a very simple model was set up. It consists of 4 identical stages. In all stages the stock is closed-loop controlled to a nominal stock. Therefore the only decision which can be made in the entire supply chain is the quantity of an order. In a first simulation run a suitable order strategy will be defined. Good results can be realized, if an order is splitted up in two: A customers order and a stock order. In a second run this strategy will be applied to orders, which contain a seasonal trend an a random part. It will be shown that the bullwhip effect can be minimized with the applied order strategy.

1 INTRODUCTION

Dynamic behavior of the material flow in a supply chain is influenced by the order policy of each particular company of a supply chain. Sometimes, these orders are affected by seasonal trends and random influences. In combination with not defined interaction of all companies the bullwhip effect occurs. It has been described first by (Forrester 1958). It is the increasing of a small variation in the requirements of a customer to an enormous oscillation with the manufacturer at the beginning of a supply chain. In many articles, this phenomenon is only described in general terms without a mathematical definition (i.e. Erlach 2010 and Dickmann 2007). It is questionable if the bullwhip effect can be avoided at all (Bretzke 2008). A mathematical justification for this thesis is not given in that paper.

To minimize the bullwhip effect, cooperation between all members in a supply chain is necessary. Basically, informations about i.e. orders of customers have to be provided to all subsuppliers in the supply chain. Most of the supply chains do not have this kind of cooperation. Therefore in the following examination the particular stages are acting independent.

A very simple model of a supply chain without any cooperation between the particular members has been published on the ECMS2013 (Barbey 2013). The target of this simulation was to develop strategies for a closed-loop control of each stage of a supply chain. These controlling strategies have been applied to a seasonal trend in this simple simulation model. (Barbey 2014).

This model will be used with a controlling strategy, which includes a kind of cooperation between the members of the supply chain (Barbey 2015). The model is designed in the following manner:

The model consists of four identical stages according fig. 1. The behavior of each stage is the same. The time to place an order is 1 time unit (TU). The time for delivery is 3 time units. Therefore the replenishment lead time to fill up the stock for one stage is the sum of both, 4 time units. If a customer places an order, the lead time for the entire supply chain is 16 time units to deliver the material from the very beginning to the end of the supply chain. To be able to fulfill a customers order within the minimum lead time of 4 TU each stage needs a stock. This model is designed according the simple Forrester model.

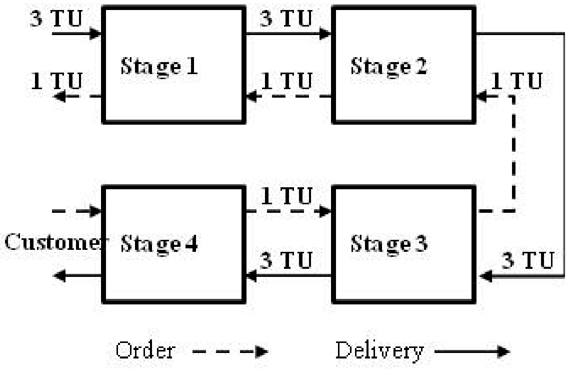


Figure 1: Model of a Supply chain
(TU= time unit)

The only decision, which can be made in this simulation, is to decide about the quantity of the order. This order has two tasks: It fulfills the predecessors order in the supply chain and compensates a difference in the own inventory. The applied controlling strategy for this decision will be described in chap. 2. The decision for an order has been taken each time unit. It is obvious that these parameters do not simulate a real supply chain. Normally the lead time to get material is much shorter than the time for the

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

next order. However, this simulation demonstrates with this short order period the bullwhip effect in a more impressive manner.
To demonstrate the bullwhip effect clearly, all other influences like delay in delivery or empty stock have been eliminated.

2 APPLIED ORDER STRATEGY

The dynamic behavior of a supply chain and different order strategies have been discussed in several papers (Barbey 2013 and Barbey 2014).

The order strategy, which is applied in this examination, is described below. The best strategy to fulfill a customers order is:

Order in = order out

This strategy works perfect, if the orders are constant over time, or they have included only a random part. If there is any kind of trend, this strategy leads to a deviation of the stock from the nominal stock in each stage of the supply chain. To fill up the stock to the nominal stock an additional quantity has to be ordered. This additional order has nothing to do with a customers order, it is only related to the behavior of a particular stage of the supply chain. Therefore it should be handled as a second order, the “stock order”. Now this stock order is independent from the customers order, i.e. in terms of the delivery time. The delivery time is only influenced by the decision how long the compensation of the stock will take. An example of this order strategy is given in fig. 2 and fig. 3.

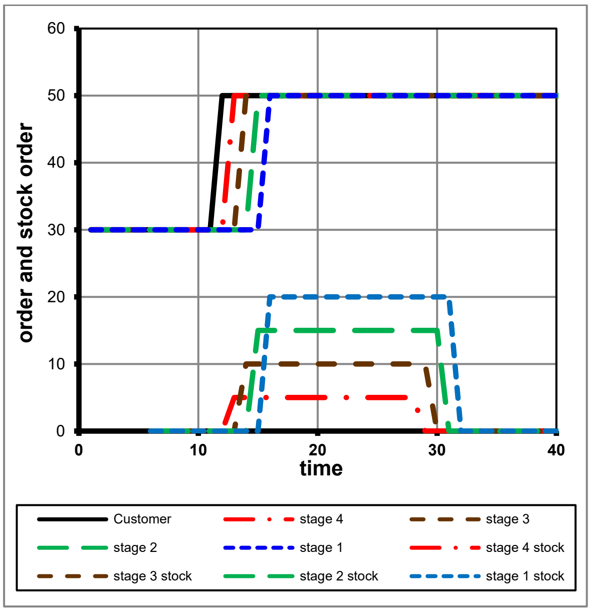


Figure 2: Order strategy of the closed-loop controller: Customers order and stock order with a compensation time of 16 time units

Fig. 2 shows the in=out strategy for the increase of the customers order from 30 to 50 units in the upper part. The lower part shows the stock order. The customers order is the same for all stages only with a time difference of one time unit, which was defined as the time to place an order (fig. 1). For the stock order the decision is to compensate the stock difference within 16 time units in all stages. The stock order increases from stage to stage. This is obvious because a stage has to compensate its own stock difference and all differences of the stages downstream. Therefore the bullwhip effect is only created by the stock orders.

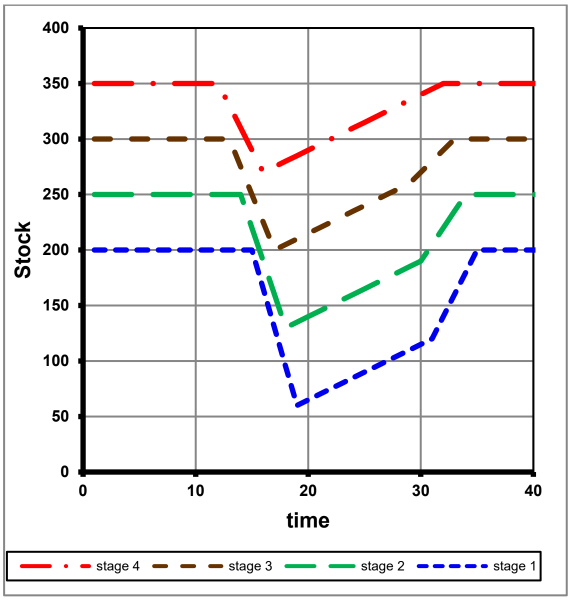


Figure 3: Stock compensation to the nominal stock within 16 time units

Each stage upstream has a higher stock difference (fig. 3). The bullwhip effect occurs in the stock difference too. However, each stage is able to compensate the stock difference in the same time.

3 RANDOM ORDERS

Normally orders are not constant over time. For the next simulation it is assumed, that there is a random part additional a constant order. The orders have a uniform distribution in a range of 200. It is obvious that there is no suitable controlling strategy for a constant stock, if the orders are random. The best strategy is in=out. With this strategy the stock has any values between a maximum and a minimum. In fig. 4 is the difference between the maximum and the minimum marked with a horizontal line. If now a controlling strategy acc. Chap. 2 is applied, a decision of the compensation time has to be done. In fig. 4 the compensation time has been changed from 2 to 122. For short compensation times there is a enormous bullwhip effect in the stock difference. The orders of the customer at the end of the the supply chain are uniform distributed in a range of 200. This order difference is marked in fig. 5 with the horizontal line. All other stages have a much higher order difference for short compensation times. There is an enormous bullwhip effect. For high compensation times the results of the in=out strategy are reached. That means for the order strategy: the longer the compensation time, the lower the bullwhip effect.

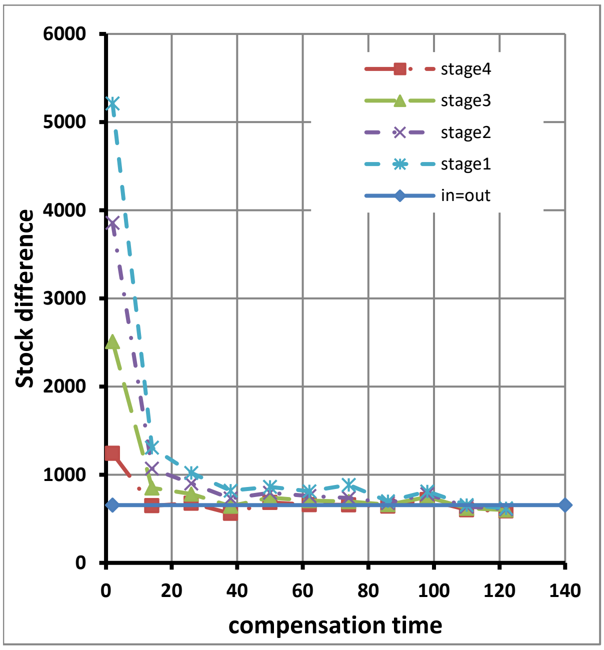


Figure 4: Stock difference for uniform distributed orders in a range of 200

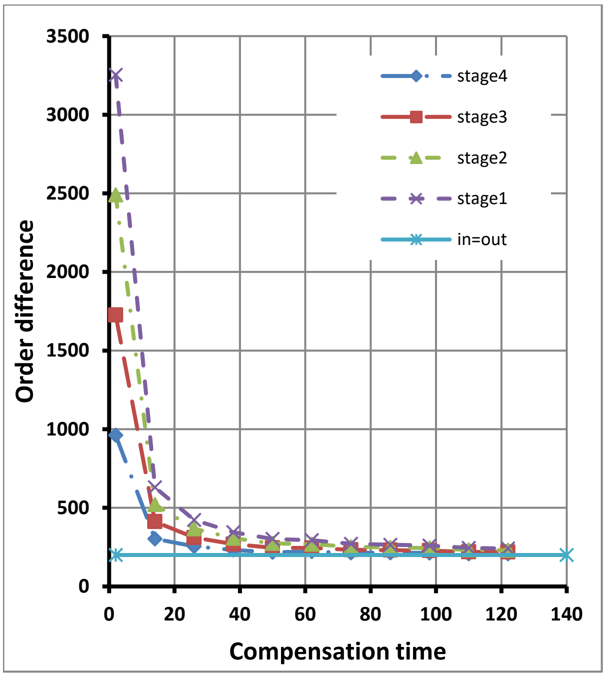


Figure 5: Order difference for uniform distributed orders in a range of 200

4 SEASONAL TREND

A seasonal trend with oscillating orders also leads to major changes in inventories. Therefore the aim must be to minimize the oscillation of the stock by an appropriate closed-loop control. If the oscillation of the stock is minimized, then the average stock is at a minimum too. A seasonal trend is simulated by a sine function very well. In this simulation the amplitude of the sine is +/- 200. Important for a seasonal trend is the period length. For this examination a period length of 300 (fig. 6 and fig. 7) and 100 (fig. 8 and fig. 9) have been applied.

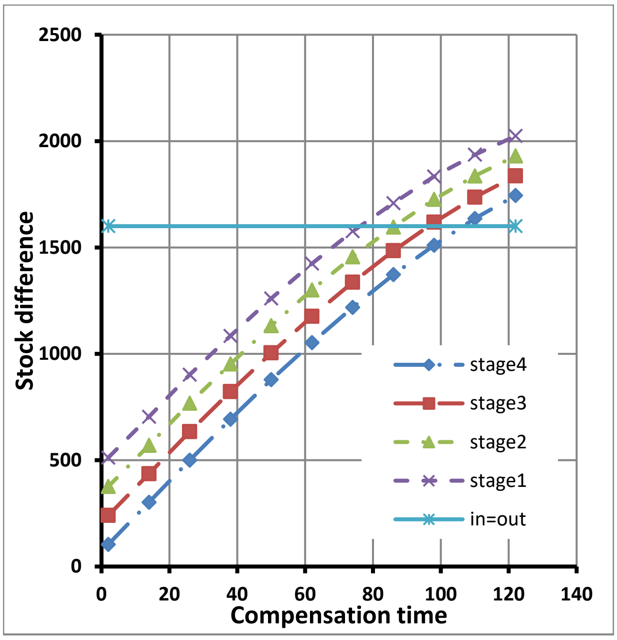


Figure 6: Stock difference at a period length of 300 for a seasonal trend

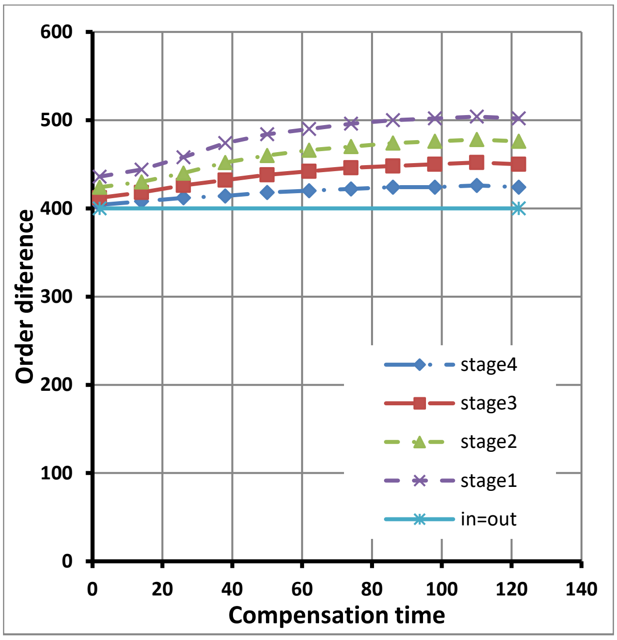


Figure 7: Order difference at a period length of 300 for a seasonal trend

For compensation times less than 70 there are better results for stage 1 than with the in=out strategy. For stage 1 is up to 110 an improvement. Is the period length of the seasonal trend shorter, here 100, the results are quite different (fig. 8 and fig. 9). For stage 1 it does not make any sense to apply another order strategy than in=out, because the stock difference is always higher. For the other stages is for short compensation times an improvement in the stock variation. But there is a enormous bullwhip effect in the the orders.

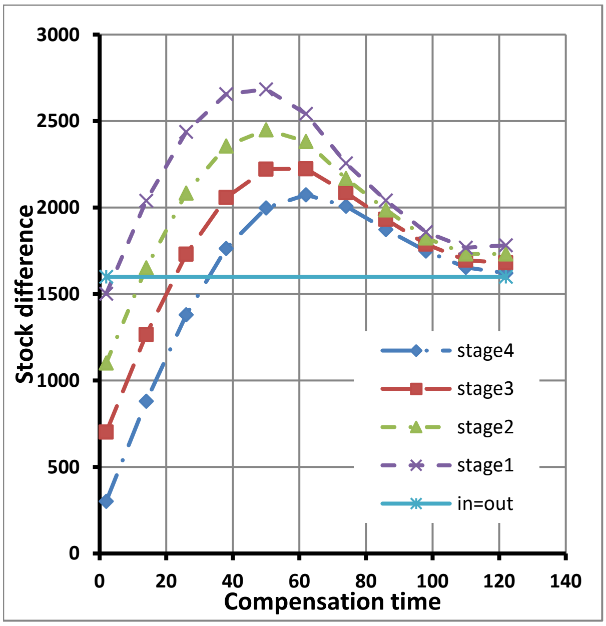


Figure 8: Stock difference at a period length of 100 for the seasonal trend

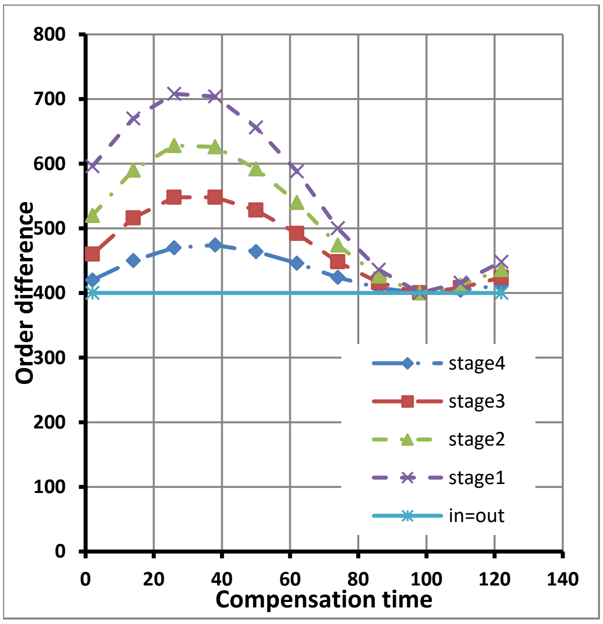


Figure 9: Order difference at a period length of 100 of the seasonal trend

5 ORDERS WITH A SEASONAL TREND AND A RANDOM PART

Random orders need a long compensation time, seasonal orders need short compensation times. The next simulation run should show, if a compensation time can be found to minimize the order difference and the stockdifference. For that the uniform distributed order with a variation of 200 and the seasonal trend with an amplitude of 200 are combined for a period length of 300 (fig. 10 and fig. 11).

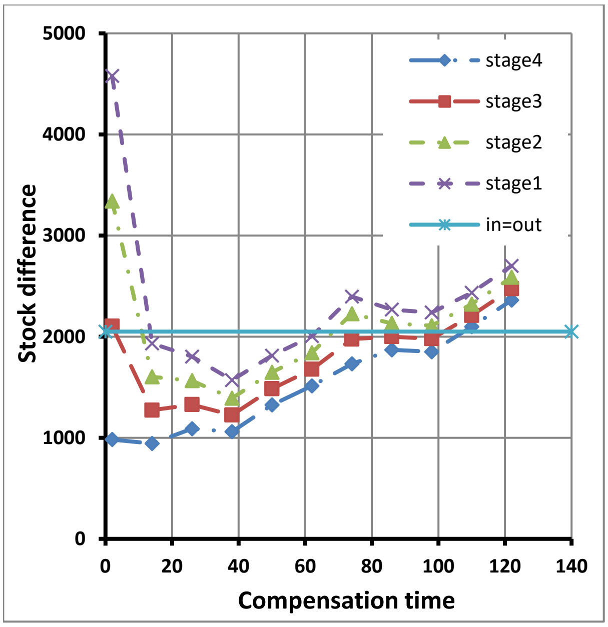


Figure 10: Stock difference at a period length of 300 for the seasonal trend and a uniform distributed random order

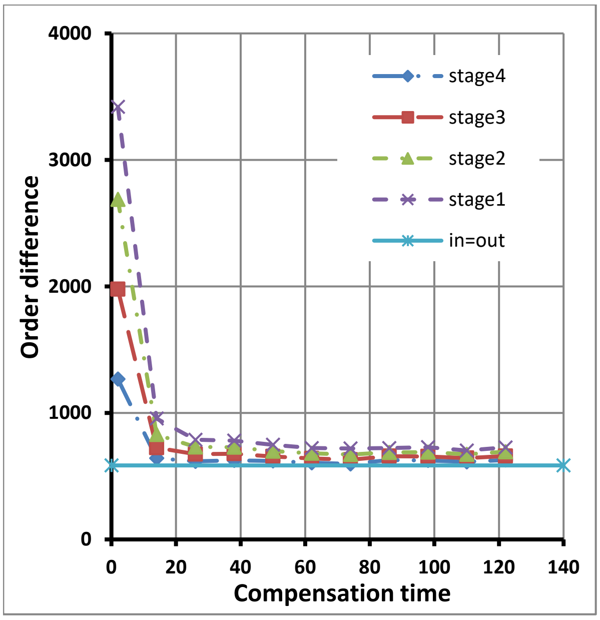


Figure 11: Order difference at a period length of 300 for the seasonal trend and a uniform distributed random order

For short compensation times is the influence of the random part very strong. To minimize the oder differences large compensation times have to be applied. For the order difference a bullwhip effect occurs at the very beginning. With higher compensation times the order difference is in the range of the in=out strategy, but never lower.

For the stock difference exsits a small range whre the results are better than with the in=out strategy. For a compensation time larger than 10 all stages are better than with the in=out strategy. For stage 1 ends this range at 60 and for stage 4 at 100.

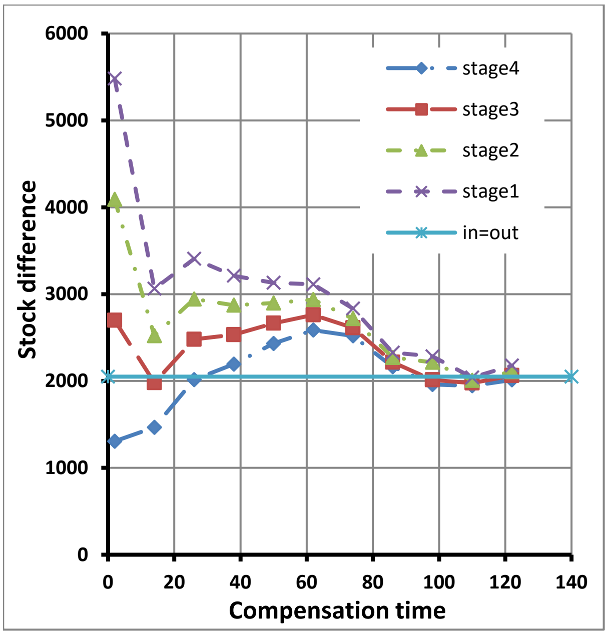


Figure 12:Stock difference at a period length of 100 for the seasonal trend and a uniform distributed random order

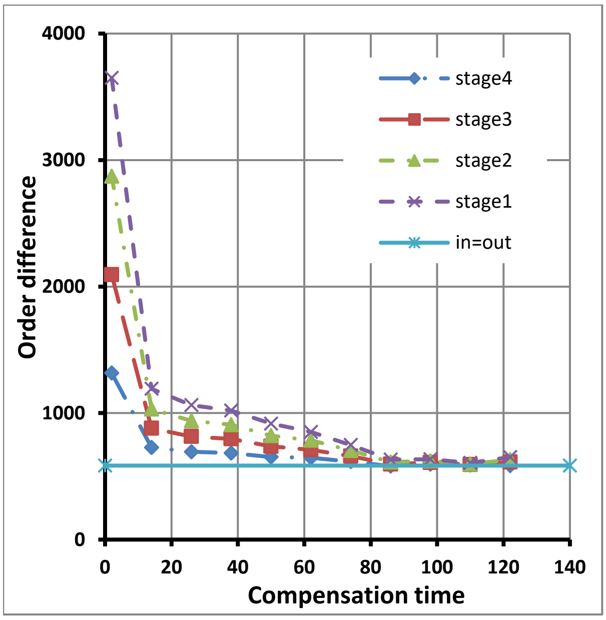


Figure 13: Order difference at a period length of 100 for the seasonal trend and a uniform distributed random order

Quite different are the results of the stock difference for a period length of 100 (fig. 12 and fig. 13). Only up to a compensation time of 20 stage 4 gets better results than with the in=out strategy. All other stages are worse.

6 SUMMARY AND CONCLUSIONS

This study is a theoretical view of the dynamics in a supply chain. For this examination a quite simple model according Forrester has been used. The advantage of a model like that is to see the main influences of the dynamic behavior of the supply chain.

The target of all stages in this examination is to keep the stock at a minimum with a seasonal trend and a uniform distributed random part in the orders. For that the orders have been splitted up in a customers order and a stock order. The customers orders have been handled with the in=out strategy. For the stock order the only decision was compensation time to bring the stock to the nominal stock.

For the random part of the orders long compensation times are neccessary. Then the results for the order difference and the stock difference are similar to the in=out strategy. The seasonal trend requires short compentation times. The stock differences are better with short compensation times than with the in=out strategy. The order differences are always higher than with the in=out strategy.

If seasonal and random orders are applied, only for long periods of the seasonal trend and a small range of the compensation time better results for the stock difference than with the in=out strategy can be found.

At this very simple model only partwise better results than with the in=out strategy can be found. It is doubtful, whether a real system with more parameters can be improved.

REFERENCES

Barbey, H.-P.: Seasonal Trends in Supply Chains. Proceedings of 28. European Conference on Modelling and Simulation (ECMS), Brescia, 2014, 748-752.

Barbey, H.-P.: Dynamic Behaviour of Supply Chains. Proceedings of 27. European Conference on Modelling and Simulation (ECMS), Alesund, 2013, 748-752.

Barbey, H.-P.: A New Method for Validation and Optimisation of Unstable Discrete Event Models, appeared in proceedings of 23. European Modelling & Simulation Symposium (EMSS), Rome, 2011.

Barbey, H.-P.: Simulation des Stabiltätsverhalten von Produktionssystemen am Beispiel einer lagerbestandsgeregelten Produktion, appeared in: Advances in Simulation for Production and Logistics Application, Hrsg.: Rabe, Markus, Stuttgart, Fraunhofer IRB Verlag, 2008, S.357-366.

Barbey, H.-P.: Application of the Fourier Analysis for the Validation and Optimisation of Discrete Event Models, appeared in proceedings of ASIM 2011, 21. Symposium Simulationstechnik, 7.9.-9.9.2011, Winterthur.

Bretzke, W.-R.: Logistische Netzwerke, Springer Verlag Berlin Heidelberg, 2008.

Dickmann, P.: Schlanker Materialfluss, Springer Verlag Berlin Heidelberg, 2007.

Erlach, K.: Wertstromdesign, Springer Verlag Berlin Heidelberg, 2010.

Forrester, J.W.: Industrial Dynamics: A major breakthrough for decision makers. In: Harvard business review, 36(4), 1958.

Gudehus, T.: Logistik, Springer Verlag Berlin Heidelberg, 2005.

AUTHOR BIOGRAPHIES

HANS-PETER BARBEY was born in Kiel, Germany, and attended the University of Hannover, where he studied mechanical engineering and graduated in 1981. He earned his doctorate from the same university in 1987. Thereafter, he worked for 10 years for different plastic machinery and plastic processing companies before moving in 1997 to Bielefeld and joining the faculty of the University of Applied Sciences Bielefeld, where he teaches logistic, transportation technology, plant planning, and discrete simulation. His research is focused on the simulation of production processes.
His e-mail address is:
`hans-peter.barbey@fh-bielefeld.de`
And his Web-page can be found at
`http://www.fh-bielefeld.de/fb3/barbey`

A DISCRETE ELEMENT MODEL FOR AGRICULTURAL DECISION SUPPORT

Ádám Kovács, János Péter Rádics and György Kerényi
Department of Product and Machine Design
Budapest University of Technology and Economics
H-1111, Budapest, Hungary
E-mail: kovacs.adam@gt3.bme.hu

KEYWORDS

Maize, corn, numerical model, decision support.

ABSTRACT

One of the main goals of precision agriculture researches is to define a suitable decision support system (DSS) for farm management with the goal of optimizing returns on inputs while preserving resources. Therefore, our long-term objective is the development of an adequate DEM model which can be adapted in the actual decision support system (DSS) to create a new model driven decision support system (MDDS) to refine operational parameters. In this study a preliminary step of our study is discussed: the predicted role of DEM modelling in decision making and a possible numerical model for corn stalk. The consequences of the study clearly demonstrate that the discrete element method is capable to reinforce agricultural decision support systems in the future and it is worth further analysing it.

INTRODUCTION

One of the main goals of precision agriculture researches is to define a preferably fully computerized decision support system (DSS) for farm management with the goal of optimizing returns on inputs while preserving resources (Chetty et al. 2014). In most cases, these kinds of decisions need more and more accurate information from independent sources. In turn, the ways of agricultural analogue, in-situ data collection are limited in time and space because of the seasonal characteristics of agricultural crops.

There is a need of an adequate method to calculate the physical and mechanical parameters of manipulated crops based on simple real time measurements. Our knowledge of the correlation of these parameters is very inadequate.

The utilization of corn plants is remarkable worldwide; the corn production in the world is almost 1000 million tons, but a comprehensive numerical model with easily calibratible parameters for the existent in-situ circumstances is missing to determine the optimal working parameters for cutting or rolling units.

Thanks to the natural diversity of physical and mechanical properties of agricultural materials, the accurate numerical modelling of these materials provides a huge challenge for researchers. In most cases during laboratorial or in situ tests, the parameters of interest usually show a wide confidence interval around the mean values. Nonetheless, these parameters could inordinately change in the same sample as well. To handle this problem a stochastic variation, as in the natural structures, should be used during the numerical simulations.

The key to this could be the discrete element method (DEM) with the Timoshenko-Beam-Bond-Model (TBBM) which has the most potential for development (Brown et al. 2014).

DEM is widely used to investigate bulk agricultural materials. The micromechanical parameters of a sunflower DEM model were calibrated based on oedometer tests so that the model can sufficiently approach the macro mechanical behaviour of the real bulk material (Keppler et al. 2011). In another study the effect of particle shape on flow was investigated by discrete element method simulation of a rotary batch seed coater (Pasha et al. 2016). The soil-tool interaction and the relations in cohesive soil were examined by using the DEM (Tamás et al. 2013).

In connection with fibrous agricultural materials fewer analyses can be found. Reduced cutting speeds; energy efficiency and cutting quality were investigated during the interactions between grass stalk and the rotation mower by DEM (Kemper et al. 2014). The removal and separation process of grapes from the stems in a destemming machine were examined by discrete element method (Coetzee and Lombard 2013). Several possible structural stalk models (chain of spheres, enhanced chain of spheres, hollow structure and solid body) for fibrous agricultural material were defined by DEM (Jünemann et al. 2013). A special solid geometrical structure of discrete element method (DEM) was analysed for corn stalks in quantitative and qualitative ways (Kovács et al. 2015).

Discrete element model application could reinforce the agricultural decision support system in countless areas: to estimate the uncollected mass and physical properties of chop for better soil mineral supplement, to prognosticate the quality of forage for more effective feeding, to validate yield maps based on simulation, to increase yield and to reduce fuel consumption during agricultural processes, etc.

Consequently, the long term objective of the study is the development of an adequate DEM model which can be adapted in the actual decision support system (DSS) to create a new model driven decision support system

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

(MDDS) to refine operational parameters. In this study a preliminary step of our study is discussed: the predicted role of DEM modelling in decision making and a possible numerical model for corn stalk.

THEORY AND BACKGROUND

Based on conventional DSS the new MDDS has been designed to get optimal machine parameters from real time and pre-calculated data acquired from a crop parameter database. The information dataset is based on in-situ and laboratory measurement data. The new method utilizes advantageously the latest technological, scientific results to be able to solve optimal adjustment problems of operational parameters of crop processing.

The new method is constituted of eight different stages which are represented on Figure 1: occurring of an optimal adjustment problem, database formation, storage of database, real time data collection, data processing, data digitalizing, decision making and act. The first step of the process is confronted by an optimal adjustment problem; an illustration of this is the definition of the optimal working parameters during an agricultural or a food production process. In parallel with that, the database formation begins. The database formation process results in a model-based crop parameter database that has two types of data: quantitative and qualitative. The next step is real time data collection in which the required in-situ input data about the investigated material, process and the actual circumstances are collected for the decision support system. Following this, in the data processing stage the collected data are transformed into the necessary form by mathematical and statistical methods. An illustration of this is when the measured densities of the specimens are recalculated to the numerical density of the digital representation of the specimens. After that, in the data digitalizing step the measured analogue data are digitalized in the numerical model. Then, the decision making is reinforced by evaluation of the digitalized in-situ data and simulation results. Based on the decision the necessary act could be accomplished and the decision making process can start again.

Considering the importance of maize in the global agriculture and the latest numerical methods from the field of agricultural simulations the new method was adapted to maize cultivation (especially for the processing of harvest-ready maize) and the discrete element numerical method (DEM) has been chosen.

MODEL FORMATION

Discrete element method (DEM) is developed to investigate bulk materials which contain separate parts. The definition of a DEM model is the following (Cundall and Hart 1993): it contains separated, discrete particles which have independent degrees of freedom and the model can simulate the finite rotations and translations, connections can break and new connections can come about in the model.

The previously described database formation stage of the model driven decision making process was separated into sub-stages, as shown on Figure 2. In this system the parameter measurement step is made up of three parts: physical properties, mechanical properties of the selected material and technical properties of the selected technological process. Following this, the collected data from laboratorial and in-situ tests were converted in the data processing stage. The physical properties of the material have to be calculated and integrated into the geometrical model of the numerical method, for instance in case of the discrete element method the physical characteristic of the material has to be divided and integrated into discrete particles. The mechanical properties were calculated during a calibration process in which the relationship among the required mechanical properties of the numerical method and the measured mechanical properties were investigated. In case of the discrete element method, the relationship among the micromechanical parameters of the contact model and macro mechanical parameters of the maize specimens were analysed. During the model formation step of data processing the measured technical properties were simplified and transformed into the system of the numerical model. Naturally, during data processing the relationships and coherence among the different input and output parameters were studied.

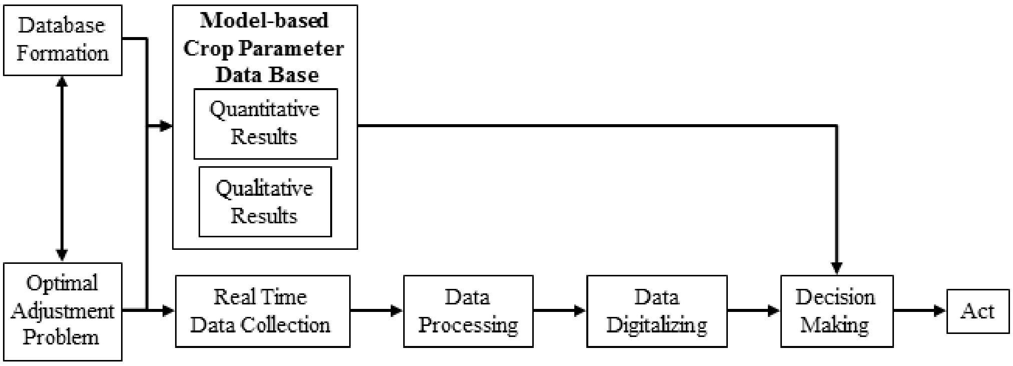


Figure 1: The model driven decision support system (MDDS)

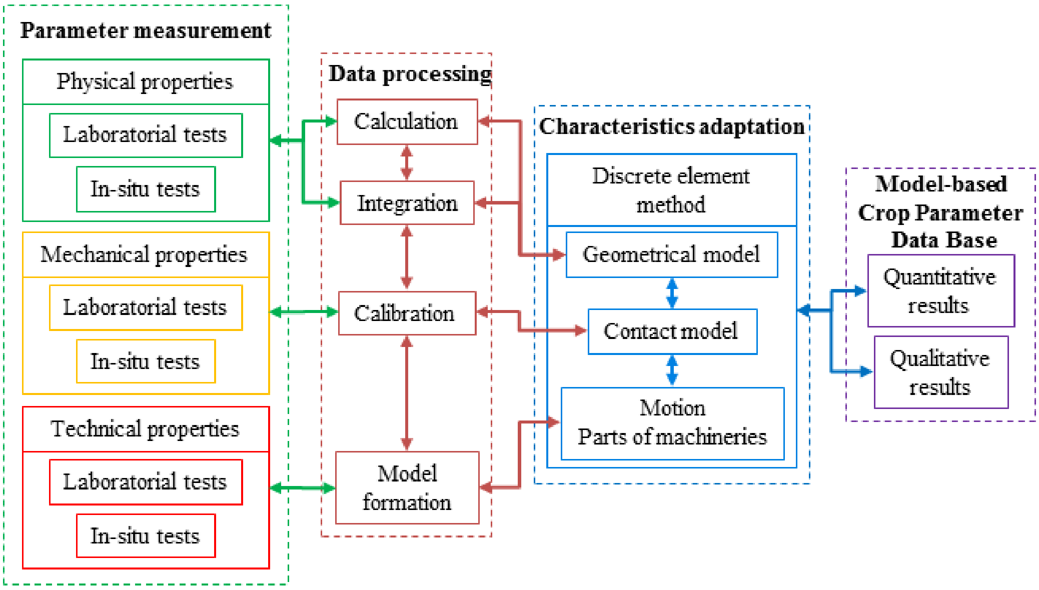


Figure 2: Detailed process of database formation for maize with discrete element method

In the stage of characteristics adaptation the analogue input data were converted into digital data by numerical modelling. The coactions of the different model parameters are one of the crucial factors during the model formation so this phenomenon has to be analyzed during the process. The numerical model provides quantitative and image base qualitative results as well that were investigated during the decision making step. Based on the quantitative and image base qualitative model-based crop parameter database, the sub-parts in the decision making process could extrapolate the optimal adjustment for the problem.

MATERIALS AND METHODS

Our study focuses on the database formation stage of the model-driven decision support system. In the stage of parameter measurement, based on harvest and product processes of harvest-ready maize the main physiological, physical, mechanical and technical parameters were determined. For all tests, specimens were collected from the same field from the middle region of Hungary (GPS coordinates: N 47.743692; E 19.613025).

The laboratorial tests were conducted in July, August, September and October, directly before harvesting. The majority of the in-situ tests were conducted simultaneously with the harvest so as to measure the actual characteristic of the processed material. In our study, root and leaves of the plant were neglected so these parts were removed from the maize stalks before the measurements.

In July and August our study particularly focused on the physical and mechanical properties of the maize stalk. To define the physiological and physical parameters and changes of the stalk mass, moisture content, length, diameter and shape of nodes and internodes of the stalk were measured and taken down. To define the mechanical properties and behaviour of the stalk two-point, three-point and four-point flexural, sideward compression, dynamic cutting and tension tests were conducted.

The results of the measures and observations weren't directly usable for the modelling method so suitable data and graphs were calculated with mathematical and statistical methods for the numerical modelling during the data processing stage.

Based on our previous study a hollow DEM geometrical structure with 18 particles was chosen for this analysis (Kovács and Kerényi 2015). This geometrical model ensures detailed analysis with lower computational costs. Thanks to the natural diversity of physical and mechanical properties of agricultural materials a stochastic variation with value 0.6, as in the natural structures, should be used during the numerical simulations (Kovács and Kerényi 2016). For the more realistic simulations the real parts and motions of the machineries were simulated. The calibration was an iterative process between the data processing and the characteristic adaptation stage where the different sets of micromechanical parameters for TBBM were tested to find the right adjusting of the numerical mechanical model.

The simulation results were evaluated by quantitative and image-based qualitative ways in order to acquire the necessary numerical data for making decisions; for instance to change the working parameters, to delay harvest etc. for the optimal maize processing.

RESULTS

The above-mentioned model-driven decision support system was tested on the physiological, physical and on the mechanical parameters through three-point flexural, sideward compression and dynamic cutting tests of the fourth internode of a maize stalk.

Digitalizing the physiological and physical parameters

Based on the evaluated physiological and physical parameters of the real stalk the digital geometrical model was formed in order to digitalize the real parameters of the stalk. During the model formation special features of the stalk (groove of the internode, surface deficiencies) were simplified and circular cross-sections were used instead of ellipsoid. The numerical geometrical model is constituted of discrete element particles that represent the physiological and physical parameters (size, shape, density) in the defined volume, as shown on Figure 3.

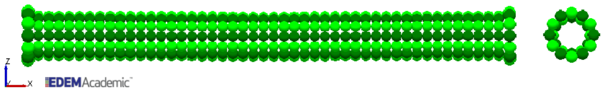


Figure 3: Digitalizing of physiological and physical parameters by the DEM geometrical model

Digitalizing the mechanical parameters and behavior

During the calibration process, the relationship between the micromechanical parameters of the DEM model and the measured macro-mechanical parameters of the real stalk were analysed based on the three-point flexural test. To represent the transversely isotropic mechanical behaviour of the real stalk a special contact system was defined among the particles. This system was developed based on our experiments and the features of the applied DEM software (EDEM 2.7, DEM Solutions Ltd.). In the hollow geometrical structure, the mechanical properties of the internode were modeled by the connections among the same type of particles (P4:P4; P5:P5) in axial direction, and were modeled by the connections among the even-numbered and the odd-numbered type of particles (P4:P5) in tangential direction. The mechanical properties among the particles of the node and the internode were defined by P1:P5; P2:P4; P7:P5; P6:P4 in axial direction and the mechanical properties of the node were defined by the P1:P2; P6:P7 connections in tangential direction, shown as on Figure 4.

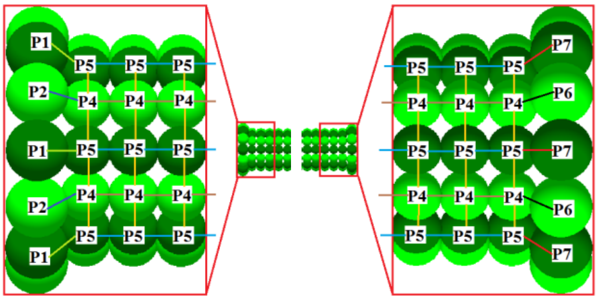


Figure 4: Digitalizing of the mechanical parameters and behaviour by DEM contact model

Quantitative results

The model was evaluated with quantitative and image-based qualitative methods. During the quantitative method the real measure diagrams, from the three-point flexural test, sideward compression and dynamic cutting were compared with the simulation diagrams.

The result from the three-point flexural simulation test is shown on Figure 5. The sudden changes of the diagram came from the DEM geometry, because the sudden connection breaks and sudden movements of the particles resulted in the observed changes on the chart. The simulated results are nearly between the minimal and maximal results of the measure. The linear section and the contraction section of the simulation data approached the measured data very well. In the plastic joint section the bonds among the particles are broken continuously, therefore the simulated data fluctuate strongly and the mean values are beneath the measured value. Taking into consideration the elastic mechanical behaviour of the applied DEM bonded model (Timoshenko-Beam-Bond-Model) these results are acceptable. For more accurate results in the plastic joint section a new bonded is needed that could simulate plastic deformations.

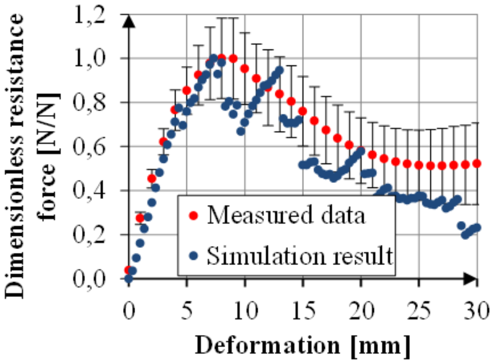


Figure 5: Simulated and measured characteristics of three-point flexural test

The simulation result of the sideward compression test is generally below the measured data of the real specimen, Figure 6. Up to the 35% deflection rate, the simulation characteristics corresponds to the measured data very well, but between the 35% and the 70% deflection rate, the results are far below the measure characteristics thanks to smaller size of the applied particle radiuses. With larger particle size a higher compression resistance force could be generated in the model. In the end of the simulated compression process a dramatic increase of the curve can be observed.

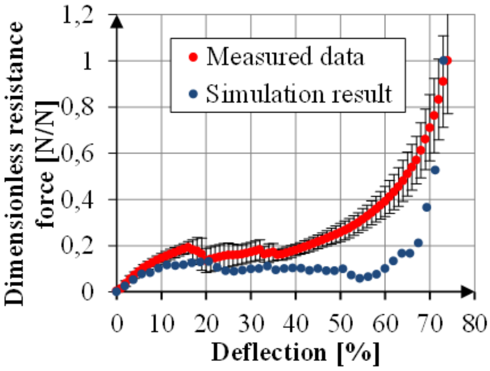


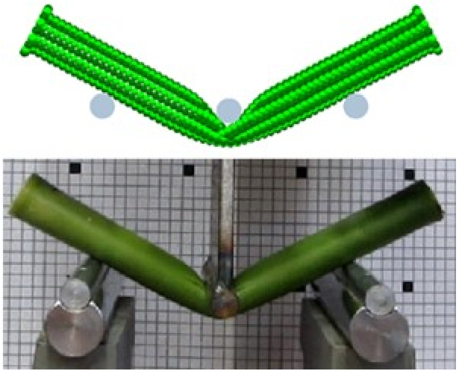
Figure 6: Simulated and measured characteristics of sideward compression test

The measured dynamic cutting work and the radius of the confidence interval (P=95%) were 21.92±6.8 J and simulated dynamic cutting work was 14.89 J, so the difference from the average measured results is about 32%, which is an acceptable difference if the high variation and biodiversity among the same species of agricultural crops is taken into consideration.

Image-based qualitative results

During the image-based qualitative evaluation cross-section deformations, crashes, breaks of the model were compared with the observed experiences of the real specimens.

In course of the image-based qualitative evaluation of simulation results of three-point flexural test, the shape of the modelled sample with the real specimen at the point of maximum deformation and the residual deformation of modelled sample with the real specimen was compared, see Figure 7. The bended cross-sections of the modelled specimens were crashed under the loading anvil just as the real specimen and the bended shapes of the modelled specimens were near the same as the real specimen. The residual shape of the specimen bended a little back (thanks to the more extended crashed zone) after the simulated bending process.



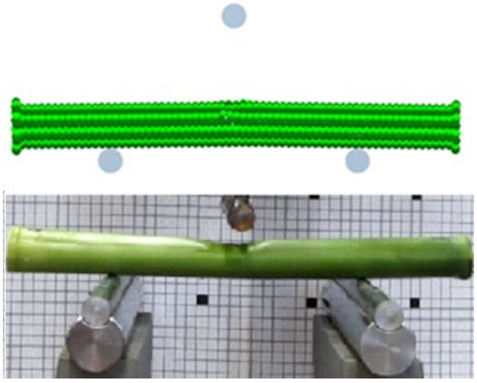


Figure 7: Simulated and real residual shape of specimens at the point of maximum deformation (above) and in the end of the three-point bending test (bottom)

In case of the qualitative evaluation of the sideward compression test the deformations and the damages of the cross-section were analysed, see Figure 8. In the first stage, the initial shape of the specimens can be observed; while in the second stage, elastic deformation took place until the first break appeared in vertical direction. During the third stage, another break appeared in horizontal direction; while in the fourth stage, that is the end of the compression process, the model mimics well the behaviour of the real specimen. In the last stage, compression clamps returned to the starting point and an elastic deformation was conducted on the real specimen and in the model specimen as well.

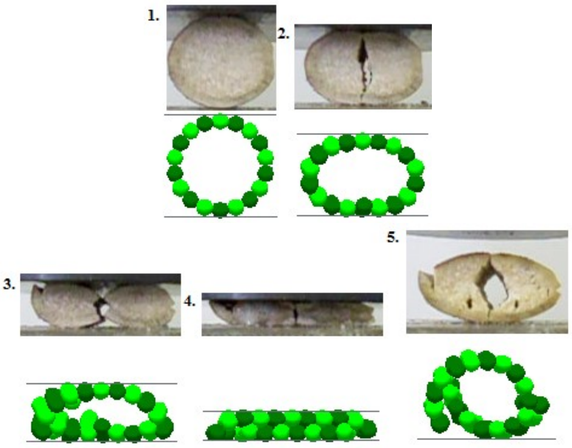


Figure 8: Simulated and real deformations and crashes of specimens during the sideward compression test

During the qualitative evaluation of the dynamic cutting test the surface of cut was analysed. In the model the cutting knife broke out the bottom part of the specimens, in turn, on the real specimen this phenomenon was not observed, the surface of cut was fully straight, see Figure 9.

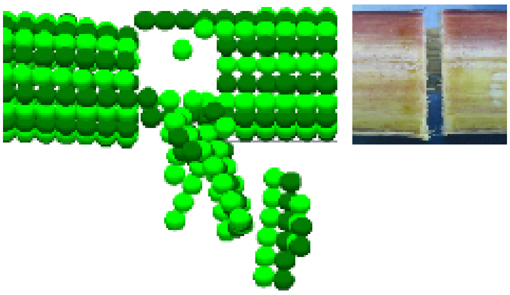


Figure 9: Simulated and real surfaces of dynamic cutting test

CONCLUSIONS

Our study focuses on the digitalizing of analogue data from laboratorial and in-situ measures of maize and for that process the discrete element method has been chosen. Laboratorial three-point flexural, compression, dynamic cutting tests and tests in connection with harvesting were conducted to define the main physiological, physical and mechanical parameters and behaviour of corn stalks and ears from mid-summer to the harvesting date.

During the model formation to select the right model that can best simulate the mechanical and physical properties and behaviour of the plant, the DEM simulations of the laboratorial and in-situ tests were conducted. With modifications of the geometry structure and the input parameters of the contact model, the right composition was found.

Based on the results the following conclusions could be formulated:

(1) Theoretical course of model driven decision support system is usable to analyse the relationship among the physical and mechanical parameters of maize and operational parameters of farm machineries during harvesting.

(2) The discrete element numerical models are capable of simulating the interactions among the parts of the machines and different parts of maize and compare the simulation and experimental results.

(3) The discrete element numerical models are usable to form a model-based crop parameter database.

In the future, the current results and models can be adapted to more detailed simulations on finding the optimal harvesting parameters of machineries.

REFERENCES

Brown N. J.; J. F. Chen; J. Y. Ooi. 2014. "A bond model for DEM simulation of cementitious materials and deformable structures." *Granular Matter*, No. (2014) 16, 299-311.

Chetty V.; N. Woodbury; S. Warnick. 2014. "Farming as feedback control." In *American Control Conference (ACC)* (Portland, Oregon, USA), 2688-2693.

Coetzee C. J.; S. G. Lombard. 2013. "The destemming of grapes: Experiments and discrete element modelling." *Biosystems Engineering*, No. 114(3), 232–248.

Cundall P. A.; R. D. Hart. 1993. "Numerical Modeling of Discontinua." *Analysis and Design Methods*, No. 1993 231-243.

Jünemann D., S. Kemper; L. Frerichs. 2013. "Simulation of stalks in agricultural processes - Applications of the Discrete Element Method." *Landtechnik*, No. 68(3), 164–167.

Keppler I.; L. Kocsis; I. Oldal; A. Csatár. 2011. "Determination of the discrete element model parameters of granular materials." *Hungarian Agricultural Engineering*, No. 23/2011, 30-32.

Kovács Á.; K. Kotrocz; GY. Kerényi. 2015. "The adaptability of discrete element method (DEM) in agricultural machine design." *Hungarian Agricultural Engineering*, No. 27/2015, 14-19.

Kovács Á.; GY. Kerényi. 2016. "Comparative analysis of different geometrical structures of discrete element method (DEM) for fibrous agricultural materials." In *4th CIGR International Conference of Agricultural Engineering* (Aarhus, Denmark, June 26-29), 1-8.

Kovács Á.; GY. Kerényi. 2016. "Stochastic variation in discrete element method (DEM) for agricultural simulations". *Hungarian Agricultural Engineering*, No. 30/2016, 31-38.

Pasha M.; C. Hare; M. Ghadiri; A. Gunadi; P. M. Piccione. 2016. "Effect of particle shape on flow in discrete element method simulation of a rotary batch seed coater." *Powder Technology*, No. 296, 29–36.

Tamás K.; I. J. Jóri; A. M. Mouazen. 2013. "Modelling soil–sweep interaction with discrete element method." *Soil and Tillage Research*, No. 134, 223–231.

AUTHOR BIOGRAPHIES

ÁDÁM KOVÁCS was born in Debrecen, Hungary and went to Budapest University of Technology and Economics, where he studied agricultural machine design and obtained his MSc. degree in 2016. Currently he is a Ph. D. student in the same institution and his topic is discrete element modeling of maize. He worked for the WIGNER Research Center for Physics at the Department of Plasma Physics for two years, where he designed diagnostic devices for fusion reactors. His e-mail address is: kovacs.adam@gt3.bme.hu and his Web-page can be found at http://gt3.bme.hu/en.

JÁNOS PÉTER RÁDICS is assistant professor at Budapest University of Technology and Economics where he received his MSc degree. He completed his PhD degree at Szent István University

Gödöllő. His main research is simulation of soil respiration after different tillage methods, and he also takes part in the DEM simulation research group of the department. His e-mail address is: radics.janos@gt3.bme.hu and his Web-page can be found at http://gt3.bme.hu/radics.

GYÖRGY KERÉNYI studied agricultural machine design at Szent István University, Gödöllő and after that he went to Budapest University of Technology and Economics, where he obtained his PhD degree in 1997. Currently he is an associate professor and deputy head of Department of Product and Machine Design in the same institution and his research topic is numerical methods in agricultural machine design. His e-mail address is: kerenyi.gyorgy@gt3.bme.hu and his Web-page can be found at http://gt3.bme.hu/en.

INTEGRATED OPTIMIZATION OF TRANSPORTATION AND SUPPLY CONCEPTS IN THE AUTOMOTIVE INDUSTRY

Corinna Maas
Institute for Materials Handling, Material Flow, Logistics
Technical University of Munich
Boltzmannstr. 15, 85748 Garching bei München, Germany
maas@fml.mw.tum.de

Andreas Tisch
Institute for Materials Handling, Material Flow, Logistics
Technical University of Munich
Boltzmannstr. 15, 85748 Garching bei München, Germany
andreas.tisch@mytum.de

Carsten Intra
Executive Board Production & Logistics, Research & Development
MAN Truck & Bus AG
Dachauer Str. 667, 80995 Munich, Germany
carsten.intra@man.eu

Johannes Fottner
Institute for Materials Handling, Material Flow, Logistics
Technical University of Munich
Boltzmannstr. 15, 85748 Garching bei München, Germany
fottner@fml.mw.tum.de

KEYWORDS

Inbound logistics; supply concept; transportation concepts; cost optimization; CO2 emissions

ABSTRACT

A growing cost pressure due to increasing transportation costs and a changing environment call for a flexible adaptability of the inbound logistics processes in the automotive industry. Therefore, these processes need to be continuously reviewed for efficiency potentials. This paper provides an optimization model that allows for the integrated cost assessment of supply and transportation concepts. Additionally, the idea of green logistics is addressed by including the costs for carbon dioxide (CO2) emissions in the optimization model. The model is applied to an industrial case of a commercial vehicle manufacturer. The results show that delivery frequencies and the consideration of the entire material flow of a transport relation are main influence factors. The integration of CO2 emissions shows that the emissions can be reduced while only slightly increasing logistics costs.

MOTIVATION

Globalization, expansion of new markets, and fast changing environments are core challenges in the automotive industry leading to increasing logistics costs (Göpfert, 2013). Inbound logistics, which is the link between suppliers and manufacturers, is particularly concerned by these developments. This is because the inbound logistics embraces transportation and the respective costs. These transportation costs encompass a high share of the total logistics costs (Bravo and Vidal, 2013). Besides increasing transportation costs, progressive environmental pollution is a prominent issue nowadays. Larger transport distances are one main driver for environmental pollution. Thus, a continuous improvement of the inbound processes is required. Consequently, inbound logistics should be configured cost efficiently and at low emissions using different transport and supply concepts. Literature shows that these objectives are mainly addressed exclusively. Hoen *et al.* (2014) integrate CO2 emissions in the selection process of transport concepts. This is also one of many approaches for the quantitative selection of transportation concepts. For supply concepts, very few quantitative approaches exist and focus rather on the in-house logistics costs than on the inbound costs (see e.g. Wagner and Silveira-Camargos, 2011). The available qualitative approaches in literature for selecting those concepts focus either on supply or transportation, and do not address costs or emissions. Supply concepts are often chosen based on the parts' characteristics, such as the value or fluctuations in consumption (see e.g. Wagner and Silveira-Camargos, 2011). A qualitative approach to selecting the transportation concept is a decision tree using different criteria, such as delivery frequency or supplier location (see e.g. VDA-5010). To meet the gap in research we have provided an approach that combines the selection of transportation and supply concepts into one quantitative model based on costs and emissions.

MODEL STRUCTURE AND ELEMENTS

Model Framework

Inbound logistics is defined as "Activities associated with receiving, storing, and disseminating inputs to the product, such as material handling, warehousing, inventory control, vehicle scheduling, and returns to suppliers" (Porter, 2004, 39f.). To conduct the described activities, different concepts for each activity are needed, such as transportation concepts and supply concepts (see Figure 1).

A supply concept defines the configuration of the logistics process from the supplier to the manufacturer. Supply concepts include direct delivery and in-stock delivery. The latter implies that there is at least one warehousing stage included in the process. Direct delivery concepts embrace Just-in-Time (JIT) and Just-in-

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Sequence (JIS) (VDA-5010). JIT means the delivery of homogenous parts just in time they are needed at the assembly line. JIS additionally orders the parts corresponding to the production sequence (Wagner and Silveira-Camargos, 2011).

A transportation concept includes the description of the transport process and the necessary logistics service. We distinguish between three different concepts: Direct relation, milk run, and hub and spoke. Direct relation is a single-stage transport chain. Thus, the material is delivered directly from the supplier to the recipient. A milk run is used for smaller delivery volumes, because partial loads from different suppliers are consolidated into a full load. The hub and spoke concept contains three stages: In the pre-run the goods are transported to a hub; in the hub the goods are re-sorted; and in the main-run the re-sorted goods are directly delivered to the recipient. In the automotive industry, hub and spoke concepts are often implemented using area forwarding. The area forwarder is instructed to collect all part loads from a certain supplier area, to combine these loads and to forward them to the OEM. How the area forwarder is exactly combining the part loads and which tours are taken is not transparent to the OEM. Therefore area forwarding represents a black box for the OEM in terms of operating processes. The difference between hub and spoke and area forwarding lies in the occurring expenses and the process transparency (Schulte, 2009).

For a consistent understanding, we briefly explain the used terminologies and the considered inbound process: The combination of a part and a charge carrier is defined as a shipping unit. A packing batch is the number of parts the respective charge carrier contains. Thereby the volumetric weight can be determined, which is needed to calculate the freight rates. A shipping unit is transported from the supplier to the manufacturer by means of transport using a transportation concept. A transport relation is defined by the combination of a supplier and the receiving manufacturer plant. The parameters of the material planning, such as lead times, safety stocks, etc., represent restrictions that need to be met. Each part is delivered by a combination of supply and transportation concepts described above. This combination determines the subsequent in-house processes.

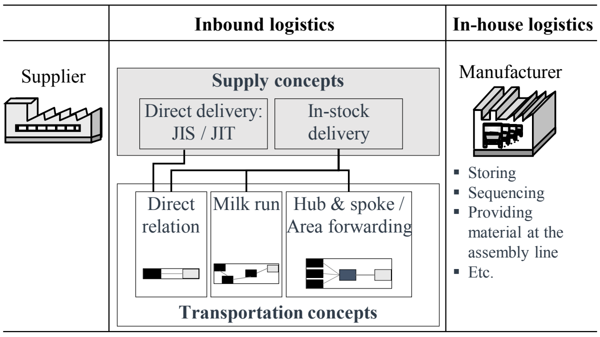


Figure 1: Combinations of Transportation and Supply Concepts

Optimization Approach

The objective of this work is to identify saving potentials within the inbound process by selecting appropriate supply and transportation concepts. Additionally, CO2 emissions should be minimized. We only consider combinations of the most common supply and transportation concepts described before (see Figure 1). A JIS or JIT delivery only works with the transportation concept direct relation. In-stock delivery, however, is combinable with each of the four transportation concepts (hub and spoke and area forwarding are two different concepts in terms of occurring costs). Thus, there are three main optimization scenarios possible: First, the switch from direct delivery to in-stock delivery and vice versa. To compare both supply concepts adequately, in-stock delivery is always combined with the cost optimal transportation concept. The second scenario addresses the optimization of the transportation concept within in-stock delivery. The third scenario refers to the choice of the optimal delivery frequencies for area forwarding and hub and spoke.

To identify the most economic inbound concept, the described model framework must be transferred into a cost model. The main challenge is assigning the arising expenses to the different logistics activities. One appropriate approach addressing this issue is activity-based costing. The approach deals with the assignment of overhead costs to upstream and downstream activities from production, such as storing, transportation, etc. (Weber, 2012). Hence, the entire process described in the model framework is modelled by the relevant cost types. The selection of the cost types is geared toward the material flow process (Wagner and Silveira-Camargos, 2011). Although we focus on inbound concepts, in-house logistics is still taken into account. This is because the subsequent processes from inbound logistics have an impact on the cost calculation. We include all costs that are crucial to distinguish between the different concepts. Furthermore, depending on each logistics concept, the derived cost functions differ (Wagner and Silveira-Camargos, 2011). Table 1 displays these different compositions and the considered cost types.

Cost Types

In the following, we explain the different cost types listed in Table 1. The sum of these costs we call inbound logistics costs. The transportation costs depend on the used transportation concept. For all transport concepts, one cost part is defined by the full load. It is calculated by the transportation cost rate multiplied by the distance to the supplier and the number of transports. The transportation cost rate, however, varies between the concepts, due to different tariff arrangements. To calculate the costs for a milk run, we add stopping costs. The distance is based on the defined tour of the milk run. For area forwarding, the transportation costs are either calculated by full load or partial load. This depends on the volumetric weight of the respective load. The partial load is calculated by multiplying the weight of a shipping unit, the cost rate per kilogram and the number of shipping units. The cost rates per kilogram depend on the distance to the supplier and the transported kilograms per delivery, which leads to a highly complex problem description. The transportation costs for hub and spoke consist of the pre-run and the main-run. For the pre-run, the same costs as for the area forwarding are assumed. The main-run can either be calculated like the partial load or the full load (i.e. direct collection). Empty container deliveries to the supplier are executed through the network of main supplier plants. To estimate the costs for the return of empties, inbound transportation costs are multiplied by the convertible ratio of the containers.

The inventory and storage costs can both be divided into fix and variable costs. In this case, fix costs depend on the safety time and variable costs on the daily inventory. The inventory fix costs are calculated by multiplying the price of a shipping unit by the interest rate per day, the number of days the shipping units are in circulation, and the number of shipping units. The variable inventory costs are defined by the price of the shipping unit multiplied by the interest rate per day and the daily inventory. The daily inventory arises from the delivery and the occurring demands. Note that only the inventories of the manufacturer are included, that is transit and in-house inventories. The storage costs are only relevant for in-stock deliveries. The fix storage costs are calculated by the price for warehouse space, the space of a charge carrier adapted by the stacking ratio, the safety time, the costs for storing and un-storing, as well as the number of shipping units. For the variable storage costs, the daily cost rate based on the space utilization is multiplied by the daily inventory.

The miscellaneous costs include trailer and container rental fees, sequencing costs, and costs for service providers. Trailer rental fees depend on two aspects: The buffer time and the trailer range. The buffer time is the time a trailer spends on average on the trailer yard which is coupled with the number of transports. The trailer range is measured in days and is determined by the trailer content (i.e. the number of shipping units) and the consumption of the respective part. The trailer rental fees only incur for direct deliveries. For the container rental fees, we calculate the cost rate for a charge carrier multiplied by the time the charge carriers are in circulation and the number of shipping units. The circulation time is two times the transit time from supplier to manufacturer plus the time the charge carrier is standing at the supplier's plant. Sequencing costs are composed by the sequencing price per part, the batch size, and the number of shipping units. For in-stock delivery it has to be checked whether the parts need to be provided homogeneously or sequentially at the assembly line. The costs for service providers occur in in-house logistics when a trailer has to be moved from the trailer yard to the dock and vice versa. They depend on the number of internal transports (counted twice for forwarding and returning) and the cost rate of the service provider. These costs only occur for direct deliveries.

Table 1: Inbound Concepts with Assigned Cost Types

	JIT	JIS	In-stock delivery			
	Direct relation	Direct relation	Direct relation	Milk run	Hub and spoke	Area forwarding
Transportation costs						
Full load (direct collection)	X	X	X	X	(X)	X
Partial load						X
Pre-run					X	
Main-run					X	
Stopping costs				X		
Empty returns	X	X	X	X	X	X
Inventory and storage costs						
Inventory (fix)	X	X	X	X	X	X
Inventory (variable)	X	X	X	X	X	X
Storage (fix)			X	X	X	X
Storage (variable)			X	X	X	X
Miscellaneous costs						
Trailer rental fees	X	X				
Container rental fees	X	X	X	X	X	X
Sequencing costs		X	(X)	(X)	(X)	(X)
Service provider costs	X	X				

X Cost type is relevant for this inbound concept
(X) Cost type is not always relevant for this inbound concept

CO2 Emissions

The calculation of CO2 emissions follows the consumption-based approach according to the European Standard EN 16258, i.e. a "well-to-wheel" approach. Greenhouse gas emissions are expressed by CO2 equivalents (CO2e), which is a standardized "measurement against which the impacts of releasing (or avoiding the release of) different greenhouse gases can be evaluated" (Kontovas and Psaraftis, 2016, p. 45). In the calculation, one can distinguish between an empty run and a full run. The emissions of an empty run are constant per kilometer because they depend on the tare weight of the freight vehicle. We assume a linear development of the emissions from an empty to a full run. To determine the CO2e emissions realistically, the calculation is done separately per transportation concept. For area forwarding we assume an average capacity utilization of the transports for each area to calculate the emissions.

To include CO2e emissions in the cost optimization, we multiply the CO2e emissions by a cost rate (Euro per ton of CO2e emission). Since there are no taxes on CO2 emissions in Germany so far, we assumed the highest tax rate for transport fuels within Europe, which is the tax rate of Finland with US$ 66 per ton CO2e (World Bank Group and ECOFYS, 2016). Including CO2e emissions in the optimization model is optional. Nevertheless, CO2e emissions are included for two reasons: First, to demonstrate the consideration of a non-monetary objective, and second to comprise green logistics due to its importance.

OPTIMIZATION MODEL

Requirements, Assumptions, Boundaries

By analyzing the different cost and CO2e emission functions, it became obvious that there are three main drivers: The number of executed transports, the number of shipping units, and the sum of daily stock. In a wider sense, the trailer rental fees can be regarded as inventory costs for direct deliveries and the costs for service providers can be considered as an extended transportation within the manufacturer's plant. Hence, the objective dimensions of the optimization model consist of inventory and storage costs, transportation costs, and CO2e emissions. The objective is to reduce each value, taking into account the correlated trade-offs.

Furthermore, we want to point out some dependencies: The trailer rental fee, which depends on the trailer range, cannot be assigned to only one specific cost driver. It is rather a combination of the number of transports and the respective transported quantities interacting with the predominant demands. Area forwarding and hub and spoke do not apply a fix cost rate, but costs depend on how much is shipped in each run (i.e. the transportation costs for each transport of a transport relation can differ). Furthermore, the stocks depend on the time-based and quantity-based shipments. Consequently, the inbound costs per supply and transportation concept do not rely on a single shipment. Thus, a multi-period consideration is necessary.

Additionally, we assume the following: All process flows behave ideally; the network of suppliers is given; a shipping unit is the smallest indivisible unit; a year is defined by 48 weeks and 5 working days; the used freight vehicle corresponds to a mega trailer.

The following aspects are not part of the optimization: Emergency concepts or extra tours; quality issues; a lack of delivery reliability; network optimization; operative control; upfront investments allowing for the use of alternative inbound concepts; the definition of possible milk run tours. To include milk runs in the optimization, each tour must be defined separately in advance.

Minimization Problem

The objective of the model is the identification of optimization potentials within the inbound logistics process based on a monetary valuation. The developed cost accounting shows that two aspects are mainly relevant for determining the costs: Firstly, the chosen type of inbound concept and, secondly, the order quantity. The order quantity is not only crucial for the respective transportation tariff, but also indirectly for the inventories. Hence, the model to be developed corresponds to a lot sizing problem. In literature, several approaches exist that address this decision. In this paper, we apply a mathematical optimization model that minimizes the logistics costs. We used a decomposition approach to solve the problem efficiently. First, we optimize each combination of supply and transportation concept separately. The objective function includes all cost types that depend on the order quantity. This excludes fix storage and inventory costs, as well as container rental fees and sequencing costs. Fix storage and inventory costs and sequencing costs depend on the process and occur per shipping unit. Container rental fees depend on the frequency for return of empties – figuratively and according to the assumptions, these costs are also process-related. Second, we compare the logistics costs of the different partial solutions (including all cost types) and choose the most cost-efficient concept combination.

Before formulating the optimization problem, we want to stress that the staggered transportation tariffs lead to a non-linear problem. Since linear optimization problems are easier to solve, linearity is defined as a requirement. To obtain a linear model, the decision model is adapted to the determination of the order quantity at a certain point in time for each staggered tariff. For a better understanding of the mathematical formulation, see the notation overview in Table 2.

Table 2: Notation Overview

Decision variables			
$q_{T,i,k}$	Number of transports in period *i* using tariff *k*		
$q_{U,i,k}$	Number of shipping units in period *i* using tariff *k*		
Specific weights for the different supply and transportation concepts			
λ_S	Weighting by which the sum of the daily inventory is included in the objective function; varies per supply and transportation concept		
$\lambda_{T,i,k}$	Weighting by which the number of transports in period *i* with tariff *k* is included in the objective function; varies per supply and transportation concept		
$\lambda_{U,i,k}$	Weighting by which the number of shipping units in period *i* with tariff *k* is included in the objective function; varies per supply and transportation concept		
Parameters			
$d_{U,i}$	Demand for a shipping unit in period *i*		
lb_k	Lower bound of a tariff *k* (in volumetric weight)		
ub_k	Upper bound of a tariff *k* (in volumetric weight)		
$p_{T,k}$	Transport cost rate for a transport (including stopping costs) using the corresponding transport tariff *k*		
$q_{MIN,U}$	Minimum order quantity of a shipping unit		
$q_{SS,U}$	Safety stock of a shipping unit		
$v_{U,k}$	The value of a shipping unit that is used as assessment basis for the transport tariff *k* (in volumetric weight)		
w_{FV}	Load capacity of a freight vehicle (in volumetric weight)		
w_U	Volumetric weight of a shipping unit		
Indices and abbreviations			
i	Index for the period	*S*	Inventory / stock
j	Indexvariable iterating the periods	*SS*	Safety stock
k	Index for the transport tariff	*T*	Transport
FV	Freight vehicle	*U*	Shipping unit
MIN	Minimum		

The objective function is displayed in formula (1). We call it a general model for the case of a single part. The objective function holds for each combination of the inbound concepts. We differentiate between the alternative concept combinations, compounded by the cost types, with the three weights λ_S, $\lambda_{T,i,k}$ and $\lambda_{U,i,k}$.

$$\min_{q_{T,i,k};q_{U,i,k}} \sum_{i\in\mathbb{R}^I}\sum_{k\in\mathbb{R}^K} \lambda_{T,i,k}\cdot q_{T,i,k} + \lambda_S \cdot \sum_{i\in\mathbb{R}^I}\left[\sum_{j=1}^{i}\left[\sum_{k\in\mathbb{R}^K} q_{U,j,k} - d_{U,j}\right] + \frac{1}{2}\cdot d_{U,i}\right] + \sum_{i\in\mathbb{R}^I}\sum_{k\in\mathbb{R}^K} \lambda_{U,i,k}\cdot q_{U,i,k} \qquad (1)$$

Subject to the constraints:

$$\sum_{j=1}^{i}\left(\sum_{k\in\mathbb{R}^K} q_{U,j,k} - d_{U,j}\right) \geq q_{SS,U} \qquad \forall\, i \in \mathbb{R}^I \tag{2}$$

$$q_{U,i,k} \leq \left\lfloor \frac{ub_k}{v_{U,k}} \right\rfloor \cdot q_{T,i,k} \qquad \forall\, i \in \mathbb{R}^I, k \in \mathbb{R}^K \tag{3}$$

$$q_{U,i,k} \geq \left\lceil \frac{lb_k}{v_{U,k}} \right\rceil \cdot q_{T,i,k} \qquad \forall\, i \in \mathbb{R}^I, k \in \mathbb{R}^K \tag{4}$$

$$\sum_{k\in\mathbb{R}^K} q_{U,i,k} > \left\lfloor \frac{w_{FV}}{w_U} \right\rfloor \cdot \left(\sum_{k\in\mathbb{R}^K} q_{T,i,k} - 1\right) \qquad \forall\, i \in \mathbb{R}^I \tag{5}$$

$$\sum_{k\in\mathbb{R}^K} q_{U,i,k} \geq q_{MIN,U} \qquad \forall\, i \in \mathbb{R}^I \tag{6}$$

$$q_{U,i,k} \geq 0 \qquad \forall\, i \in \mathbb{R}^I, k \in \mathbb{R}^K \tag{7}$$

$$q_{T,i,k} \geq 0 \qquad \forall\, i \in \mathbb{R}^I, k \in \mathbb{R}^K \tag{8}$$

$$q_{U,i,k} \in \mathbb{Z} \qquad \forall\, i \in \mathbb{R}^I, k \in \mathbb{R}^K \tag{9}$$

$$q_{T,i,k} \in \mathbb{Z} \qquad \forall\, i \in \mathbb{R}^I, k \in \mathbb{R}^K \tag{10}$$

Constraint (2) ensures that the inventory does not fall below the safety stock. Constraints (3) and (4) guarantee that the specific tariff is only applied when the upper or lower bound is not exceeded or undercut, respectively. The boundaries are specified per area and OEM individually. For a case study these boundaries are provided by the OEM. The available tariffs and the corresponding boundaries are defined in a way that the maximum capacity utilization of the freight vehicle is implicitly included. Constraint (5) ensures that for each period a new transport is only triggered when the previous transport is completely full. Constraint (6) states that each order quantity of every period has to correspond to at least the minimum order quantity. Constraints (7) to (10) describe the non-negativity and integer conditions.

Model Extensions

For the application to an industrial case, some adaptions need to be made. These adaptions are implemented by adding further constraints and additional decision variables. Thereby the model's complexity is increased, but still the linear character of the model has not changed.

So far the model does not include restrictions for warehouse capacities. This could lead to the fact that the model accepts high inventories in favor of transportation costs savings. In practice, the data of warehouse capacities are often not available or at least difficult to obtain. Hence, we decided to include delivery frequencies in the model that are often used to control order quantities and inventories. To implement delivery frequencies, additional restrictions are added to the model. These restrictions are the following: The delivered quantity should not exceed the sum of demands until the next delivery point. The maximum quantity of a delivery of one week should not exceed the sum of demands of this week. A delivery is only permitted if the delivery frequency admits it. And the delivery frequencies for the considered period are equal for each week. The mathematical formulation of these restrictions goes beyond the scope of this paper. Furthermore, for direct deliveries, a minimum capacity utilization of each transport is defined, since there are no pre-defined numbers of deliveries per period for those supply concepts. Thereby, we eliminate the case when a full load is exceeded by only one shipping unit and thus an additional full load is billed because of this single shipping unit.

So far, the model considers the case of a single part on a transport relation. In reality though, the material flow of an entire transport relation is considered, i.e. all parts that a supplier delivers to the manufacturer. Thereby, more reliable cost statements can be drawn due to the weight staggered tariff systems. To integrate all parts of a transport relation, we suggest combining them into one reference shipping unit. This is done by acquiring the weighted averages of the part's characteristics for all parts of the transport relation. The reference shipping units are formed while pre-processing the available data. Moreover, safety stocks are set to zero. The process flows are ideal and thus the safety stock is never permitted to undercut (although in reality the safety stock serves in emergency cases). The order quantity decision does not depend on safety stocks. Delivery times and call-off orders are not explicitly regarded, since these are parameters for the operative planning. The presented model addresses a more tactical level, striving for a precise statement about the inbound logistics costs for each inbound concept combination – i.e. not solely average cost calculations. Parts with an order quantity smaller than the minimum order quantity, are monetarily not interesting and negligible. Therefore minimum order quantities are also set to zero.

The optimization problem is formulated in Microsoft Excel. The individual optimization problems are solved using the Gurobi Optimizer 6.5.1 via the open source plug-in "OpenSolver" (OpenSolver, 2017).

APPLICATION TO INDUSTRIAL PRACTICE

Case Description

We consider the case of a commercial vehicle manufacturer: The company decided to relocate the production of bus chassis from the initial plant to either plant A or B. Note that the data presented in this paper is completely anonymized. It is merely used to demonstrate the applicability of the model. From the company's perspective, the alternative plant A was favored.

The analysis uses the following data structure: All parts of the relevant production portfolio of the bus chassis are considered. The data basis is a production period of one month extrapolated to one year. It is assumed that the network of suppliers of those parts remains the same after relocating the production, as contracting new suppliers may take some time. Inter-plant transports are not included, i.e. deliveries of components from other production plants of the company. The analysis covers 452 in-stock delivered parts (i.e. the supply concept of the initial plant) from 142 different suppliers. Furthermore, uniformly distributed demands are assumed due to the data basis of one month. Fluctuating demands are generally applicable to the model, but only reasonable when considering a sufficiently large number of periods. The objective of the model is to reveal optimization potential instead of supporting operative control.

Case Results

The objective of the case study is twofold: First, the optimization model should be applied to a real practical problem. Second, one of the three optimization scenarios should be exemplified. To solve the relocation problem, the inbound logistics costs for all three plants need to be compared. For the cost calculation the most efficient combination of supply and transportation concept for each part and transport relation is used. Due to little quantities that are procured from each supplier, direct delivery is for none of the plants preferred over in-stock delivery. The first optimization scenario (i.e. the choice between supply concepts) is therefore neglected. The inbound logistics costs are always calculated with the optimal transportation concept for each reference shipping unit. Thus, the second optimization scenario is applied, but will not be discussed in detail for the different parts. Instead, this chapter stresses four aspects: The effect of reference shipping units, the effect of the optimization of delivery frequencies (i.e. the third optimization scenario), and the effect of included CO2e emissions in the optimization. The results of the first three analyses are illustrated in Figure 2. Since all parts are delivered in-stock, cost types related to direct delivery are not displayed in the results.

In analysis 1, the overall inbound logistics costs were calculated for all three plants. Reference shipping units are built per supplier over all 452 parts (i.e. only bus chassis parts). The transportation costs have the highest share with an average of 49%. Storage and inventory costs only have a share of 17% on average. For the initial production plant, the transportation costs' share was the lowest at 44%, which can be explained by smaller distances to the suppliers. From an inbound costs perspective, the company's tendency favoring plant A over B can be supported.

The second analysis focuses on the effect of reference shipping units while comparing only the initial plant and plant A. Here the reference shipping units are extended by the entire material flow of each supplier (i.e. all parts of the suppliers and not only bus chassis parts are considered). The results in Figure 2 show the costs of the initial plant and plant A with reference shipping units considering either only bus chassis parts (1) or all parts of the suppliers (2). The costs difference of -1.9% of plant A is not as high as the cost difference of -21.8% of the initial plant. This can be explained by the fact that the suppliers are mainly new suppliers for plant A. In contrast, the initial plant can achieve synergies through higher transport volumes because the same suppliers deliver parts for other components. The transportation costs therefore even decrease by 28%. These numbers show the importance of sourcing suppliers in accordance with the production network.

In the third analysis, the effect of optimized delivery frequencies for plant A is examined. The upper bar shows the logistics costs when using the initial delivery frequencies. The lower bar displays the logistics costs with optimized delivery frequencies. The reference shipping units consider all parts of the suppliers. The optimized delivery frequencies result in 10.6% less logistics costs. 71% of these cost savings can be drawn from only ten of the considered suppliers. Three suppliers still comprise 37% of the cost savings. Note that the adaption of the delivery frequencies is not readily possible in practice, as smoothing effects for incoming goods and demand fluctuations also need be considered.

The last analysis focuses on the CO2e emissions. We run the optimization model for plant A once including CO2e emission in the objective function and then without emissions. We found that the CO2e emissions could be reduced by 1.14% per year, whereas logistics costs only increased by 0.03%. Pre-studies of other transport relations had shown that there is a lot of potential in reducing CO2e emissions while logistics costs increase hardly noticeably.

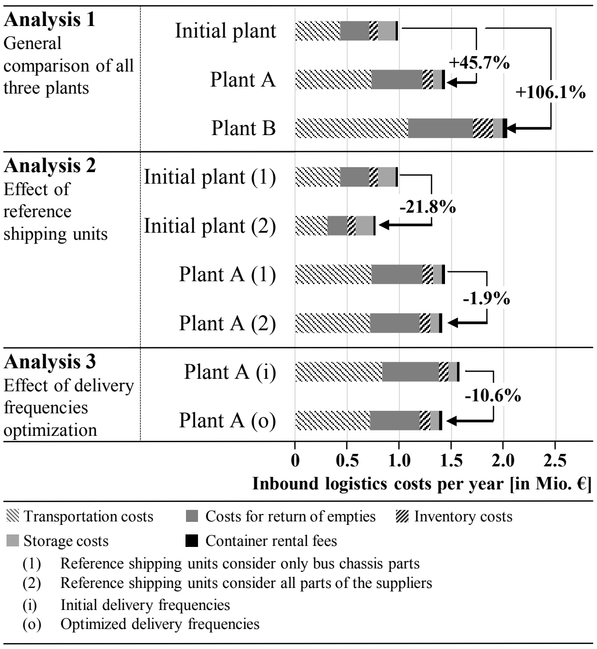


Figure 2: Results from Different Analyses

CONCLUSION AND OUTLOOK

This paper proposes an optimization model for evaluating the most efficient combination of supply and transportation concepts. Additionally, the idea of green logistics is included by adding CO2e emissions to the objective dimensions. The objective function of the model is cost oriented. We applied activity-based costing for the different processes to model the inbound concepts. The complexity of the model arises from the weight and distance staggered transport tariffs as well as the inclusion of delivery frequencies. The optimization model is a linear, multi-period, integer model that is able to use deterministic and dynamic demands with the objective of determining the optimal order quantity.

The use cases for the developed model are broad: Identifying saving potentials in existing inbound processes, selecting the inbound concept for new sourced parts, or supporting strategic management decisions. The latter complies with the case presented here. The main findings were the following: The idea of reference shipping units was identified as highly relevant, because the inbound logistics costs were calculated more precisely than with a single-part view. Delivery frequencies are equally relevant. The findings for CO2 emissions show that a reduction is possible without increasing logistics costs significantly. To summarize the innovation of this work, three aspects need to be stressed: The integrated combination of supply and transportation concepts; the implementation of staggered transportation tariffs; and the application of reference shipping units.

The presented model still leaves room for future investigations: The model should be extended by more capacity restrictions, such as available sequencing area or warehouse capacity, in order to gain more detailed results. In our optimization, the only means of transport is a mega trailer. To model more different inbound concepts, further means of transport should be considered. Additionally, modelling other combinations of supply and transportation concepts may be interesting, e.g. a JIS milk run. Although the model detects saving potentials within the inbound concepts, the final decision for changing the inbound process requires the evaluation of necessary investments and the future development of the concerned parts' characteristics.

REFERENCES

Bravo, J.J. and Vidal, C.J. (2013), "Freight transportation function in supply chain optimization models. A critical review of recent trends", *Expert Systems with Applications*, Vol. 40 No. 17, pp. 6742–6757.

EN 16258, *Methodology for calculation and declaration of energy consumption and greenhouse gas emissions of transport services*, European Committee for Standardization.

Göpfert, I. (2013), *Automobillogistik: Stand und Zukunftstrends,* 2nd edition, Springer Gabler, Wiesbaden.

Hoen, K.M.R., Tan, T., Fransoo, J.C. and van Houtum, G.J. (2014), "Effect of carbon emission regulations on transport mode selection under stochastic demand", *Flexible Services and Manufacturing Journal*, Vol. 26 No. 1-2, pp. 170–195.

Kontovas, C.A. and Psaraftis, H.N. (2016), "Transportation Emissions: Some Basics", in Psaraftis, H.N. (Ed.), *Green Transportation Logistics*, *International Series in Operations Research & Management Science*, Vol. 226, Springer International Publishing, Cham, pp. 41–79.

OpenSolver (2017), "Guide to Solvers. Gurobi", available at: http://opensolver.org/guide-to-solvers/ (accessed 1 April 2017).

Porter, M.E. (2004), *Competitive advantage: Creating and Sustaining Superior Performance,* 1. Export Edition, Free Press, New York, London.

Schulte, C. (2009), *Logistik: Wege zur Optimierung der supply chain, Vahlens Handbücher,* 5th edition, Vahlen, München.

VDA-5010, *Standardbelieferungsformen der Logistik in der Automobilindustrie*, available at: https://www.vda.de/de/services/Publikationen/Publikation.~497~.html (accessed 15 January 2017).

Wagner, S.M. and Silveira-Camargos, V. (2011), "Decision model for the application of just-in-sequence", *International Journal of Production Research*, Vol. 49 No. 19, pp. 5713–5736.

Weber, J. (2012), *Logistikkostenrechnung: Kosten-, Leistungs- und Erlösinformationen zur erfolgsorientierten Steuerung der Logistik,* 3rd edition, Springer Vieweg, Berlin.

World Bank Group and ECOFYS (2016), "Carbon Pricing Watch 2016", available at: https://openknowledge.worldbank.org/bitstream/handle/10986/24288/CarbonPricingWatch2016.pdf?sequence=4&isAllowed=y (accessed 25 January 2017).

AUTHOR BIOGRAPHIES

CORINNA MAAS was born in 1986 in Marl, Germany and went to the University of Cologne where she studied business administration and international management until 2012. Since 2014 she has worked as a research assistant at the Institute for Materials Handling, Material Flow, Logistics at Technical University of Munich. She is part of a cooperation project with the commercial vehicle manufacturer MAN Truck & Bus.

ANDREAS TISCH was born 1989 in Munich, Germany. He studied a combination of mechanical engineering and business administration at the Technical University of Munich. This work was generated in connection with his master's thesis at the Institute for Materials Handling, Material Flow, Logistics in 2016. Since 2017 he has worked for the BRUNATA-METRONA Group.

CARSTEN INTRA was born in 1971 in Koblenz, Germany. He studied mechanical engineering in addition to economics at the RWTH University Aachen. He obtained his doctoral degree at the Laboratory for Machine Tools and Production Engineering of the RWTH University Aachen. He joined MAN Nutzfahrzeuge AG in 2001. Since April 2012 he has been a Member of the Executive Board for Production & Logistics at MAN Truck & Bus AG, Munich and also an Executive Board Member for Research & Development since November 2015.

JOHANNES FOTTNER was born in 1971 in Munich, Germany. He went to the Technical University of Munich where he studied mechanical engineering. He obtained his doctoral degree in the research field of technical logistics in 2002. In 2008 he became managing director of MIAS Group. Since 2016 he has been back at the Technical University of Munich as a professor for technical logistics.

MODELING AND SIMULATION OF COOPERATION AND LEARNING IN CYBER SECURITY DEFENSE TEAMS

Pasquale Legato and Rina Mary Mazza
Department of Informatics, Modeling, Electronics and System Engineering
University of Calabria
Via P. Bucci 42C - 87036, Rende (CS), Italy
e-mail: {legato, rmazza}@dimes.unical.it

KEYWORDS
Simulation optimization, cyber security, team formation and cooperation.

ABSTRACT

Cyber security analysts may be organized in teams to share skills and support each other upon the occurrence of cyber attacks. Team work is expected to enforce the mitigation capability against unpredictable attacks addressed against a set of cyber assets requiring protection. A conceptual model for evaluating the expected performances of cooperating analysts by reproducing their learning process within a team is proposed. Analytical approaches to solve the underlying state-space model under stochastic evolution and discrete-event simulation are both discussed. The basic assumption is that a set of regeneration points corresponds to skill achievement through learning. A Simulation-based Optimization (SO) tool ranging from the inner level modeling of the cooperation-based learning process to the outer assignment of analysts to assets is then presented. Team formation may be supported by the SO tool for obtaining the team composition, in terms of individuals and skills, that maximizes system performance measures. Numerical results are reported for illustrative purposes.

INTRODUCTION

In today's rapidly changing threat landscape, cyber attacks are becoming much more common and damaging. To keep pace with this change, one must start by recognizing that cyber defence is as much about people as it is about technology. As a matter of fact, governments, organizations and industry worldwide are seeking ways to both manage and improve the expertise of their human resources in order to prevent, mitigate and recover from cyber attacks (NATO 2016).

Generally speaking, with respect to human resource management, a company may decide to deploy its cyber defense security analysts according to different *modus operandi*. Cyber defense analysts may be called to *i*) work alone according to traditional individualistic approaches or *ii*) in consultation with other analysts who are committed to a common mission and are willing to share the knowledge that is necessary to fulfill that mission (Kvan and Candy 2000). In the former case, it is a matter of being in charge of one's own achievements, concentration and schedule when deciding what to do and when to do it. In the latter, it is about teaming two or more individuals, with complementary background and skills, who organize their efforts in a mutual supportive way, share experience and complete common tasks.

Testing and evaluating the suitability of either of these two alternatives in a cyber attack scenario, under stochastic attack arrivals and mitigation services, requires a systematic approach in order to *i*) evaluate overall attack tolerance with regard to system performance degradation and *ii*) assess the effectiveness of using cooperation-driven learning among teammates as a countermeasure against cyber attacks.

To this purpose, a state-space model is exploited to mimic the learning process of an analyst when working in consultation with other analysts. This learning model, applied to the cooperating members of the same team, allows to account for the acquisition of new skills or the growth of expertise on pre-existing skills. So, cooperation-driven learning becomes a countermeasure against attacks by increasing positive attack mitigation in whatever be scenario. Both the analytical and simulation solution of the model are discussed as possible evaluation tools within a more general and powerful framework (Legato and Mazza 2016) aimed at optimizing the benefits from cooperation-based learning.

The paper is organized as follows. A description of the cooperation and learning process is provided in section 2. The conceptual model of the attack arrival and mitigation process and how it can account for cooperation-based learning among analysts is introduced in section 3. Analytical and simulation methodologies for model solution are discussed in section 4, while section 5 focuses on the choice to embed an SO procedure in the overall solution framework. The SO tool is presented in section 6 and conclusions are drawn in section 7.

COOPERATION AND LEARNING

When working in a team, cooperating with colleagues is at the basis of the learning process experienced by any given analyst during his/her daily task of defending an assigned cyber asset against unpredictable attacks. The capability of an analyst to gain and, in turn, transfer

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

knowledge depends on which skills he/she bears, along with the specific level held for each skill. In the attempt to quantify measurable knowledge, here we consider a four-level scale: no level of skill, basic level, intermediate level and expert level. These levels are in a continuum meaning that anyone standing at any of these levels can pursue progressive skill acquisition and, thus, learn through the continuum until he/she has reached the expert level. Put in other terms, unskilled analysts can learn from basic level analysts, intermediate level analysts and expert level analysts; basic level analysts can learn from intermediate level analysts and expert level analysts; intermediate level analysts can learn from expert level analysts. Obviously, expert level analysts can only act as hand-on knowledge workers, in addition to their individual role as cyber attack mitigation units.

Let us now consider a possible evaluation program according to which credits are awarded to analysts for every attack their skills allow them to mitigate. A credit is a measure of security performance whose amount depends on the type of attack – the more dangerous the attack, the higher the amount of credits rewarded. Whether working alone or in cooperation with others, an analyst's behavior is meant to collect as many credits as possible over time. If the analyst works alone and is skilled to manage an incoming threat, mitigation occurs according to a service time that depends on the type of attack. As a result of attack mitigation, the analyst is rewarded with the entire credit and free again to face new threats. If no such skill is held by the analyst, the lack of ability to mitigate the malicious attack may produce a negative impact on the entire system and likely cause a loss of overall performance, unless he/she works in consultation with other analysts. Practically speaking, the analyst may consult with skilled team members, if any, and thus start acquiring the necessary knowledge to manage the attack. If such a teammate exists and is not already engaged in other activities, attack mitigation may begin; otherwise, the analyst must wait. As a result of the ongoing interaction process, the "enquiring" analyst starts to accumulate knowledge (e.g. one scalar unit for each mitigation completed under cooperation or for each time unit spent in a cooperation state) because of the learning process he/she is undergoing. The team members that took part in the knowledge-sharing process then share the related credits. If none of the team members hold the appropriate skill to manage the attack, similar to the work alone *modus operandi*, this lack produces a negative impact on the entire system and causes a loss of overall performance.

A new skill achieved through the learning process may be recognized to the analyst after a (positive) periodic verification, provided that a fixed threshold on the number of attacks mitigated in cooperation has been reached. This skill recognition is the result of an examination whose timeframes and procedures are approximately scheduled by the senior management. Our focus is to investigate the extent to which there is a practical possibility of building a stochastic model of an analyst's evolution through his/her work by means of cooperation-based learning activities and then solving it by analytical approaches and simulation techniques.

CONCEPTUAL MODEL

From a conceptual point of view, the attack arrival and mitigation process experienced by an asset which is protected by a suitably skilled analyst may be represented by the client-server paradigm. The attack is a client who arrives randomly to a given asset and should be mitigated (serviced) by the analyst dedicated to that asset. A queued attack represents a non-detected status, but, as the attack "waits" to be detected, usually the damage delivered to the asset becomes bigger. Once detected, the attack either receives a mitigation service from the dedicated analyst or the analyst is forced to ask for cooperation from a colleague dedicated to another asset. Whatever be the case, the time required by the analyst(s) to mitigate the attack is random and may grow larger due to both a greater time of detection and a delay in starting the mitigation caused by a lack of cooperation.

The client-server model at hand is illustrated in Figure 1 for a couple of assets with two cooperating analysts. Clearly, it results in a non-standard queueing system not only because of the correlation among service duration and waiting time, but also because, when called to cooperate with a teammate, a server-analyst appears to be on "vacation" to his/her asset and to other colleagues as well.

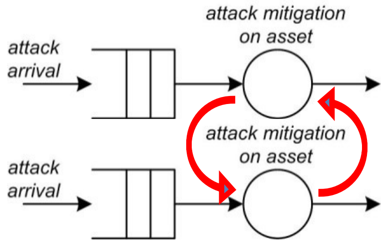


Figure 1: The Client-Server Model for the Attack Arrival and Mitigation Process under Cooperation

After using the above client-server system to represent working and cooperation among analysts, now we need to model the underlying learning process. A possible learning model should quantify, by means of a scalar quantity, the amount of knowledge incorporated by an analyst as a result of his/her attack mitigation experience under different scenarios and management policies.

In this respect, state-space models have already proven to be successful (Distefano et al. 2012) in capturing dynamic effects in reliability and availability studies. So, we choose to model the analyst's evolution over time by means of a sequence of states, each representing the different conditions in which an analyst may be found. These conditions are: idle, busy, waiting for colleague, teaching and learning, as illustrated in Figure 2.

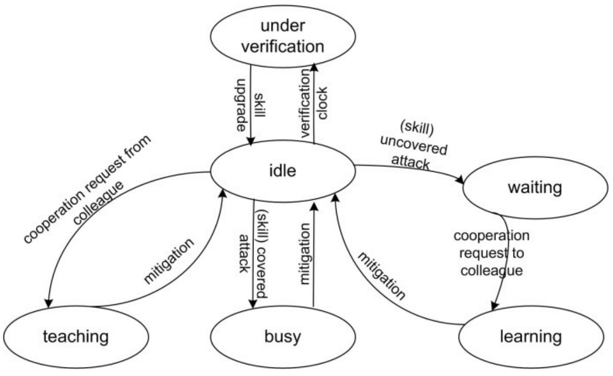


Figure 2: State-based Representation of Analyst Evolution

In principle, the learning state could be completely characterized by a mathematical real-valued function capturing the instantaneous measure of the learning growth. Similarly to the instantaneous hazard function in reliability modeling (Trivedi 2002), the rate of learning may be dependent or not from the age in a proper state. As for the skill upgrade policy, the real domain suggests that skill upgrade is the result of a successful verification of the knowledge gain achieved by an analyst between two successive verification instants. Verification activities occur rather cyclically along the analyst's working life and is aimed at evaluating the (cumulative) time spent by the analyst in the cooperation-based learning state (through repeated visits).

METHODOLOGIES FOR SOLUTION

Markov and generalized Markov models (Kulkarni 2009) are at the basis of the analytical tractability of a state-space based stochastic model. Nowadays, efficient tools are available to also manage the case where a large set of states need to be considered (Trivedi and Sahner 2009). However, besides the size of the model at hand, the mathematical tractability of the state-space model in Figure 2 relies upon the existence of points within the process where the memoryless property occurs. In our case, these points may occur at the time epochs of any given verification, followed by the upgrade of the analyst's skill level. Moreover, the further assumption that any future state of the analyst between two successive regeneration points can only depend from his/her state at the latest regeneration point, leads us to the concept of Markov renewal sequences. Therefore, Markov regenerative processes (Logothetis et al. 1995) appear to be the most powerful analytical tool for the stochastic modeling of the cooperation-based learning process of our interest. In particular, MRGPs allow generally distributed clock times for skill verification and upgrade. To this respect, one should formulate the MRGP through the definition of the related three matrix valued functions concerning transition probabilities, global kernel and local kernel (op. cit.). In particular, the local kernel matrix describes the behavior of the MRGP between two consecutive regeneration time points. It is required to compute the steady-state solution of the MRGP, provided that the discrete-time Markov chain embedded at the regeneration points is finite, aperiodic and irreducible and, therefore, returns the unique steady-state solution for the state visit ratios.

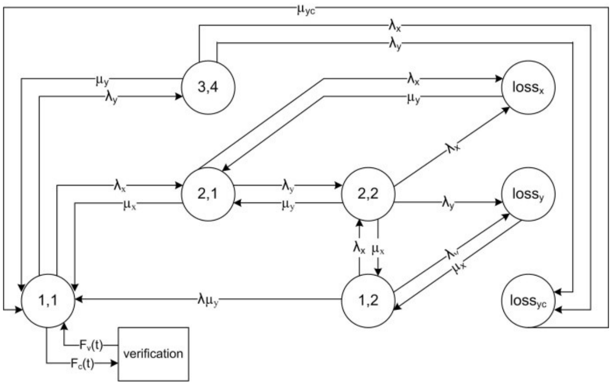


Figure 3: State-Diagram for a Simple Model

For illustrative purposes, let us consider the simple model in Figure 3 in which two cyber assets are both subject to two different attacks - attack x, attack y with rates λ_x and λ_y, respectively. Asset 1 is protected by analyst 1 who bears only the skill required to mitigate x-type attacks with mitigation rate μ_x. Similarly, asset 2 is protected by analyst 2 who bears only the skill required to mitigate y-type attacks with mitigation rate μ_y. Cooperation between the two analysts may occur when one requires the skill of the other to cover the attach on his/her asset. In this example, we assume that mitigation, with or without cooperation, should start immediately; otherwise, a "loss" (i.e. degradation) in system performance occurs. Thus, the state of an analyst is one of the following: 1 – idle, 2 – busy, 3 – learning, and 4 – teaching. Observe that the "verification" state, refers to skill verification for both the analysts. Its occurrence is regulated by the distribution function of the verification clock time, $F_c(t)$, while its time to positive completion is regulated by the distribution function $F_v(t)$ and leads to a new regeneration cycle under an upgraded skill level.

The state-diagram amenable to an MRGP-based analysis is described as follows. State (1,1) means that both analysts are idle and verification occurs only in this state; state (2,1) means that analyst 1 is busy because of the occurrence of an x-type attack on his/her own asset; (2,2) means that both analysts are busy mitigating an x-type and y-type attack on their respective assets; state (1,2) means that analyst 2 is busy because of the occurrence of an y-type attack on his/her own asset; (loss_*) represents the impossibility to provide mitigation for neither attacks, whether separately (loss_x, loss_y) or in cooperation (loss_{yc}, loss_{xc}); (3,4) means that analyst 1 is learning by cooperating with analyst 2 in mitigating an y-type attack on his/her own asset;); (4,3) means that analyst 2 is learning by cooperating with analyst 1 in mitigating an x-type attack on his/her own asset. For simplicity, in Figure 3 we represent only half of the entire model, i.e. the part referred to attack and mitigation activities occurring on asset 1. The corresponding activities on asset 2 are defined by analogy.

In principle, the state-diagram embedded in the regeneration cycle could be a continuous-time Markov chain provided that we assume exponential distributions for both attack occurrences and mitigation services, under independent sojourn times per visit in any given state. One could relax the exponential assumption to get a semi-Markov embedded process. However, the independence assumption on the sojourn time per visit would somehow limit the representation of the underlying memory effect of the learning activity accomplished by correlated successive returns in the same learning state.

This stated, it is our belief that providing a simulation framework for both performance evaluation and optimization is worth becoming the major goal of our current research contribution. Simulation allows us to obtain greater flexibility in setting a more realistic queuing-based description of the management policies regarding team formation and analyst cooperation in the cyber security real domain (Poste Italiane, the Italian national postal service) that has stimulated our work. Here, dealing with at least 10 different assets and ten analysts is quite common. Moreover, the optimal assignment of skilled analysts to assets may be pursued by embedding a suitable meta-heuristic based search process for better feasible assignments and, thus, learning by cooperation within a simulation based optimization tool.

Within the simulation framework, the quantitative analysis of the analyst's cooperation-based learning process between two (and also over many) successive regeneration epochs is carried out by regenerative discrete-event simulation (Shedler 1992). The effectiveness of regenerative simulation relies upon the practical possibility of replicating a suitable number of sample trajectories containing regeneration epochs and, therefore, estimate the expected performance measures of interest by both intra- and inter-cycle sample means. In particular, we may estimate the long-run probability of finding an analyst mitigating an attack in cooperation with a colleague. This probability is given by the ratio between the expected time spent by the analyst in a learning state divided by the expected duration of the overall time needed by that analyst to change his/her skill level:

$$P[\text{analyst in learning state}] = \frac{E[\text{time spent in learning state}]}{E[\text{time between change of skill level}]}$$

So, the simulation tool may also be used to validate analytical tools based on MGRPs against real policies, data and statistical distributions. This will be the subject of a companion paper.

SIMULATION-BASED OPTIMIZATION

In the illustrative example presented in the previous section, the analysts have already been assigned to a specific cyber asset and their teaming into a group of two is the only option available. When the number of analysts, assets, skills and skill levels grows larger, analyst-asset assignment and team formation is far from being so straightforward. Due to all the possible combinations, one may have to first *generate* teams by means of a search process and then *evaluate* them by means of an evaluation process. This situation may benefit from introducing a simulation-based optimization procedure (SO) (Fu and Nelson 2003) in the framework under development. In an SO approach a structured iterative approach calls an optimization algorithm to decide how to change the values for the set of input parameters (e.g. analyst-asset assignment and team formation) and then uses the responses generated by simulation runs to guide the selection of the next set. The logic of this approach is shown in Figure 4.

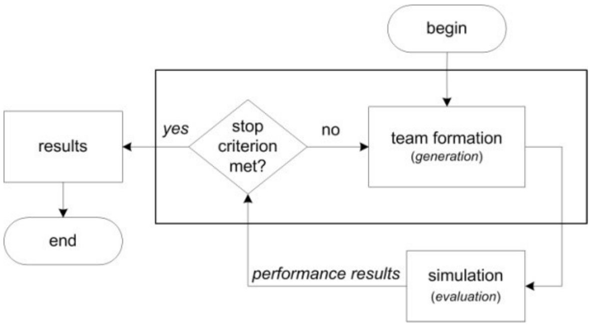


Figure 4: The Logic of Simulation-based Optimization

On the *generation* side, the first step in assembling cyber defence teams consists in assigning the analysts to the cyber assets, at least one for each asset. In the second step, the actual team formation is carried out by deciding the maximum number of teammates that can form a team, along with the specification of which colleagues should be part of the team. This step is guided by the idea of grouping analysts with complementary skills in order to cover a wider range of attacks and, thus, favor the learning process among as many teammates as possible. On the other hand, too many unskilled members on a team may prevent the skilled teammates from protecting their own assets because too busy in answering support requests from teammates.

During team generation, should an exhaustive coverage of all the possible system combinations be not reasonable, nor affordable from a computational point of view, then metaheuristic-based approaches would have to be addressed. Here we use a *simulated annealing* procedure (Kirkpatrick et al. 1983) which was first introduced by developing the similarities between combinatorial optimization problems and statistical mechanics. In the field of metal sciences, the annealing process is used to eliminate the reticular defects from crystals by heating and then gradually cooling the metal. In our case, a reticular defect could be seen as grouping analysts in teams that are not able to "properly" protect cyber assets and, thus, guarantee a given quality of service level when the above assets undergo an attack. Technically speaking, the annealing process is aimed to generate feasible teams of analysts, explore them in a more or less restricted amount and, finally, stop at a satisfactory solution. To avoid getting caught in local minima, during the exploration process a transition to a worse feasible solution can occur with probability

$$p=exp(\Delta/T)$$

where Δ is the difference between the values of the objective function (measure of learning) of the current solution (state) θ and the candidate solution θ_t and T is the process temperature. A prefixed value of T determines the stop of the entire process and it usually decreases according to a so-called *cooling schema*. Unfortunately, in the literature there is no algorithm that can determine "correct" values for the initial temperature and cooling schema, but, as suggested by empirical knowledge simple cooling schemas seem to work well (Ingber 1993).

In the following, some pseudo-code is given for the original SA algorithm for a minimization problem.

```
Algorithm: Simulated Annealing
$\theta \leftarrow$ initial solution
for time = 1 to time-budget do
    T $\leftarrow$ cooling-schema[time]
    if T=0 then
        Present current solution as the estimate of
        the optimal solution and stop
    Generate a random neighbor $\theta_t$ of the current
    solution $\theta$ by performing a move.
    $\Delta = f(\theta) - f(\theta_t)$
    if $\Delta > 0$ then
        $\theta \leftarrow \theta_t$
    else
        $\theta \leftarrow \theta_t$ (with probability $p=exp(\Delta/T)$)
end for
```

When customizing the SA algorithm to our problem, some choices need to be made.

To begin with, choosing the proper cooling schema has great impact on reaching a global minimum. In particular, it affects the number and which analysts are assigned to a team (solutions) that will be evaluated by running the SA algorithm. To this end, the so-called simple mathematical cooling schema $T_{i+1}=\alpha \cdot T_i$ has been tested, and the best results are returned for an initial temperature $T_0=100$ and a decreasing rate $\alpha \cong 0.9$.

The "move" definition for neighborhood generation is very context-sensitive. For our problem, a move must be defined with respect to the feasibility (or lack thereof) of a team by taking into account the analysts' skills. Some examples of moves are:

- move analyst *l* from team *i* to team *j* (*i* <> *j*);
- swap analyst *l* and analyst *k* (*l* <> *k*), originally assigned to team *i* and team *j* (*i* <> *j*), respectively.

As far as the stopping criteria are concerned, designers can choose among the following possibilities:

- stop when the algorithm has reached a fixed number of iterations *n* or an upper bound on the available time-budget;
- stop when the current solution has not been updated in the last *m* iterations;
- stop when the cooling schema has reached a *lower bound* on the temperature.

Although we do not use this algorithm to perform an exhaustive search of the sample space, nor are we provided with any sort of control running on which part of the feasible set is being explored, the solutions returned as final output are likely to belong to the set of optimal global solutions (Banks et al. 2000) that under the cooperation-based learning policy allows to deliver:

- knowledge gain;
- percentage of attacks mitigated;
- resource (analyst) utilization;
- number of credits gained;
- number of cyber defense security analysts per team;
- cyber defense security team composition in terms of skill types and levels held by every single analyst assigned to every single team.

On the *evaluation* side, the simulation model for the attack arrival & mitigation process has been conceived according to an attack-centric point of view: attacks are entities flowing through a cyber network that may damage cyber assets and call for mitigation by (a group of) skilled analysts who seize and/or release resources while doing so.

An attack is defined by a record:

```
type attackrecord
  Eventtype
  Arrivaltime
  Analyst
  teammate(s)
  Asset
  Attacktype
  Operationtime
  Queue
end type
```

The primary attack events (in italics) of the simulation model are listed in Table 1, along with their effect on the system state in terms of actions and resources seized and/or released. Each event marks the beginning or the end of a given model activity and must be counted only once. An event always triggers the beginning (end) of a specific activity, but, for the sake of shorter notation, any "begin" ("end") prefix (suffix) is omitted from the event name.

Table 1: Events of the Discrete-Event Simulation Model

Event	Actions	Resources	
		Seize	Release
attack_arrival	schedules *attack_outcome*	analyst	-
	schedules *attack_outcome*	teammate	
	queues request on analyst	-	-
	queues request on teammate	-	-
	updates statistics	-	-
attack_outcome	schedules *attack_arrival*	-	analyst
	schedules *attack_arrival*		teammate
	unqueues request on analyst	analyst	-
	unqueues request on teammate	teammate	-
	updates statistics	-	-

THE SO TOOL

The SO tool is illustrated in Figure 5. It has been designed and implemented in compliance with all the conventional steps used to guide a thorough and sound simulation study (Banks et al. 2000).

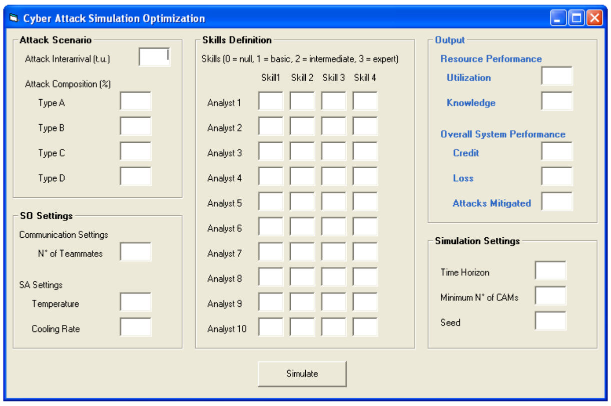


Figure 5: Snapshot of the SO Tool

All experiments have been run on a personal computer equipped with a 2.26Hz Inter Core™2 duo processor and 3 GB of RAM.

The GUI panel in Figure 5 has been conceived to meet the suggestions given by the cyber security senior management of Poste Italiane. It allows to easily specify the input data and SO parameters by means of proper sections. In the particular case at hand, we consider an attack scenario in which 10 analysts are used to defend 10 cyber assets against 4 types of attacks. Both attack interarrivals (in time units) and composition (in percentage) with respect to different types of attacks need to be specified. The skills of the 10 analysts are then defined by specifying for every analyst which skills he/she features and the level of competence for each skill (0=no skill, 1=basic, 2=intermediate, 3=expert). After inserting the maximum number of teammates in a group (here ranging between 1 and 10), the input stage is then completed by providing the SA and simulation settings. These are, respectively, the initial temperature along with the cooling rate of the SA procedure, the overall time horizon, the minimum number of cooperative attack mitigations (CAMs) completed per verification cycle and the simulation seed.

We now define the scenario for the preliminary set of experiments in order to evaluate the management of analysts and the resulting system performance. The expected measure of learning, skill upgrade, best team(s) composition (in number and skills) are returned for the former, while system credit and system loss are returned for the latter.

On average, attacks occur every 100 time units according to an exponential renewal process (λ, the average interarrival rate, is thus equal to 1/100). Arrivals are characterized by a combination of 4 different types of attacks (i.e. 70% type A, 15% type B, 10% type C and 5% type D). The 10 analysts are able to provide attack mitigation according to their own skills which are reported in Table 2.

Table 2: Skill Types and Levels of the Cyber Defense Security Staff

Analyst/Skill	A	B	C	D
analyst 1	0	0	1	0
analyst 2	3	0	1	0
analyst 3	0	2	0	3
analyst 4	0	0	3	0
analyst 5	3	0	0	0
analyst 6	3	3	0	0
analyst 7	0	1	0	3
analyst 8	1	0	0	0
analyst 9	3	3	3	0
analyst 10	3	3	0	0

In the given scenario, analysts respond to attacks by working alone (n° of teammates=1) or in cooperation with other analysts (n° of teammates>1). The rate (μ) of the attack mitigation activity depends on the type of attack, the skill level held by the analyst and if mitigation occurs alone or in cooperation with other analysts. In the later case, mitigation times are inflated by 30%.

The initial temperature and the cooling rate of the SA scheme are set equal to 100 and at least 0.948, respectively, so that at least 100 different team-formation and assignment configurations are considered for the given scenario. The time horizon is fixed at 14400 time units and both point estimates and 95% confidence intervals can be obtained for the measures of learning, system credit and system loss. Here, for clarity of illustration, in Figures 6 through 8 we prefer plotting the central value within the interval estimates to show some preliminary numerical results.

Let us start by considering the (average) measure of learning (i.e. one scalar unit for each mitigation completed under cooperation). Figure 6 shows that this measure grows approximately linearly with the number of analysts per team. In other terms, thanks to cooperation, the cumulative number of attacks mitigated by all the analysts in the fixed time horizon goes from 0 (no cooperation) to 180 (complete cooperation among the 10 analysts).

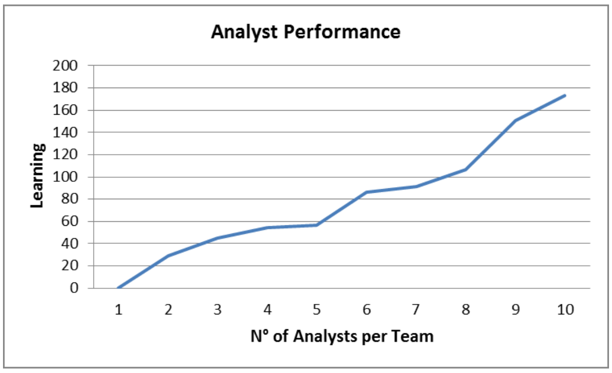


Figure 6: Trend of Analyst Learning

From the analyst's individual point of view, the learning benefit is resumed in Table 3 which, for each analyst, reports his/her skill upgrade achieved through cooperation along the time horizon.

Table 3: Skill Upgrade per Analyst

Analyst	Initial Skills	Final Skills
1	C	A, B, C
2	A, C	A, B, C, D
3	B, D	A, B, C, D
4	C	A, B, C
5	A	A, B, C
6	A, B	A, B
7	B, D	A, B, C, D
8	A	A, B, C
9	A, B, C	A, B, C, D
10	A, B	A, B

As one may see from Table 4, this knowledge growth is accomplished in conjunction with a specific asset-analyst assignment and subsequent team formation in which analysts with complementary skills have been teamed together.

Table 4: Details of the Asset-Analyst Assignment and Team Skills when Max No. of Teammates=7

Analyst	Asset	Teammates	Team Skills
1	6	2, 4, 5, 6, 8 & 10	A, B, C
2	9	1, 3, 4, 5, 7 & 10	A, B, C, D
3	10	2, 4, 5, 6, 7 & 9	A, B, C, D
4	7	1, 2, 3, 5, 6 & 10	A, B, C, D
5	8	1, 2, 3, 4, 6 & 9	A, B, C, D
6	4	1, 3, 4, 5, 7 & 8	A, B, C, D
7	2	2, 3, 6, 8, 9 & 10	A, B, C, D
8	5	1, 6, 7, 9 & 10	A, B, C, D
9	1	3, 5, 7, 8 & 10	A, B, D
10	3	1, 2, 4, 7, 8 & 9	A, B, C, D

As for the remaining performance measures, let us first consider the system credit recalling that the more dangerous the attack, the higher the amount of credits rewarded. In this set of experiments, A-type attacks are the less dangerous (1 credit rewarded per mitigation), while D-type attacks are the most dangerous (4 credits rewarded per mitigation). As shown in Figure 7, the behavior of the system credit follows a bathtub curve as the number of analysts per team grows. For small number of teammates, the benefit of cooperation is surmounted by the waiting times experienced by the requiring analysts when asking (a limited number) of skilled teammates for support. For middle-size teams, system credit remains rather stable. This is likely due to the greater number of skills and, thus, attacks covered by the teammates which affects the waiting times in a positive way. For a large number of analysts per team, both cooperation and waiting times benefit from the availability of all the skills against incoming attacks.

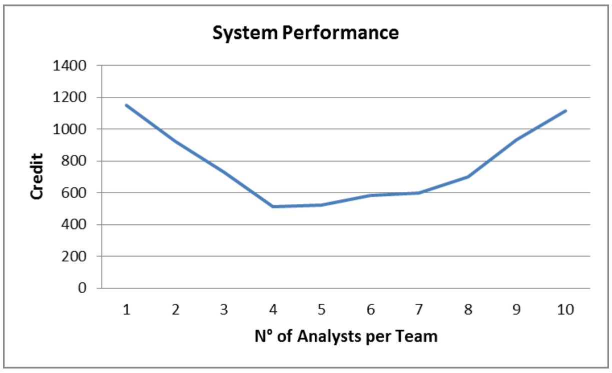


Figure 7: Trend of System Credit

As for system loss, this is a measure of the number of attacks that cannot be mitigated by the analysts due to a lack of the skills required to fulfil this purpose. By analogy with system credit, it is equal to 1 per non mitigated A-type attack and reaches 4 per non mitigated D-type attacks.

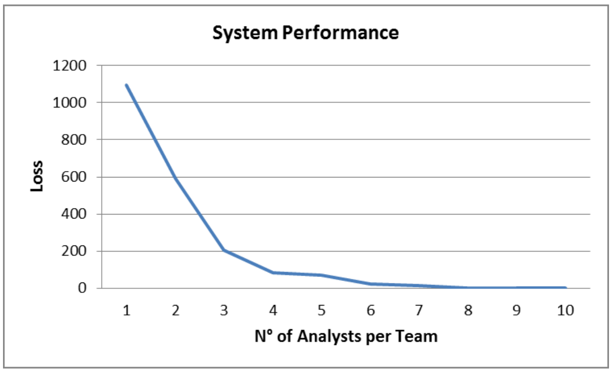


Figure 8: Trend of System Loss

Figure 8 shows that in the scenario under examination system loss is totally overcome when the number of teammates is equal to or greater than 7. As previously stated, this corresponds to the possibility of always finding a free skilled teammate upon request by an analyst.

CONCLUSIONS

A client-server system with multiple cooperating servers bearing different skills and learning capabilities has been discussed and illustrated. The centrality of an MRGP as an evaluation tool for the analysis has been highlighted. Then, a more general simulation tool has been proposed with the aim of pursuing the optimality of dynamic team formation to favor cooperation and learning among security analysts each dedicated to their own asset. The simulation-based approach allows for an effective evaluation and optimization of the whole organizational process, starting by the assignments of analysts to assets and reproducing the occurrence of attacks followed by cooperation and learning. The tool may also incorporate a relaxed MRGP aimed at reproducing the learning process of an analyst when working in consultation with other analysts. The learning model, applied to the cooperating members of the same team, allows to account for the acquisition of new skills or the growth of expertise on pre-existing skills. We have shown how the tool may be used to *i*) evaluate overall attack tolerance, in terms of credit and loss measures, with respect to system performance and *ii*) assess the effectiveness of using cooperation-driven learning as a countermeasure against cyber attacks, in terms of new skills achieved by analysts.

ACKNOWLEDGEMENTS

This work was partly supported by the Cyber Security Technological District funded within the Italian National Operational Programme for Research and Competitiveness 2007-2013 under grant number PON03PE_00032_2.

REFERENCES

Banks, J., J.S. Carson, B.L. Nelson and D.M. Nicol. 2000. *Discrete-Event System Simulation*. 3rd Edition. Prentice-Hall, Inc., Upper Saddle River, New Jersey.

Distefano, S., F. Longo, F., and K.S. Trivedi. 2012. "Investigating Dynamic Reliability and Availability through State-Space Models". In *Computers & Mathematics with Applications*, 64, No.12, 3701-3716.

Fu, M. and B. Nelson. 2003. Guest Editorial. *ACM Transactions on Modeling and Computer Simulation* 13, No.2, 105-107.

Ingber, L. 1993. "Simulated Annealing: Practice versus Theory". *Mathematical Computer Modelling* 18, No.11, 29-57.

Kirkpatrick, S., C.D. Gelatt and M.P. Vecchi. 1983. "Optimization by Simulated Annealing". *Science*, New Series, 220, No.4598, 671-680.

Legato, P. and R.M. Mazza. 2016. "A Simulation Optimisation-based Approach for Team Building in Cyber Security". *International Journal of Simulation and Process Modelling* 11, No.6, 430-442.

Logothetis, D., K.S. Trivedi and A. Puliafito. 1995. "Markov Regenerative Models". In: *Proceedings of the IEEE International Computer Performance and Dependability Symposium* (Erlangen, DE, April 24-26), 134-142.

Kulkarni, V. G. 2009. *Modeling and Analysis of Stochastic Systems*. 2nd edition. Chapman & Hall, London.

Kvan, T. and L. Candy. 2000. "Designing Collaborative Environments for Strategic Knowledge in Design". *Knowledge-Based Systems* 13, No.6 (Nov), 429-438.

NATO. 2016. Cyber Defence. http://www.nato.int/cps/en/natohq/topics_78170.htm [Last updated: 17 Jan. 2017, accessed on 14 Jul. 2016].

Shedler, G. S. 1992. *Regenerative Stochastic Simulation*. Academic Press - Elsevier, Oxford.

Trivedi, K. S. 2002. *Probability and Statistics, with Reliability, Queuing and Computer Science Applications*. 2nd edition. John Wiley & Sons, Inc., NY, NY.

Trivedi, K. S. and R.A. Sahner. 2009. "SHARPE at the Age of Twenty Two". *ACM SIGMETRICS Performance Evaluation Review* 36, No.4, 52-57.

AUTHOR BIOGRAPHIES

PASQUALE LEGATO is an Associate Professor of Operations Research in the Department of Informatics, Modeling, Electronics and System Engineering (DIMES) at the University of Calabria, Rende (CS, Italy). He has been a member of the Executive Board of the University of Calabria as well as university delegate for the supervision of associations and spin-offs from the University of Calabria. He has been involved in several EEC funded research projects aimed at the technological transfer of SO procedures and frameworks in logistics. He is a member of the INFORMS Simulation Society. His research activities focus on predictive stochastic models for cyber security, queuing network models, stochastic simulation and the integration of simulation techniques with combinatorial optimization algorithms. His e-mail address is: legato@dimes.unical.it and his web-page can be found at http://wwwinfo.dimes.unical.it/legato.

RINA MARY MAZZA is the Research Manager of the Department of Informatics, Modeling, Electronics and System Engineering (DIMES) at the University of Calabria, Rende (CS, Italy). She graduated in Management Engineering and received a PhD in Operations Research from the above university. She has a seven-year working experience on knowledge management and quality assurance in research centers. She has also been a consultant for operations modeling and simulation in container terminals. Her current research interests include discrete-event simulation and optimum-seeking by simulation in complex systems. Her e-mail address is: rmazza@dimes.unical.it.

NUMERICAL DISCRETE ELEMENT SIMULATION OF SOIL DIRECT SHEAR TEST

Krisztián Kotrocz
György Kerényi
Department of Machine and Product Design
Budapest University of Technology and Economics
H-1111, Muegyetem rkp. 3-9., Budapest, Hungary
E-mail: kotrocz.krisztian@gt3.bme.hu

KEYWORDS
soil model, direct shear test, 3D DEM, discrete element simulation.

ABSTRACT

One of the most common methods to measure soil mechanical properties (namely cohesion and internal friction angle) is direct shear box test. In this paper the development of a three-dimensional (3D) discrete element soil model for simulation of a cohesive soil's direct shear test is presented. The aim was to calibrate the properties of the Hertz-Mindlin with bonding contact model available in EDEM software to the results of real direct shear box test. The cohesion and internal friction angle were calculated from the equation of the Mohr-Coulomb line of the soil model. Results show that direct shear laboratory test can be simulated very well using discrete element method (DEM). The model's calculated cohesion and internal friction angle values and the corresponding mechanical properties of real cohesive soil were in good agreement with relative error of 4.44 percent and 4.66 percent, respectively.

INTRODUCTION

In the last few decades the development of agricultural machines led to increase the size and mass of the machines and therefore the stress applied into agricultural soils as well. In soil-wheel interaction normal- and shear stress is generated in the soil by both driven and non-driven wheels. In designing agricultural machines and their equipment (e. g. running gear and tillage tools) engineers need to know the mechanical properties, namely cohesion and internal friction angle of the soils. These properties depend mostly on the soil's moisture content and bulk density (Sitkei 1967).

The most common methods to measure soil's shear strength parameters are direct shear and triaxial laboratory tests (McKyes 1985). Direct shear test can be easier to conduct and therefore is widely applied in agricultural researches (Dirgeliene et al 2014). From the results of the test cohesion and internal friction angle can be identified using the Mohr-Coulomb criterion.

In addition in the 20th-21st century the information technology has been evolved a lot allowing to create and use numerical simulations for modelling real materials. The most known method is Finite Element Method (FEM) which is used mainly in simulation of homogeneous materials (e. g. steels and plastics). Efforts have been made for modelling soil with FEM (Mouazen and Neményi 1998; Chi and Kushwaha 1990) but FEM is not suitable for modelling granular materials (e. g. soils). Granular assemblies consist of individual elements and sub-assemblies with non-continuous displacements which cannot be simulated in FEM properly. In Discrete Element Method (DEM), developed by Cundall and Strack (Cundall and Strack 1979) the material is modelled as a group of individual particles therefore DEM can be a correct choice to simulate soil material. DEM has been used in several research works to study the dynamic motion of lunar (Nakashima et al 2010) and Mars wheel (Knuth et al 2012) or to simulate soil-tool interaction as well (Tamás et al 2013). Many researchers used DEM to simulate the direct shear test of real soil (Tamás et al 2013 and Sadek et al 2011) as well but non of them used the Hertz-Mindlin with bonding contact model in their simulations.

In this paper DEM was used to simulate cohesive soil's behaviour under direct shear test. The software used to perform the simulations was EDEM 2.7 Academic, available from DEM Solutions Ltd. Numerical direct shear test were performed and to calibrate the contact properties of the soil model a new process was used and will be presented in the paper.

MATERIALS AND METHODS

In discrete element simulations the whole process is divided into small timestep of dt. The displacements of each elements are calculated from the forces and moments acting on them using Newton's 2nd law in every single timestep of dt. The forces and moments can be determined from the overlaps of the particles according to the used contact model. Therefore contact models play an important role in discrete element simulations because the material properties can be modelled properly by using the correct contact model with sufficient contact properties between the elements. In this paper the Hertz-Mindlin with bonding contact model available in EDEM 2.7 Academic sofware was chosen to simulate cohesive soil material. This contact

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

model is based on the work of Hertz (Hertz 1882), Mindlin (Mindlin 1949), Mindlin and Deresiewicz (Mindlin and Deresiewicz 1953) and Potyondy and Cundall (Potyondy and Cundall 2004). In this contact model the forces acting on an element are divided into normal and tangential (shear) directions. The normal force can be calculated as follows:

$$F_n = \frac{4}{3} \cdot E^* \cdot \sqrt{R^*} \cdot \delta_n^{\frac{3}{2}}$$

where δ_n is the normal overlap of the contacting elements, E^* and R^* are the so-called equivalent Young's modulus and equivalent particle radius, respectively and can be calculated from the properties of the two contacting elements using Equation (1) and Equation (2).

$$E^* = \left(\frac{1-\nu_1^2}{E_1} + \frac{1- {}_{2}^{2}}{E_2} \right)^{-1} \tag{1}$$

where ν_1 and ν_2 are the Poisson ratio and the E_1 and E_2 are the Young's modulus of the first and second contacting element, respectively.

$$R^* = \left(\frac{1}{R_1} + \frac{1}{R_2} \right)^{-1}. \tag{2}$$

where R_1 and R_2 are the radius of the first and second contacting element, respectively.

In addition tangential force can be transmitted from one particle to another. The magnitude of this force is calculated by Equation (3).

$$F_t = -S_t \cdot \delta_t \tag{3}$$

where δ_t is the tangential overlap of the contacting elements and S_t is the tangential stiffness calculated as follows:

$$S_t = 8 \cdot G^* \cdot \sqrt{R^* \cdot \delta_n} \tag{4}$$

In Equation (4) the G^* is the equivalent shear modulus which is determined from the shear modulus and the Possion ratio of the contacting elements, G_1; G_2 and ν_1; ν_2, respectively.

$$G^* = \left(\frac{1-\nu_1}{G_1} + \frac{1-\nu_2}{G_2} \right)$$

The tangential force is limited by its maximum value which can be determined from the Coulomb fiction:

$$F_t \le F_n \cdot \mu_s$$

where μ_s is the static frictional coefficient between the elements.

There are additional damping forces acting in normal, F_n^d and tangential directions, F_t^d as well. These can be calculated with the following formulas:

$$F_n^d = -2 \cdot \sqrt{\frac{5}{6}} \cdot \beta \cdot \sqrt{S_n \cdot m^*} \cdot v_n^{rel} \tag{5}$$

$$F_t^d = -2 \cdot \sqrt{\frac{5}{6}} \cdot \beta \cdot \sqrt{S_t \cdot m^*} \cdot v_t^{rel} \tag{6}$$

In Equation (5) and Equation (6) m^* is the equivalent mass, S_n is the normal stiffness, β is given below, v_n^{rel} and v_t^{rel} are the relative normal and tangential velocities of the contacting elements, respectively. These can be determined by using the following equations:

$$m^* = \left(\frac{1}{m_1} + \frac{1}{m_2} \right)^{-1}$$

$$S_n = 2 \cdot E^* \cdot \sqrt{R^* \cdot \delta_n}$$

$$\beta = \frac{\ln e}{\sqrt{\ln^2 e + \pi^2}}$$

where e is the so-called coefficient of restitution.

In addition there are bond forces and moments in the contacts to bond the particles together and to simulate the soil cohesive behaviour. The following force- and moments are summed to the corresponding Hertz-Mindlin components:

$$\Delta F_n = -S_n^B \cdot A \cdot \Delta\delta_n \tag{7}$$

$$\Delta F_t = -S_t^B \cdot A \cdot \Delta\delta_t \tag{8}$$

$$\Delta M_n = -S_n^B \cdot J \cdot \Delta\theta_n \tag{9}$$

$$\Delta M_t = -S_t^B \cdot \frac{J}{2} \cdot \Delta\theta_t \tag{10}$$

In Equation (7) to Equation (10) the S_n^B and S_t^B are the bond normal- and tangential stiffness, $\Delta\delta_n$ and $\Delta\delta_t$ are the relative normal- and tangential displacements, $\Delta\Theta_n$ and $\Delta\Theta_t$ are the relative normal- and tangential rotations of the contacting elements, respectively and:

$$A = \pi \cdot R_B^2$$

$$J = \frac{1}{2} \cdot \pi \cdot R_B^4$$

are the area and the polar moments of inertia of the bond's cross section, respectively (R_B is the radius of the bond). $\Delta\delta_n$, $\Delta\delta_t$, $\Delta\Theta_n$ and $\Delta\Theta_t$ are calculated from the bond formation time of t_{Bond} when the bond forces and moments are set to zero. In addition these bonds can be broken when the normal- and tangential bond stresses exceed their limits (bond normal- and tangential strength):

$$\sigma_{\max} = \frac{-\Delta F_n}{A} + \frac{2 \cdot M_t}{J} \cdot R_B$$

$$\tau_{\max} = \frac{-\Delta F_t}{A} + \frac{M_n}{J} \cdot R_B$$

There is another important parameter, namely the contact radius ($R_{contact}$) wich determines the point where two particles become in contact from. While using bonds between the particles this parameter have to been set up larger than the real radius of the particle in order to allow to transmit tensile force between the elements while they are no more in contact physically.

The laboratory direct shear tests

Direct shear tests were conducted in the Szent István University of Gödöllő to measure the frictional mechanical properties of real cohesive soil. The soil samples were collected with core cylinders near Mohács, Hungary in the November of 2015 and were transported later to the Szent István University where ELE 26-2112/01 direct shear apparatus - shown in Figure 1 - was used to measure the soil mechanical properties.

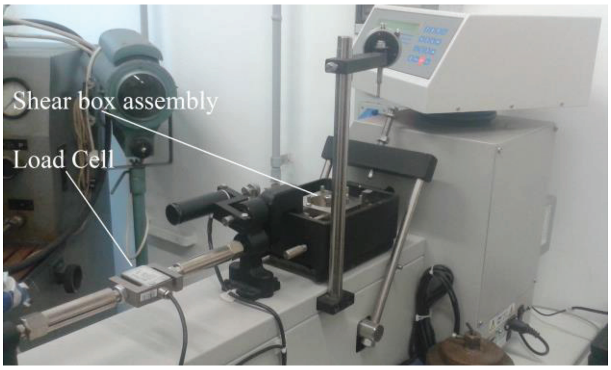


Figure 1: The used ELE 26-2112/01 type direct shear apparatus

The shear box with dimension of Ø6.4e-02x2.54e-02 m was filled with soil sapmles and the vertical load of 200 N, 400 N and 600 N was added to the samples, respectively. After that the horizontal speed of 5 mm min^{-1} which is equal to 8.3e-02 mm s^{-1} was set up and the shear process was started. During the measurements the horizontal displacement of the top section and the shear force (is necessary to push out the top section on the bottom section of the shear assembly) were measured by the built in ASCELL TC type load cell. Note that the vertical displacement of the samples were not measured during the measurements.

Finally the results were evaluated from the shear force-horizontal displacement diagrams. In case of each normal load of *N* there is a maximum shear force of *T* as it can be seen in the results section of the paper. From this data the Mohr-Coulomb line of the soil can be drawn and the mechanical properties of the soil were calculated using the following equation (Terzaghi 1943):

$$T = c \cdot A + N \cdot \tan\varphi$$

where *c* means the cohesion, *A* is the sheared area and *φ* is the soil's internal friction angle.

Development of the 3D discrete element model

A 3D discrete element model was developed to simulate direct shear tests of cohesive soil. First the same geometry was created virtually to that of real direct shear apparatus. Therefore two cylinders with radius of 3.2e-02 m and height of 1.27e-02 m each were set into the model and than were filled with spherical elements. After the elements settled down by Earth gravity and reached the equilibrium state (the maximum velocity of the elements was under the value of 1e-05 m s^{-1}) the vertical force was added to the model by a disk made of spherical elements as well. Between these particles there were no contacts calculated during the simulations they are only for allowing to add the vertical load to the soil model. Finally the top cylinder was moved horizontally with the speed of 8.3e-01 mm s^{-1} which is 10 times higher than the speed used in real direct shear tests. This was done to minimize the calculation time because in case of lower horizontal speeds the value of the shear force was not change significantly according to our earlier results but it takes for weeks to complete the simulations.

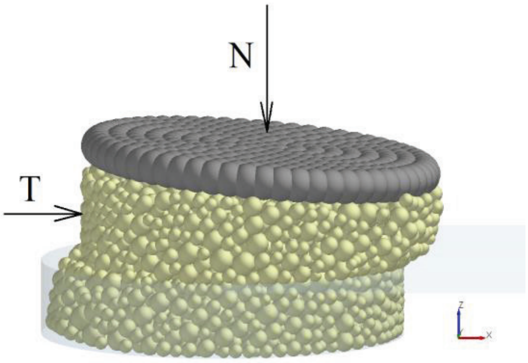


Figure 2: The 3D discrete element model of soil direct hear test with the disk

The timestep of 1e-05 was chosen which is approximately 11.7 % of the Rayleigh timestep. According to the EDEM user manual the timestep have to set up 20 % or lower to the Rayleigh timestep to guarantee good results from the simulations. During each direct shear simulations the horizontal force and the displacement of the top section and the vertical force of the disk (e. g. the normal load of the sample) were calculated and saved in each 5e-02th timestep. Finally the results were evaluated using Microsoft Excel 2016 software by drawing the corresponding shear force-horizontal displacement diagram and from that the Mohr-Coulomb line of the soil model as well.

While using the Hertz-Mindlin with bonding contact model the contact properties in Table 1 have to been set up before the simulation. To obtain these data the following steps were used:

- Step 1: Determination of the elements' Young's modulus. In this paper the Young's modulus was choosen according to the gradient of the shear force-displacement diagram in case of normal load of 200 N (see Figure 4).
- Step 2: Determination of the bond strength. In this paper the bond tangential strength was choosen to be approximately equal to soil's cohesion and the bond normal strength was choosen to be twice as much than the tangential strength presenting the shear stress is twice as dangerous as the tensile stress.
- Step 3: Determination of the elements' contact radius. In this paper the contact radius was choosen to be 1.3 times greater than the real particle's radius.
- Step 4: Calculation of the bond stiffness from the bond strength, contact radius and other geometrical properties of the elements.
- Step 5: Calibration of the bond radius to the maximum shear force.

The aim of this process is to determine or calculate the contact properties so that the bond breaks between two elements because it exceed its stress limit values and not because the two contacting elements get too far from each other so their contact radiuses do not overlap each other. Therefore in step 4 the stiffnesses of the bond should be calculated with Equation (11).

$$S_n^B = S_t^B = \frac{\tau_{\max}}{\Delta u} \qquad (11)$$

where Δu can be determined as follows:

$$\Delta u = 2 \cdot R_{contact} - NO$$

where *NO* is the average normal overlap of the elements at the bond formation time (t_{Bond}). With the data given in Table 1 the bond normal- and tangential stiffness were calculated with Equation (11) and were exactly 1.193e+07 Pa m^{-1}. Because of the computational error of the discrete element simulation this value was rounded up to 1.2e+07 Pa m^{-1}.

Table 1: The properties of the discrete element model derived from the 3D direct shear simulations

Parameter	Value
Geometrical properties	
Particle radius distribution (m)	2e-03…4.5e-03
Contact radius ($R_{contact}$) (m)	2.67e-03…6e-03
Initial porosity (before the normal load applied) (%)	0.415
Properties of the Herzt-Mindlin with bonding contact model	
Bulk density (kg m^{-3})	1800
Shear modulus (Pa)	2.88e+06
Poisson ratio (-)	0.3
coefficient of restitution (e) (-)	0.5
Friction coefficient between ball and ball (μ_{ball}) (-)	0.6
Friction coefficient between ball and walls (μ) (-)	0.5
Bond radius (R_B) (m)	1.2e-03
Bond normal stiffness (S_n^B) (Pa m^{-1})	1.2e+07
Bond shear stiffness (S_t^B) (Pa m^{-1})	1.2e+07
Bond normal strength (Pa)	7.738e+4
Bond shear strength (Pa)	3.869e+4

The data in Table 1 are the results of many calibrational simulations.

RESULTS AND DISCUSSION

The results of real direct shear tests and 3D discrete element simulations will be presented together in this section.

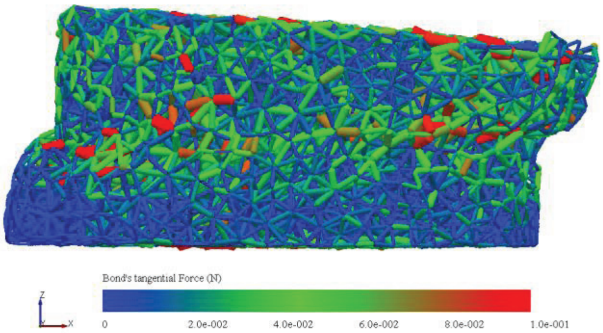


Figure 3: The bonds between the elements with the calculated bond tangential forces in case of 3D discrete element simulation of normal load of 200 N

In Figure 3 the bonds between the particles are presented and are coloured according to the bond tangential force in case of simulation of normal load of 200 N. It can be seen that the higher bond forces arise in the middle of the model (near the shear zone) and at the top of the model where the disk with the normal load is contacting with the soil praticles. This observation can be experienced in case of real soils as well where the highest stresses arise close to the shear zone of the sample.

In Figure 4 the results of real direct shear test as dotted lines and results of discrete element simulations as continuous lines are presented. The gradient of the shear force-displacement curve of simulation is similar to that of measurement in case of normal load of 200 N and in range of displacements higher than 2e-03 m. This was expected according to the step 1 of the calibration process. Small difference between these curves can be seen in range of displacements smaller than 2e-03 m. This can be experienced probably because there are sudden changes in the shear force values in range of displacements smaller than 2…3e-03 m in case of each test which can be related to the error of direct shear measurements. Figure 4 shows greater differences between the simulation and measurement curves in case of higher normal loads especially in range of small displacements but comparing the maximum values of the shear force in the simulations and in the measurements they are close to each other in case of all normal loads. This means that the failures of the soil models in the simulations are very similar in point of view of shear force value to the failures of real soils. The maximum shear forces were summarized in Table 2 as well. These data can be used to draw the Mohr-Coulomb line of the soil models and real soils as well.

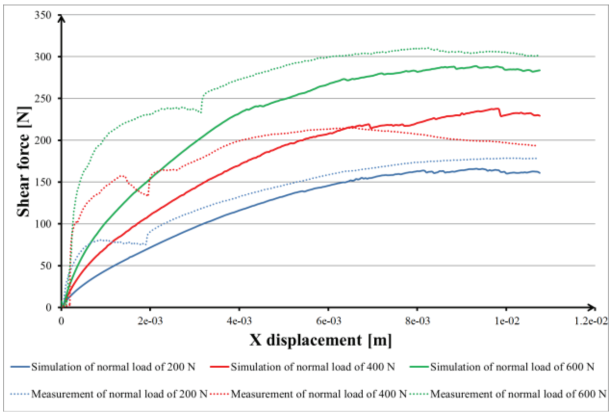


Figure 4: Results of the direct shear simulations

Table 2: The maximum of the shear force in the three different simulations and direct shear tests

Normal load (N)	Maximum of the shear force in simulation (N)	Maximum of the shear force in measurement (N)
200	165.99	179.03
400	237.56	215.23
600	288.72	310.54

These lines can be seen in Figure 5 where the results of the simulations are illustrated as red points and red dotted line and the results of the measurements are illustrated as blue points and blue dotted line.

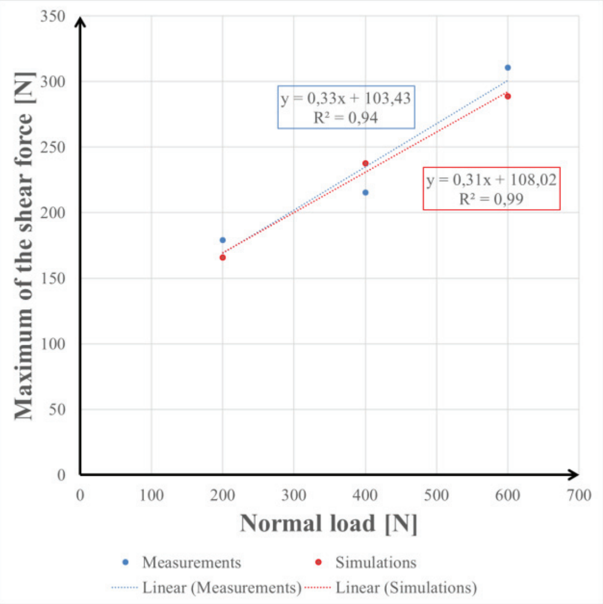


Figure 5: The calculated Mohr-Coulomb line of the real soil and of the discrete element soil model

Linear trendlines were fitted to the measurement and to the calculated simulation values with high R^2 value of 94% and 99%, respectively using the Ordinary Least Squares available in Microsoft Excel 2016 software. Figure 5 shows very small differences between the Mohr-Coulomb lines of simulation and measurement therefore similarities in the values of cohesion and internal friction angle can be expected as well. The relative error of these properties (*RE*) was calculated using the following formula:

$$RE = \frac{N_{simulation} - N_{measurement}}{N_{measurement}} \cdot 100$$

where $N_{simulation}$ is the value of the mechanical properties from the simulation and $N_{measurement}$ is the value of the properties from the measurement. The results were summarized in Table 3.

Table 3: The calculated mechanical properties of the soil model and of the real soil.

	Discrete element soil model	Real soil	Relative error (%)
Cohesion (c) (Pa)	3.82e+4	3.66e+4	4.44
Internal friction angle (φ) (°)	17.06	17.89	4.66

According to Table 3 the mechanical properties of real soils and the same properties of soil models are close to each other. The relative errors are under the value of 5% which means that the presented calibrational process can be used to calibrate the contact properties of the Hertz-Mindlin model in case of direct shear simulations.

In the future additional calculations should be performed to obtain more similar shear force-displacement curve in simulations to that of real measurements especially in range of small displacements. To do this it is necessary to invenstigate to effect of changing the Poisson ratio on the value of shear force.

ACKNOWLEDGEMENT

The authors will gratefully acknowledge the assistance of the staff of Institute of Process Enginnering, Szent István University of Gödöllő to provide their direct shear apparatus for the measurements.

REFERENCES

Chi, L. and Kushwaha, R. L. 1990. “A non-linear 3-d finite element analysis of soil failure with tillage tools.” *Journal of Terramechanics*, 27(4), 343-366.

Cundall, P. A. and Strack, O. D. L. 1979. “Discrete numerical model for granular assemblies.” *Geotechnique*, 29(1), 47-65.

Dirgéliené, N.; Amsiejus, J.; Norkus, A. and Skuodis, S. 2014. “Comparison of sandy soil shear strength parameters obtained by various construction direct shear apparatuses.” *Archieves of Civil and Mechanical Engineering* 14, 327-334.

Hertz, H. 1882. "On the contact of elastic solids." *J. reine und angewandte Mathematik* 92, 156-171.

Knuth, M. A.; Johnson, J. B.; Hopkins, M. A.; Sullivan, R. J. and Moore, J. M. 2012. “Discrete element modeling of a Mars Exploration Rover wheel in granular material.” *Journal of Terramechanics*, 49, 27-36.

McKyes, E. 1985. *Soil Cutting and Tillage*. Elsevier, New York, USA.

Mindlin, R. D. 1949. "Compliance of elastic bodies in contact." *Journal of Applied Mechanics* 16, 259-268.

Mindlin, R. D. and Deresiewicz H. 1953. "Elastic spheres in contact under varying oblique forces." *ASME*, September, 327-344.

Mouazen, A. M. and Neményi, M. 1998. “A review of the finite element modelling techniques of soil tillage.” *Mathematics and Computers in Simulation*, 48 (1), 23-32.

Nakashima, H.; Fujii, H.; Oida, A.; Momozu, M.; Kanamori, H.; Aoki, S.; Yokoyama, T.; Shimizu, H.; Miyasaka, J. and Ohdoi, K. 2010. “Discrete element method analysis of single wheel performance for a small lunar rover on sloped terrain.” *Journal of Terramechanics*, 47, 307-321.

Potyondy, D. O.; Cundall, P. A. 2004. “A bonded-particle model for rock”. *International Journal of Rock Mechanics & Mining Sciences*, 41, 1329-1364.

Sadek, M. A.; Chen, Y. and Liu, J. 2011. “Simulating shear behavior of a sandy soil under different soil conditions.” *Journal of Terramechanics*, 48, 451-158.

Sitkei, Gy. 1967. *Mezőgazdasági gépek talajmechanikai problémái* (*The soil mechanics problems of the agricultural machines*). Akadémiai Kiadó, Budapest, Hungary (in hungarian).

Tamás, K.; Jóri, J. I. and Mouazen, A. M. 2013. “Modelling soil-sweep interaction with discrete element method.” *Soil & Tillage Research*, 134, 223-231.

Terzaghi, K. 1943. *Theoretical Soil Mechanics*. John Wiley and Sons, New York, USA.

AUTHOR BIOGRAPHIES

KRISZTIÁN KOTROCZ was born in Salgótarján, Hungary and went to the Budapest University of Technology and Economics, where he studied mechanical engineering and obtained his MSc degree in 2012. After that he started his PhD studies and worked in the Budapest University of Technology and Economics, Department of Machine and Product Design where he is an assistant lecturer curretnly. His research area id soil modelling using discrete element method. His e-mail address is: kotrocz.krisztian@gt3.bme.hu and his Web-page can be found at http://gt3.bme.hu/en.

GYÖRGY KERÉNYI studied agricultural machine design at Szent István University, Gödöllő and after that he went to Budapest University of Technology and Economics, where he obtained his PhD degree in 1997. Currently he is an associate professor and deputy head of Department of Product and Machine Design in the same institution and his research topic is numerical methods in agricultural machine design. His e-mail address is: kerenyi.gyorgy@gt3.bme.hu and his Web-page can be found at http://gt3.bme.hu/en.

MODELLING PREFERENCE TIES AND EQUAL TREATMENT POLICY

Kolos Cs. Ágoston
Department of Operations Research
and Actuarial Sciences
Corvinus University of Budapest
H-1098, Budapest, Fővám tér 8., Hungary
Email: kolos.agoston@uni-corvinus.hu

Péter Biró
Hungarian Academy of Sciences
H-1112, Budaörsi út 45, Budapest, Hungary
Department of Operations Research and
Actuarial Sciences
Corvinus University of Budapest
H-1098, Budapest, Fővám tér 8., Hungary
Email: peter.biro@krtk.mta.hu

KEYWORDS

College admission problem; Integer programming; preference ties; equal treatment policy

ABSTRACT

The college admission problem (CAP) has been studied extensively in the last 65 years by mathematicians, computer scientists and economists following the seminal paper of Gale and Shapley (1962). Their basic algorithm, the so called deferred acceptance mechanism always returns a student optimal stable matching in linear time, and it is indeed widely used in practice. However, there can be some special features which may require significant adjustments on this algorithm, or the usage of other techniques, in order to satisfy all the objectives of the decision maker. The college admissions problem with ties and equal treatment policy is solvable with an extension of the Gale and Shapley algorithm, but, if there are further constraints, such as lower quotas, there exist no efficient way to find a stable solution. Both of these features are present in the Hungarian higher education matching scheme and a simple heuristic is used to compute the cutoff scores. Integer programming is a robust technique that can provide optimal solutions even when we have multiple requirements. In this paper we develop and test a new IP formulation for finding stable solutions for CAP with ties and equal treatment policy. This formulation is more general than the previously studied ones, and it has better performance, as we demonstrate with simulations, mostly because of its pure binary nature.

INTRODUCTION

The so-called Gale-Shapley algorithm (Gale and Shapley 1962) became widely used in the past decades for solving various kinds of two-sided matching problems under preferences, such as resident allocation to hospitals, college admissions (CAP in the following) and school choice. This algorithm was also the core mechanism studied in the corresponding theoretical research, see a comprehensive overview on the algorithmic (Manlove 2013) and game theoretical aspects (Roth and Sotomayor 1990) of this topic. However, in many practical applications there may be special features which make the problem more challenging to solve. For instance, in the Hungarian higher education admission scheme these special features are the presence of ties, lower quotas, common quotas and paired applications, where each of the latter three features makes the problem NP-hard to solve (Biró et al. 2010).

In a recent paper we formulated integer linear programmes to tackle these special issues one by one (Ágoston et al. 2016). One important finding of that paper was that even an NP-hard problem, such as the college admission with lower quotas, can be possible to solve with IP techniques for really large instances if we use some clever preprocessing heuristics. However, some of our formulations turned out to be inefficient to solve large instances, and we have not considered the cases when multiple special features are present at the same time. In this paper we give a new formulation for the special feature of ties with equal treatment property and we show that this new formulation, based only on binary variables, is easier to solve for large instances than the other natural formulation with cutoff scores. This will give us a chance to tackle also the combined case with IP technique, where both ties and lower quotas are present, which is relevant in the Hungarian application.

In CAP, students give their strict preferences over colleges (or programmes) where they apply. Universities rank the applicants according to their scores. Scores are usually based on secondary school grades and entrance exams. It may happen that two (or more) applicants have the same score for a college, so a *tie* may occur in the ranking of the university. There are many ways around the world, how the ties are handled see (Biró and Kiselgof 2015). In most countries the ties are broken in some way. In Spain the scoring method is very fine, so no tie may occur. In Ireland a random number is generated for each student and preferences of universities are determined by considering both the scores and the generated random number. In Turkey the ties are broken according to the age of the students. However, in some other countries the equal treatment policy is used, which means that the applicants with

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

the same score have to be treated equally, either all of them are accepted or all of them are rejected. In Hungary the college quotas cannot be exceeded, so the last group of students with the same score whose admission would lead to the violation of the quota is always rejected. In Chile the equal treatment policy is more permissive, college quotas can be exceeded with the last group of students with equal score. The Gale-Shapley algorithm can be extended for finding student-optimal and student-pessimal stable solutions in case of ties for both the Hungarian policy (Biró 2008) and for the Chilean policy (Biró and Kiselgof 2015).

In (Ágoston et al. 2016) we formulated an IP for finding stable matchings for the Hungarian policy with using the cutoff scores as variables. This formulation turned out to be inefficient when solving large instances. In this paper we propose a new formulation using only binary variables on the applications. We show that with an appropriate objective function this IP can be used for finding the student-optimal stable solution. We also give techniques to reduce the size of the problem and thus speed up the solution. Our hope is that by building on this new formulation we will also be able to solve the combined case with ties and lower quotas for realistic instances.

Related works

In the classical assignment problem the two sides of the market has to be assigned in order to maximize or minimize the total utility or cost, respectively. Both LP solutions and efficient algorithms have been used for solving this problem, e.g. the so-called Hungarian method (Kuhn 1955). In the corresponding stable matching problems, which is called as stable marriage problem in the one-to-one case and college admissions problem in the many-to-one case, we build our LP formulations on the assignment problem, but we need to add constraints that provide stability with respect to the agents' submitted preferences.

The stable marriage problem was first investigated by Vande Vate (1989) as an LP problem. He formulated an LP and he proved the integrality property, i.e. that that extreme points of the feasibility set are all integers and they correspond to the set of stable matchings. Rothblum (1992) gets the same result for a more general problem. He defines another stability constraint which can be generalized to CAP. Baïou and Balinski (2000) investigate this formulation for CAP, and they show that this formulation allows fractional solutions to be extreme points, but they propose an alternative formulation which satisfies the integrality property. The formulation given in (Baïou and Balinski 2000) is a pure binary problem, i.e. all of the decision variables are binary. Also in this paper we can find a different formulation for CAP which satisfies the integrality property. However this second formulation has a large number of constraints.

In case of ties, the literature mostly focuses on the concept of weak stability, where the rejection of a student is considered fair if the quota of the college is filled with students of greater or the same score. Here, the problem of finding a maximum size weakly stable matching is NP-hard (Manlove 2013), but still this solution concept is used e.g. in the resident allocation in Scotland. IP technique was developed for solving this problem in Irving and Manlove (2009).

The concept of stable matching with ties and equal treatment property corresponding to the Hungarian policy was defined in (Biró 2008) and studied in (Biró and Kiselgof 2015) and (Fleiner and Jankó 2014). A mixed integer LP (MILP) with the cutoff scores being the variables was formulated in (Ágoston et al. 2016)

PRELIMINARIES

In the college admissions problem (CAP) let $A = \{a_1, \dots a_n\}$ be the set of applicants and let $C = \{c_1, \dots, c_m\}$ be the set of colleges. Let u_j denote the upper quota of college c_j. Regarding the preferences and priorities, let r_{ij} denote the rank of college c_j in a_i's preference list, meaning that a_i prefers c_j to c_k if $r_{ij} < r_{ik}$. For the sake of simplicity, an applicant's most preferred college gets rank 1, the second gets rank 2, and so on. The maximum of r_{ij} is denoted by $\overline{r}$. Let s_{ij} be an integer representing the score of a_i at college c_j, meaning that a_i has priority over a_k at college c_j if $s_{ij} > s_{kj}$. In the classical CAP by Gale and Shapley the scores of the students are different at every college, so the rankings by the universities are strict. We denote the set of applications by E. A *matching* M is a subset of applications, such that every student is allocated to at most one college and the number of allocated students at a college is less than or equal to its quota. A college is said to be *saturated* in a matching if its quota is filled. A matching is *stable* if for any unselected application either the student is allocated to a better college of her preference or the college filled its quota with better students. In the classical CAP a stable matching can be described with a set of cutoff scores $\overline{t} = [t_1, \dots t_m]$, where each student is allocated to the best college in her list where she achieves the cutoff score. A natural choice for such a set of cutoff scores for a stable matching is when t_j is the lowest score of the allocated students at c_j if c_j is saturated and zero otherwise. The relation of the set of cutoff scores and stable matchings for the classical CAP was studied in details in (Azavedo and Leshno 2016).

If ties are allowed in the scores of the students and we use the equal treatment policy with no quota violation then stability of a matching can be defined through cutoff scores. A set of cutoff scores is *H-stable* if no college with positive cutoff score can decrease its cutoff score without violating its quota in the corresponding matching. A matching is H-stable if it corresponds to a H-stable set of cutoff scores. Note that if no ties occur then this stability definition is equivalent to the Gale-Shapley stability for CAP. Let us refer to the CAP with ties as CAPT, and the above described model as H-CAPT, that is the college admission problem with ties and equal treatment policy with higher score-stability. Biró and Kiselgof (2015) proved that the natural extension of the Gale-Shapley student proposing algorithm always produces a student-optimal H-stable solution, which corresponds to a set of minimal cutoff scores, i.e. there is no other H-stable set of cutoff scores where even one college could have a lower cutoff score. Note that this also applies then in this case the number of admitted students is as high as possible and that every student gets the best possible place of her preference. However, we shall also mention that this mechanism does not remain strategy-proof, as in the classical model.

INTEGER PROGRAMMING FORMULATIONS

When describing a linear programme for CAP and H-CAPT, we introduce binary variables $x_{ij} \in \{0,1\}$ for each application by applicant a_i to college c_j, as a characteristic function of the matching, where $x_{ij} = 1$ corresponds to the case when a_i is assigned to c_j. The feasibility of a matching can be ensured with the following sets of constraints.

$$\sum_{j:(a_i,c_j)\in E} x_{ij} \leq 1 \;, \qquad \forall a_i \in A \tag{1}$$

$$\sum_{i:(a_i,c_j)\in E} x_{ij} \leq u_j \;, \qquad \forall c_j \in C \tag{2}$$

For CAP, the stability can be provided with the following set of constraints (see for example Baïou and Balinski 2000).

$$\left(\sum_{k:r_{ik}\leq r_{ij}} x_{ik}\right)\cdot u_j + \sum_{h:(a_h,c_j)\in E, s_{hj}\geq s_{ij}} x_{hj} \geq u_j \qquad \forall (a_i,c_j)\in E \tag{3}$$

Constraints (1) and (2) are called feasibility constraints and constraints (3) are called stability constraints. The feasible set S_s contains all nonnegative solutions which satisfies constraints (1), (2) and (3). Integer points of S_s corresponds to stable matching for CAP, and vica versa. However, note that this formulation does not have the integrality property, since S_s may have non-integer extreme points, as illustrated in (Baïou and Balinski 2000). Objective function is arbitrary in this case, but if we minimize or maximise the sum of the ranks of the allocated students then we can obtain the student-optimal and student-pessimal solutions, respectively.

If we consider the possibility of ties in the rankings (CAPT) then constraints (3) ensures only the so-called weak stability condition of the matching, where the rejection of an application can be explained with the saturation of the quota with students at least as good as the student concerned. Example 1 describes why a weakly stable solution for CAPT is not necessary H-stable as well.

Example 1: We have one college and two applicants, who have the same score. The upper limit for the college is one. In this case none of the applicants can be assigned to the college for H-CAPT, but according to constraints (1), (2) and (3) it is a feasible solution (and thus weakly stable for CAPT) if we allocate one of them to the college.

In case of H-CAPT, the college quota will not necessarily be full in a H-stable matching, even if there are more than enough first applications submitted to the college, as we have also seen in Example 1. So stability constraints (3) are not appropriate in case of ties. We formulated another model in (Ágoston et al. 2016), where the cutoff scores are the main variables, but that model is not a pure binary model and according to our simulations the running times of the solver for that model are high even for a small instances.

In our new formulation that we investigate in this paper, we keep feasible constraints (1) and (2) and change (3) to

$$\sum_{k:r_{ik}\leq r_{ij}} x_{ik} \geq x_{hj} \quad \forall (a_i,c_j),(a_h,c_j)\in E, s_{ij}\geq s_{hj} \tag{4}$$

Intuitively the above constraint means that if a student a_h is allocated to college c_j then every student a_i, who has a score at c_j at least as high as a_h has, must also be allocated to c_j or to a better college of her preference. Let S_p contain all nonnegative solutions that satisfy constraints (1), (2) and (4).

Example 2 shows the difference between S_s and S_p for CAP.

Example 2: We have two applicants, a_1 and a_2, and three colleges ,c_1,c_2 and c_3. The first applicant's preferences are $c_1 \succ c_2 \succ c_3$; the second applicant's preferences are $c_3 \succ c_1$. The first applicant' scores for the three colleges are $(1;2;2)$, the scores of the second applicant are $(2;1;1)$. The quotas are 2 for c_1 and 1 for c_2. The set S_s contains only one point: $x_{11} = 1$ and $x_{23} = 1$, whilst S_p has many extremal points:

$$\begin{pmatrix} x_{11}\\ x_{12}\\ x_{13}\\ x_{23}\\ x_{21}\end{pmatrix} = \begin{pmatrix} 1\\0\\0\\1\\0\end{pmatrix};\begin{pmatrix} 0\\0\\0\\0\\0\end{pmatrix};\begin{pmatrix} 0\\1\\0\\0\\0\end{pmatrix};\begin{pmatrix} 0\\0\\1\\0\\0\end{pmatrix}$$

$$\begin{pmatrix} x_{11}\\ x_{12}\\ x_{13}\\ x_{23}\\ x_{21}\end{pmatrix} = \begin{pmatrix} 0\\0\\0\\0\\1\end{pmatrix};\begin{pmatrix} 0\\1\\0\\1\\0\end{pmatrix};\begin{pmatrix} 0\\0\\0.5\\0.5\\0\end{pmatrix}$$

Our first remark is that constraints (3) do not yet ensure the stability of a matching, since the empty matching is also a feasible solution. With regard to CAP, the integer solutions in S_p correspond to the so-called *envy-free matchings* defined in (Wu and Roth 2016). In such a matching no student has a *justified envy* over another student, meaning that there exist no two students a_i and a_h, and a college c_j such that a_i is matched to c_j, but a_h has a higher score there than a_i and she is matched to a less preferred college (or no college at all). In other words, a matching is envy-free for an instance of CAP if and only if it is a stable matching for the adjusted quotas, where the quota of each college is equal to the number of students assigned there in the matching.

The same statement holds for H-CAPT, the integer solutions in S_p correspond to H-stable matchings with regard to adjusted quotas, where the quota of each college is equal to the number of students assigned there in the matching. Let us call these matchings *H-envy-free matchings*, that is envy-free matchings with regard to the equal treatment property. In this definition a student would also feel justified envy if she is no assigned to a college and any better of her preference, whilst a student with the same score is admitted there.

In the following two propositions we summarize the connections between the extreme points of S_p, the envy-free matchings, and the matchings induced by cutoff scores.

Proposition 3: For an instance of CAP, the set of envy-free matchings are the integer extreme points of S_p and these are also the feasible matchings that can be induced by a set of cutoff scores.

Proof: We have already seen that the envy-free matchings are the integer extreme points of S_p, and vice versa. Now we suppose that M is an envy-free matching and we show that it can be induced by a set of cutoff scores. Indeed, if we choose the score of the weakest admitted student at each college to be the cutoff score of that college then M is induced by these cutoff scores. In the other direction, let t be a set of cutoff scores and let M be the induced feasible matching (which does not violate the quotas of the colleges). It is easy to see that M is an envy-free matching by definition, since the students admitted to a more preferred college from a student a_i's perspective have all higher scores there than a_i has. ■

Proposition 4: For an instance of CAPT, the set of H-envy-free matchings are the integer extreme points of S_p and these are also the feasible matchings that can be induced by a set of cutoff scores.

Proof: The proof is the same as above. ■

The structure of envy-free matchings for CAP has been studied in (Wu and Roth 2016). In particular they showed that the student-optimal stable matching for the instance of CAP is also student-optimal in the set of envy-free matchings. Regarding CAPT, we know that there exists a student-optimal H-stable matching, obtained by the extended Gale-Shapley algorithm (Biró and Kiselgof 2015). In the following proposition we show that this matching is also optimal for the set of H-envy-free matchings.

Proposition 5: For an instance of CAPT, the student-optimal H-stable matching is also optimal for the students in the set of H-envy-free matchings.

Proof: Suppose for a contradiction that there exists a matching, where at least one student gets a better college than in the student-optimal H-stable matching, M_s, and suppose also that it is not Pareto-dominated with any other such matchings. Let the corresponding set of cutoff scores be $\overline{t}$. This matching cannot be a H-stable matching, as that would contradict with the student-optimality of M_s in the set of H-stable matchings, thus there must exist at least one college, c_j, where we can increase the cutoff score by admitting more students, but without violating the quota. Let this new set of cutoff scores be denoted by $\overline{t}'$. The induced matching by $\overline{t}'$ is feasible, and by Proposition 4 it is also H-envy-free. Moreover, every new student admitted at c_j improved their assignment compared to M and nobody else received a worse college, so M was not Pareto optimal among the set of H-envy-free matchings, a contradiction. ■

Now, we describe how one can obtain the student-optimal H-stable matching using an appropriate objective function.

Lemma 6: Let we set objective function as

$$\max \sum_{k=1}^{\overline{r}} \alpha_i \sum_{r_{ij}=k} x_{ij} \ , \qquad (5)$$

where $\alpha_1 > \alpha_2 > \cdots > \alpha_{\overline{r}}$. The optimal integer solution of (5) subject to constraints (1), (2), (4) and $x_{ij} \geq 0$ is the student-optimal H-stable matching.

Proof: From Proposition 4 we know that any feasible solution corresponds to a H-envy-free matching. Proposition 5 implies that the student-optimal H-stable matching, M_s is also optimal in the set of H-envy-free matchings. Therefore there is no other H-envy-free matching where even one student can get a better college, so the objective function (5) must be maximized for M_s. ■

NUMERICAL RESULTS

In this section we investigate how our proposed IP model behaves numerically. We replicate some results from (Ágoston et al. 2016) and extend it with new results.

For the numerical modeling we used a desktop computer with 2.33 GHz Intel Pentium processor and 6 GB RAM. Operating system is Windows 7 Enterprise.

We used the GLPK software (version 4.55 version) for solving IP problems. We kept the default parameter setting except where we explicitly mention it.

We used randomly generated samples for testing the IP models: we had n applicants and k colleges, each applicant choosing five colleges uniformly at random without replacement. So there are about $\frac{n}{k}$ first place applications at each college. We fixed the quotas at $\frac{n}{2k}$, so every quota is expected to be full. Scores are integers distributed randomly between 0 and $\overline{s}$.

First we consider the case where there are no ties ($\overline{s}$ is quite large, e.g. 10000).

Table I shows how large problems can be solved with the 'Stable Cutoff Scores IP model' (SCS-IP) defined in (Ágoston et al. 2016), where the cutoff scores are the main variables. Table II shows the running times for the basic IP model, called 'Basic College Admission IP' (BCA-IP), with objective function (5) and constraints (1), (2) and (3). We can see that considerably larger

TABLE I: Running times of SCS-IP for CAP

n	k	running time (sec)
20	10	0.2
40	10	4.0
60	10	81.3
80	10	1443.8

TABLE II: Running times of BCA-IP for CAP

n	k	running time (sec)
100	10	0.0
500	20	0.3
2500	30	28.9
7500	40	588.7

TABLE IV: Summarizing results for the speeding-up process for CAP instances. The 'OEF-IP' column shows the running time for the OEF-IP model, the same as in Table III. The '#filt.assign' columns shows how many applications can be filtered out. The 'filtered' column shows the running time of the solver for OEF-IP after filtering.

n	k	OEF-IP	#filt.assig.	filtered
160	20	1032.2	461	3.5
200	20	4868.6	573	12.7
240	20	> 3600	669	28.9
280	20	> 3600	784	45.6
320	20	> 3600	899	31.6
360	20	> 3600	987	52.3
400	20	> 3600	1083	171.6

problems can be solved with BCA-IP than with SCS-IP. However, the BCA-IP model cannot be simply extended for CAPT instances.

Table III shows the running times for our new IP model, called 'Optimal Envy-Free IP model' (OEF-IP), where we maximize (5) subject to constraints (1), (2) and (4). The running times are larger for OEF-IP than for BCA-IP, but lower than for SCS-IP. We would like to emphasize that for H-CAPT we cannot use BCA-IP, so OEP-IP remains the best approach for that.

Speeding up the solution by filtering

As we see in Table III our new formulation, OEF-IP, performs better than its alternative, but the difference is still very small. However, we have some tools that can speed up the solution, as we describe below.

Among the applications there are many which are surely not possible to accept. There are known techniques in the literature that can filter out impossible pairs (e.g. defined in (Irving and Manlove 2009), and applied in (Kwanashie and Manlove 2014)). We propose another way, which uses the known properties of the simplex method. We check each variable, one by one, by setting $x_{ij} = 1$ and solving the LP relaxation of the problem. If there is no feasible solution then we know that variable x_{ij} has to be zero.

When we checked all the variables once and could set at least one to be zero then we repeat this process for the remaining variables. We stop this filtering process when no variable can be set to be zero in a round, thus after checking each variable at most m times (where m is the number of applications). With the following instance of CAPT in Example 7 we illustrate why multiple rounds may be useful in this filtering process.

Example 7: We have 3 colleges (c_1, c_2 and c_3) and three applicants (a_1, a_2 and a_3). All the three applicants have score of 3 to c_1, score of 2 to c_2 and score of 1 to c_3. Applicant a_1 submitted application only to c_3; the other two submitted to all the three colleges, their preference orders are the same: $c_1 \succ c_2 \succ c_3$. The quota is 1 for all the three colleges. In the first round we first investigate whether $x_{1,3}$ can be one, and we find that indeed there exists such a feasible solution for the LP, namely $x_{1,3} = 1$, $x_{2,1} = 0.5$, $x_{2,2} = 0.5$, $x_{3,1} = 0.5$ and $x_{3,2} = 0.5$. However, when we check $x_{2,1}$, we see that it is not possible to set it 1, since in this case $x_{3,1}$ have to be 1 as well and we would exceed the quota for c_1. Analogously it will be clear that $x_{3,1}$ has to be zero as well. But if both $x_{2,1}$ and $x_{3,1}$ are equal to 0 then $x_{1,3}$ cannot be 1 either, so in the second round of our filtering process we can also set $x_{1,3}$ to be 0.

TABLE III: Running times of OEP-IP for CAP

n	k	running time
80	20	23,0
120	20	322.5
160	20	1032.2
200	20	4868.6

In our first simulation we simply investigate how much the running time decreases by using the filtering process (without deleting redundant constraints) for CAP instances. The running time decreases dramatically, as seen in Table IV, column 'filtered'.

Multiple rounds of the filtering process turned out to the useful indeed. For example, in case of the last row of Table IV in the second iteration we filter out another 90 variables and for this reason the running time decreases to 2.0 sec.

We now turn our attention to the H-CAPT model. We decrease $\overline{s}$ to 10, and then to 5. As we see in Table V and Table VI the running times for solving such problems are higher. This is because if there are no ties then the optimal value of the LP relaxation problem equals with the optimal value of the IP problem more often. Therefore it is more likely that we find an integer solution to the problem quickly, and in this case we can bound all the open branches. However, if there are ties, then the optimal solution of IP is significantly smaller than the optimal solution of the LP relaxation (see 'Adm. appl.' column in Table V and Table VI) resulting in longer branch-and-bound processes.

TABLE V: Running times and the number of admitted applicants for H-CAPT instances with $\overline{s} = 10$, $k = 20$, using OEF-IP with filtering.

n	without filt.	#filt.assig.	with filt.	Adm. appl.
80	78.0			30
120	1600.4			45
160	6226.7	532	0.1	56
200	> 3600	648	0.2	75

TABLE VI: Running times and the number of admitted applicants for H-CAPT instances with $\overline{s} = 5$, $k = 20$ using OEF-IP with filtering.

n	without filt.	#filt.assig.	with filt.	Adm. appl.
80	16.5			13
120	201.6			18
160	1082.4	602	0.1	24
200	> 3600	731	0.2	30

Perhaps at first sight it may seem strange that running times in Table VI are smaller than running times in Table V. However, as the number of applicants with the same score increases, we have more and more applications that we can filter out, which decreases the number of nodes in the branch-and-bound tree. As we can see, solving the case for filtered problem takes less time in case of $\overline{s} = 5$ than in case of $\overline{s} = 10$.

CONCLUSIONS

In this paper we presented a new IP formulation for college admission problem with ties under the equal treatment policy, a case present in the Hungarian higher education admission scheme. Our new IP formulation is binary, which turned out to be significantly easier to solve than our previous formulation (Ágoston et al. 2016) with integer variables. The constraints in the problem do not ensure that every (integer) solution of the problem will correspond to a stable allocation, but with an appropriate objective function the student-optimal stable solution can be obtained, as we proved. We also presented methods on how to speed-up the solution of our new formulation. Although we were able to solve much larger instances than before, it still requires further research to solve the H-CAPT problem for a real instance of the Hungarian scheme with around 100 thousands applicants.

Furthermore, we shall also investigate how one can define a fair solution for the case when besides the issue of ties the lower quotas are also present, as it occurs in Hungary. For this case with multiple special features, where one of the features (lower quotas) makes the problem NP-hard, the usage of IP technique can be especially useful. Note that in our previous work (Ágoston et al. 2016) we were indeed able to solve the feature of lower quotas successfully with IP techniques after an efficient filtering process. Our most important future plan is to formulate a combined IP for solving the H-CAPT problem with lower quotas, and develop filtering techniques that can make our approach work for large instances.

ACKNOWLEDGEMENT

Péter Biró was supported by the Hungarian Academy of Sciences under Momentum Programme (LP2016-3/2016), and by the Hungarian Scientific Research Fund - OTKA (no. K108673).

REFERENCES

Ágoston, K. Cs., Biró, P., & McBride, I. (2016). Integer programming methods for special college admissions problems. *Journal of Combinatorial Optimization.* Vol. 32, Iss. 4, pp. 1371–1399.

Azevedo, E.M., & Leshno, J.D. (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy.* Vol. 124(5), pp. 1235–1268.

Baïou, M., & Balinski, M. (2000). The stable admissions polytope. *Mathematical Programming.* Vol. 87(3), Ser. A, pp. 427–439.

Biró, P. (2008). Student Admissions in Hungary as Gale and Shapley Envisaged. *Technical Report*, no. TR-2008-291 of the Computing Science Department of Glasgow University.

Biró, P., Fleiner, T., Irving, R.W., & Manlove, D.F. (2010). The College Admissions problem with lower and common quotas. *Theoretical Computer Science*, 411:3136–3153.

Biró, P., & Kiselgof, S. (2015). College admissions with stable score-limits. *Central European Journal of Operations Research.* Vol. 23(4), pp. 727–741.

Fleiner, T., & Jankó, Zs. (2014). Choice Function-Based Two-Sided Markets: Stability, Lattice Property, Path Independence and Algorithms. *Algorithms.* Vol. 7(1), pp. 32–59.

Gale, D., & Shapley, L. S. (1962). College Admissions and the Stability of Marriage. *American Mathematical Monthly.* Vol. 69(1). 915.

Irving, R. W., & Manlove, D. F. (2009). Finding large stable matchings. *ACM Journal of Experimental Algorithmics.* Vol. 14, 1.2. pp. 1–27.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly.* 2/1-2, pp. 83–97.

Kwanashie, A., & Manlove, D. F. (2014). An Integer Programming Approach to the Hospitals/Residents Problem with Ties. *Operations Research Proceedings 2013.* Pp 263–269.

Manlove, D. F. (2013). Algorithmics of Matching Under Preferences Series on Theoretical Computer Science vol. 2, World Scientific.

Roth, A.E., & Sotomayor, M. (1990) Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis, Cambridge University Press.

Rothblum, U.G. (1992). Characterization of stable matchings as extreme points of a polytope. *Mathematical Programming*, Vol. 54(1, Ser. A), pp 57–67.

Vate, J. E. Vande (1989). Linear programming brings marital bliss. *Operations Research Letters.* Vol. 8(3). pp. 147–153.

Wu, Q., & Roth, A. E. (2016). The lattice of envy-free matchings. Mimeo

AUTHOR BIOGRAPHIES

KOLOS CS. ÁGOSTON graduated as an actuarist. He wrote his PhD thesis in insurance markets. He is now a full time lecturer at the Department of Operations Research and Actuarial Sciences, Corvinus University of Budapest. He teaches various subjects in operational research and actuarial sciences. His research topics belong to optimization problems such as cash management, cutting problems and recently college admission problem. His e-mail address is `kolos.agoston@uni-corvinus.hu`

PÉTER BIRÓ has received his PhD in mathematics and computer science at Budapest University of Technology in 2007 and then he was a postdoc at the Computer Science Department of Glasgow University for three years. In 2010 he joined the game theory research group at the Institute of Economics of the Hungarian Academy of Sciences as a research fellow, and he has been working there since, except for a one year leave in 2014 when he was a visiting professor at the Economics Department of Stanford University. Currently he is the head of the Momentum research group on Mechanism Design at the Institute of Economics and he is also a part-time lecturer at the Department of Operations Research and Actuarial Sciences, Corvinus University of Budapest. His e-mail address is `peter.biro@krtk.mta.hu`

CALIBRATION OF RAILWAY BALLAST DEM MODEL

Ákos Orosz, János P. Rádics, Kornél Tamás
Department of Machine and Product Design
Budapest University of Technology and Economics
Műegyetem rkp. 3., H-1111, Budapest, Hungary
E-mail: orosz.aakos@gmail.com

KEYWORDS
Railway ballast, DEM, Calibration, Hummel device

ABSTRACT

The ballast of the railway track is constantly changing due to dynamic forces of train traffic that results in the crushing of the rocks. The feasibility to simulate each particle makes the discrete element method (DEM) suitable for the task.

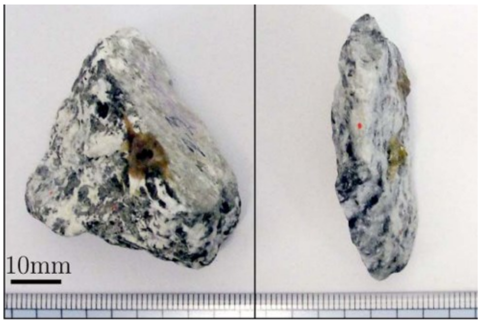
Different DEM particle models are introduced in our paper and the adequate one was chosen. This new method is based on crushable convex polyhedral elements and random shape generation via Voronoi tessellation and is implemented in the Yade DEM simulation software.

The particle geometry is validated via comparing the simplified shape of natural rough rocks and the randomly generated ones. A 3D scanner was used to digitize the natural rocks. The crushing behaviour is tested as well. The validation of interaction laws and the calibration of the micro parameters is necessary to create a DEM material model with a realistic behaviour. In the calibration process Hummel device is modelled, which provides well measurable parameters for comparing simulation and measurement results.

INTRODUCTION

In a limited number of cases, it is possible to model the railway track ballast as a continuum with the use of finite element method (Shahraki et al. 2015), however this approach does not give information about many aspects. Therefore a new approach should be utilised to simulate the railway track ballast behaviour more realistically.

In our research, the rocks of the ballast are classified into two groups based on their geometry: equant (Figure 1,a) and flat (Figure 1,b). It is well known that both of the shapes mentioned above have to be represented in the railway track ballast to endure the forces effectively. Moreover they have an optimal ratio, which is currently estimated by routine of practice. The determination of its exact value would provide many benefits.



0a) b)

Figure 1: a) Equant and b) Flat Rocks
(Asahina and Taylor 2011)

Result of the dynamic forces of the periodical loads is the fragmentation of the rocks, which changes the ratio of the different geometry types. This has a great effect on the loadability of the aggregate. That is one of the reasons to perform the complete maintenance of the railway track. A main objective of our research is to investigate this phenomenon and to determine the period between maintenances more precisely through simulating the long term behaviour of the ballast.

The full maintenance process has different stages completed with different machines. The appropriate model and simulation of the railway track ballast can be used to improve efficiency of finding the optimal machine settings and tool geometries.

Instead of the continuum approach, there is a need to simulate each rock to achieve a realistic simulation, whereby the previous questions can be answered. It is also important to take efforts in the modelling of rock breakage, because crushing of particles is necessary to bring simulation data closer to experimental data (Eliaš 2014). This makes the discrete element method suitable for the task.

This paper represents a brief introduction of the discrete element method, choosing of a discrete element material model based on particle geometry and crushability aspects, the mechanical basics of this model, a solution for random geometry generation and validation of this technique. Also an approach is introduced for calibrating the model with the device used for the Hummel measurement.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

The Discrete Element Method

The discrete element method is a numerical technique, where the simulated material is made of particles (or elements), with independent motional degrees of freedom. Interactions can be created and erased between the elements (Bagi 2007). For this reason the behaviour of the volume modelled with discrete element method depends on the definition of the elements and interaction laws. The sum of these features is called the discrete element material model (or micromechanical properties). The micromechanical properties have direct effect on the element-level behaviour, which results in the measurable, macromechanical behaviour. The aim of the study is to reproduce the measurable characteristics of the aggregate which can be achieved with a proper material model.

The exact relation between micro- and macromechanical behaviour is missing in most cases, and the reasonable parameters can only be reached by the calibration of the material model. This is done by comparing the proper measurement with its simulation, and modifying the properties of the material model until its macromechanical behaviour reproduces the test with acceptable approximation. In our research the Hummel measurement device is used to calibrate the material model.

MATERIAL MODEL

The first difficulty of creating the model was the decision about the shape of the particles. The simplest solution is to create a volume made of spheres. This also gives an advantage in computational speed. However it is known that many type of aggregates such as sand – or in our case the railway track ballast –cannot be modelled properly with spheres because of the rolling contact surfaces versus the actual friction sliding of real particle surfaces (Lane et al. 2010).

To keep the advantages of spheres but to also be able to use model elements with more complex shape, the so called clumps are used. The clumps are made of spheres with rigid connections between them. Example is shown on Figure 2,a. (Coetzee and Nel 2014).

The clump particle is still smooth and is missing the sharp corners and edges. Particle approximation can be improved by reducing the element size of the clump particle, but the highly increasing computation time have to be accepted. Therefore an extensive effort was made to simulate the aggregate with polyhedral elements. An example of the shape of these particles is shown in Figure 2,b .

There are different programs e.g. Grains3D (Wachs et al. 2012), BLOKS3D (Huang and Tutumler 2011) for handling polyhedral particles, but these are hardly accessible softwares and there is no opportunity to make modifications. Jan Eliaš also created a polyhedral particle model (Eliaš 2014), which features the possibility of crushing particles with low computation need and is implemented in the Yade open source DEM software. The available source code creates the chance to make modifications and improvements in the embedded material models, which is a significant benefit. The built-in crushing effect and the possibility of improvements are the main reasons that the crushable polyhedral material model (Eliaš 2014) was selected for further studies.

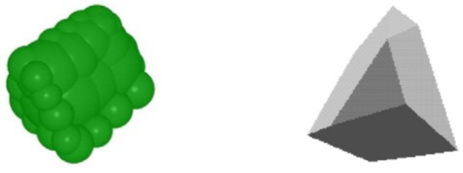

a) b)

Figure 2: Examples for Modelling Rocks with Clumps (Coetzee and Nel 2014), and Polyhedra (Huang and Tutumler 2011)

Creating polyhedral shaped elements

There are several ways to create the shape of the polyhedra. The simplest way is to create them manually, estimating the geometry of rocks with the desired precision. Advanced way is using a 3D scanner and simplifications (such method is described later in the validation section), or even automatized image processing technology has promising results (Huang and Tutumler 2011). However, what is common in these methods is that only a limited number of rocks can be processed and then reproduced. There is a need to extend the productivity of the creation process and to raise the variability of the particles. There is no pattern in the shape of the natural particles. Therefore it is feasible to generate the shapes randomly without remarkable deviation from the mined rocks. Such a solution is implemented in the chosen material model, which was inspired by Asahina and Bolander (2011) and relies on the Voronoi method (Figure 3). In 2D it is based on random points created in an area with a defined minimal distance. Then polygons are created by finding the union of apothems of the lines between the closest nuclei pairs.

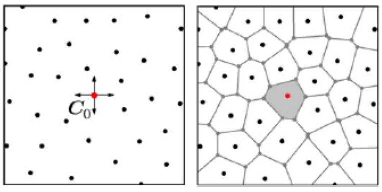


Figure 3: Voronoi Method (Eliaš 2014)

Interactions between the elements

The material of the elements is ideally rigid. The effect of plasticity is modelled within the definition of interactions (repulsive force). The model is non-cohesive. Normal and shear force is included, which arise when the polyhedrons come into contact.

The forces act in the centroid of mass of the overlapping volume. The magnitude of normal force between two particles is obtained from the magnitude of their overlapping volume by multiplying it with the factor named normal volumetric stiffness.

The direction of the normal force is perpendicular to the plane fitted by the least square method on the intersecting lines of the particle shells.
The shear force (friction) is proportional to the mutual movements and rotations of the elements. Its maximum value is regulated by the Coulomb friction model (Equation 1, where $|\mathbf{F}_s|$ is the magnitude of shear force, $|\mathbf{F}_n|$ stands for the magnitude of normal force, and φ is the internal friction angle).

$$|\boldsymbol{F}_s| \leq |\boldsymbol{F}_n| tan\varphi \qquad (1)$$

Many material models do not include velocity based damping such as the mentioned polyhedral material model (Eliaš 2014) or the model by D'Addetta et al. (2001) to dissipate kinetic energy, so it is worth considering to use an artificial damping. This numeric damping decreases the forces which increase the particle velocities and vice versa by a factor λ_d: 0-1 (Šmilauer et al. 2015).

Crushing behaviour

A further aim of the research is to simulate the fragmentation process of the rocks in the railway ballast, which leads to a great decrease in loadability. This can be obtained with a proper crushing model. (Eliaš 2014) also has the conclusion that the modelling of crushing is necessary to get a realistic behaviour of the simulated aggregate.
In the simulation crushing occurs, when the von Mises stress in the polyhedral element exceeds the so-called size-dependent strength (f_t). Using the von Mises stress is a simplification, as it is best applies to isotropic and ductile metals. However, this simplicity is an advance in discrete element modelling, as the aim is always to find the simplest material model which reproduces the aggregate (macro) behaviour with the proper micromechanical parameters.
According to Lobo-Guerrero and Vallejo (2005) the strength of the rock is dependent on its size. This phenomena is implemented in the model and is represented by Equation 2. Where f_0 is a material parameter and r_{eq} is the radius of the sphere with same volume as the polyhedra.

$$f_t = \frac{f_0}{r_{eq}} \qquad (2)$$

When crushing occurs, rocks are intentionally split into 4 pieces (Figure 4). The two mutually perpendicular cutting planes are perpendicular to the plane defined by σ_I and σ_{III} principal axes and form angles of π/4 with the planes defined by σ_I-σ_{II} and σ_{III}-σ_{II} axes.

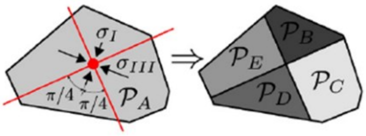


Figure 4: Crushing of Rocks (Eliaš 2014)

VALIDATION

The investigation of the properties of the polyhedral material model (Eliaš 2014) lead to a decision to use it to simulate the behaviour of railway ballast.
The first step of the process is the validation of the material model to determine if it is appropriate for the task. Two uncommon features of the selected material model are the particle geometry generation with Voronoi method and the crushing effect. The validation of these is performed by examining the particles individually. If the validation is successful, the next step is the calibration of the material parameters with the use of particle aggregate.

Particle Shape validation

To validate the randomly created polyhedral particle shapes, the generated ones were compared with the natural ones. A 3D laser scanner (NextEngine) was used to digitize the natural rock shapes. This resulted in a very complex surface geometry, which was simplified in a 3D CAD software. Simple planes were fitted on the rough surfaces using the least-square method (Figure 5). The planes were trimmed by the penetration lines. Between the manually created shapes and the randomly generated ones a high level of similarity was well observable. The average number of sides, vertices, as well as the areas of the sides and magnitudes of angles was equal with a good approximation.
A substantial simplification of the random generation is that it always creates convex polyhedra. It is easy to find concave areas on the surface of crushed rocks, which are possible stress concentration spots. The magnitude of the stress concentration effect depends on the size and shape of these concave areas and has a significant influence on the strength of the rocks.
Despite the existence of stress concentration effect at concave polyhedra, convex ones are accepted. The reason is that their strength is adjusted with the size-dependent strength material parameter and their efficient generation is also possible with Voronoi tessellation. The geometry of the crushed rocks can be estimated by the randomly generated ones relying on the Voronoi method.

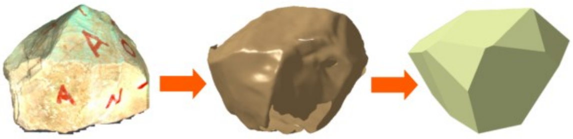


Figure 5: Steps of Rock Shape Digitalizing with 3D Scanner

Crushing test of a simple rock particle

An individual polyhedron has to crush into 4 pieces, when the von Mises stress exceeds its size-dependent strength. It occurs when the load on an element reaches a critical magnitude. The crush can occur multiple times, however, the fragmentation has to stop beside the same load, as the created new elements have greater strength because of their smaller size.

The following stages can be observed in the process (Figure 6):

1. Beginning of the load: the loading plate has not reached the rock yet.
2. The plate reaches the rock, the force starts rising.
3. Crush of the rock: when the von Mises stress exceeds the strength of the rock, it breaks into 4 pieces. After the crush, the normal force drops down to zero as the contact between the rocks and the plate interrupts.
4. Moving, sliding of the rocks. When they reach their final place, the force starts rising steeply again.
5. A second break occurs. In this case the force does not reach zero, as the contact persists.
6. Reaching the maximum force and beginning of unloading.

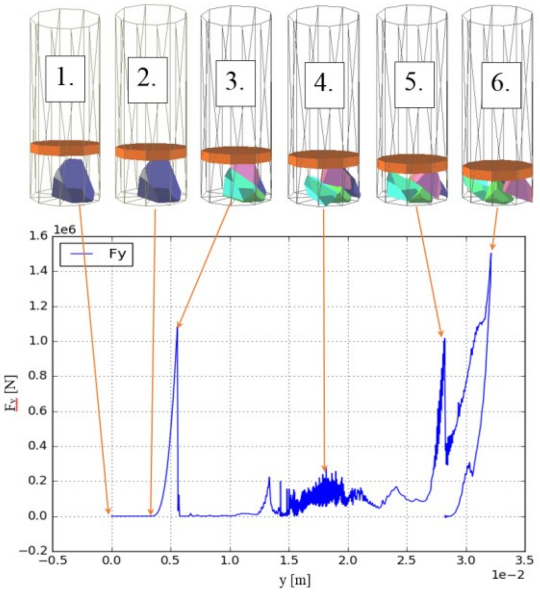


Figure 6: Loading Process of a Rock (F_y: Normal Force [N], y: Horizontal Displacement of the Plate Downwards [m])

The polyhedron breaks into 4 fragments at every crush as expected. This is certainly not a natural behaviour, however it is not an issue, because the research is more intended to create the proper model of the whole aggregate. Only the number of crushes and the further behaviour of the aggregate is important.

Despite the fact that splitting intentionally into 4 pieces is certainly not a natural behaviour, the model can be used in further applications to analyse the behaviour of an aggregate.

CALIBRATION OF THE MATERIAL MODEL

To create the model of the railway ballast with approximately the same macromechanical parameters as the natural aggregate, the proper setting of micromechanical parameters is needed. It can be obtained by calibration. During the calibration process, the simulation of a properly chosen measurement is compared to the corresponding experimental data. The adequate measurement process applies similar loads as the forces acting in natural aggregate, so the appropriate material parameters can be set. In our case, the measurement device used in the Hummel procedure is applied for the calibration process, where crushing occurs as the effect of quasi-static loads.

Hummel Device

The Hummel procedure (Hungarian Standards Institution 1983) gives information about the static loadability of construction aggregates. The standard describes the properties of the tested material, the geometry of the tool (Figure 7) and the process of loading progress. Particles of the aggregate are crushing during the test. The Hummel process enables the determination of force-displacement and particle size distribution curves under different peak loads for the tested aggregate. Tested material parameters have to be changed in order to be able to perform the test, as the 31.5/50 mm railway ballast rocks are too big for the Hummel device. Therefore, the laboratory tests and the DEM simulation is performed with 22/32 mm rocks. In the further research, grain size sensitivity simulations and measurements will be performed with smaller rocks in order to get information about the particle size dependency of the aggregate behaviour. The parameters of the 31.5/50 mm railway ballast rocks can be set by extrapolation.

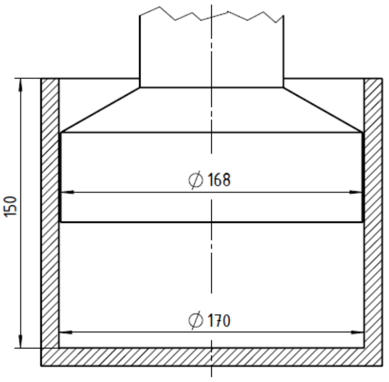


Figure 7: Hummel Device (Hungarian Standards Institution 1983)

Creating the Geometry

To obtain the optimal processing speed besides the proper geometry, the tube was approximated with a regular decagonal prism, created from triangular facets. The polyhedral particles were generated in this volume with the desired geometry (equant or flat). In the first stage of the research, equant elements were modelled. The randomly generated polyhedra are created in a cuboid volume (Figure 7/a.) inside the prism. To create a dense pack of polyhedra on the bottom of the tube model, gravitational deposition was used (Figure 8.). Gravity load is applied in the simulation and the elements fall down freely to the bottom of the prism. The height of the produced dense pack is bigger than the Hummel device has, so the elements with centroid higher than 150 mm were removed (Figure 9).

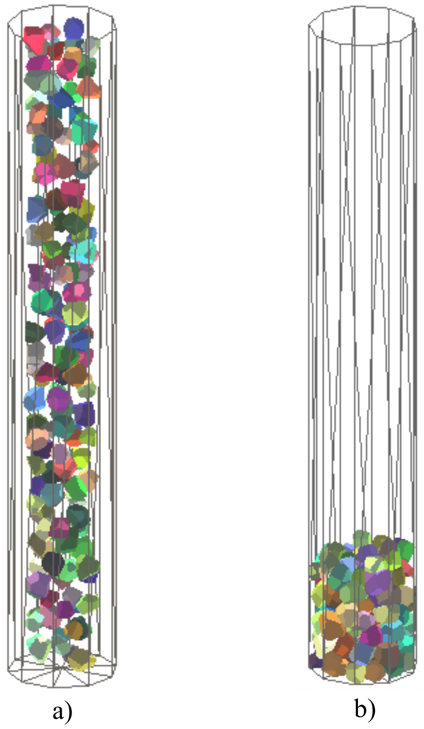


Figure 8: a) Created Polyhedra; b) the Result of Gravitational Deposition

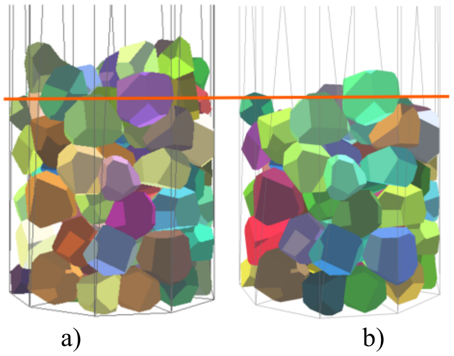


Figure 9: a) The Polyhedra Pack Before; b) and After the Removal Process

Applying the Load

The initial and the maximum load state of the simulation can be observed on Figure 10. The load is applied through the top plate which is also a polyhedron type element with disabled crushing and has a special shape to fit in the tube. The plate moves down with constant velocity, thus the magnitude of the force on the plate is displacement driven, which is constantly measured. When the pre-defined maximum force is reached, the direction of the plate movement changes and the unloading process begins. The simulation ends at the moment, when the normal force reaches zero. Data are saved in text format.

The force-displacement curve of a typical simulation is represented on Figure 11. It corresponds to the theoretical assumptions stated in the followings. In the first part of the loading process the normal force raises slowly. In this session the elements can move easily because of the high porosity of the aggregate. As the porosity decreases, the elements have less space to move, the curve becomes steeper, and also the rate of crushing events increases. In the unloading phase, the normal force drops down to zero. Crushing can cause peak forces, which eliminate in a few timesteps and have no significant influence on the loading characteristics. Therefore it is now assumed that they can be ignored, but further investigations will be done.

Considering that the simulation used preliminary material parameters, the results are acceptable and the material model and calibration method can be utilized in the further research.

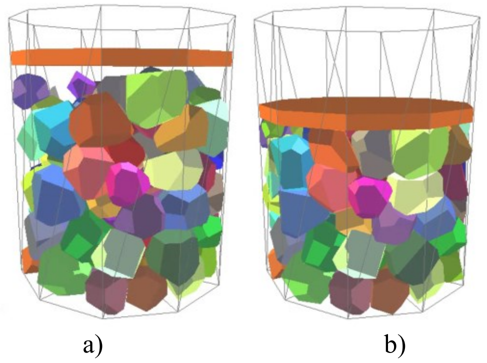


Figure 10: a) The Device Before Beginning of Load; b) at Maximum Load

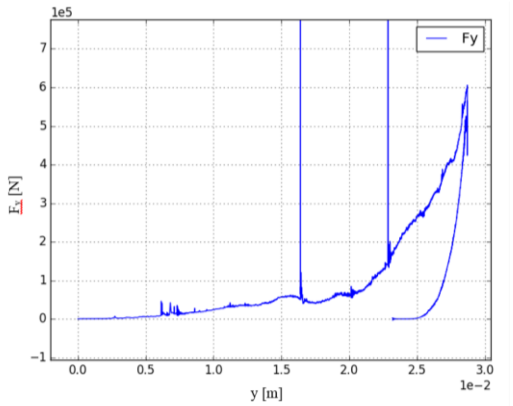


Figure 11: Normal Force (Fy [N])-Displacement (y [m]) Graph of the Hummel Device

CONCLUSIONS

In this paper an ongoing research for modelling the railway ballast is introduced. Different DEM material models were investigated, and the one featuring randomly generated convex crushable polyhedral elements was chosen. The mechanical basics of the model, random element generation via Voronoi tesselation and a crushing behaviour was investigated and validated. A calibration method was created that uses the Hummel device.

In the studied discrete element material model the geometry of the polyhedral elements approximates the shape of the rocks sufficiently, the mechanical model reproduces the behaviour of the rock aggregate, and the crushing works. The discrete element method is capable to simulate the railway ballast.

The gravitational deposition executes with the assigned micro parameters in the created geometry. The characteristics of the load-displacement curve is the same as theoretically expected. In some cases, peak forces occur during breakage, but they have no significant effect on the nature of the curve. In further studies they will be investigated in detail.

The model of the measurement based on the Hummel device is proper, so the calibration of the discrete element material model can be performed in the further stages of the research.

FURTHER RESEARCH

The next step is the static calibration of the material model. It is performed by processing the measurement data, and running a series of simulations of the Hummel device measurement with different material parameters to find the parameter combination that reproduces the experimental data with the best approximation.

A calibrated and validated material model can be used to simulate the behaviour of railway ballast and answer the technical questions discussed in the introduction.

ACKNOWLEDGEMENT

This research is a part of an R&D project supported by Magyar Államvasutak Zrt. The support is gratefully acknowledged. We are thankful for the fellow workers of the Department of Construction Materials and Technologies, BME: Miklós P. Gálos professor for coordinating the project and Gyula Emszt, department engineer and Bálint Pálinkás technician for performing the laboratory measurements.

REFERENCES

Asahina, D. and J.E. Bolander. 2011. “Voronoi-based discretizations for fracture analysis of particulate materials”. *Powder Technology* 213, 92–99.

Asahina D. and M.A. Taylor. 2011. “Geometry of irregular particles: Direct surface measurements by 3-D laser scanner”. *Powder Technology* 213, 70-78.

Bagi, K. 2007. *A diszkrét elemek módszere*. BME Department of Structural Mechanics, Budapest, 5-12.

Coetzee, C.J. and R.G. Nel. 2014. “Calibration of discrete element properties and the modelling of packed rock beds”. *Powder Technology* 264, 332–342.

D’Addetta, G.A.; F. Kun; E. Ramm and H.J. Herrmann. 2001. “From solids to granulates - Discrete element simulations of fracture and fragmentation processes in geomaterials”. In *Continuous and Discontinuous Modelling of Cohesive-Frictional Materials. 2001*, Vermeer, P.A.; S. Diebels; W. Ehlers; H.J. Herrmann; S. Luding and E. Ramm. Berlin Heidelberg 231-258.

Eliáš, J. 2014. “Simulation of railway ballast using crushable polyhedral particles”. *Powder Technology* 264, 458–465.

Huang, H. and E. Tutumluer. 2011, “Discrete Element Modeling for fouled railroad ballast”. *Construction and Building Materials* 25, 3306–3312.

Lane, J.E.; P.T. Metzger and R.A. Wilkinson. 2010 “A Review of Discrete Element Method (DEM) Particle Shapes and Size Distributions for Lunar Soil” *NASA/TM* 216-257.

Lobo-Guerrero, S. and L.E. Vallejo. 2005. “Crushing a weak granular material: experimental numerical analyses”. *Geotechnique* 55, No.3 245-249.

Hungarian Standards Institution. 1983. *Building rock materials. Strength testing of aggregates. Hummel test (MSZ 18287-3:1983).*

Shahraki, M; C. Warnakulasooriya and K.J. Witt. 2015. “Numerical study of transition zone between ballasted and ballastless railway track”. *Transportation Geotechnics* 3, 58-67.

Šmilauer, V. et al. 2015. “Yade Documentation 2nd ed.” In *The Yade Project* (http://yade-dem.org/doc/)

Wachs, A.; L. Girolami; G. Vinay and G. Ferrer. 2012. “Grains3D, a flexible DEM approach for particles of arbitrary convex shape — Part I: Numerical model and validations” *Powder Technology* 224 374–389.

AUTHOR BIOGRAPHIES

ÁKOS OROSZ was born in Hódmezővásárhely, Hungary. He is doing his mechanical engineering MSc studies at the Budapest University of Technology and Economics, where he received his BSc degree in 2016. He is also a member of a research group in the field of discrete element modelling. His e-mail address is: orosz.aakos@gmail.com and his web-page can be found at http://gt3.bme.hu/oroszakos.

JÁNOS P. RÁDICS is an assistant professor at Budapest University of Technology and Economics where he received his MSc degree. He completed his PhD degree at Szent István University, Gödöllő. His main research is simulation of soil respiration after different tillage methods, and he also takes part in the DEM simulation research group of the department. His e-mail address is: radics.janos@gt3.bme.hu and his web-page can be found at http://gt3.bme.hu/radics.

KORNÉL TAMÁS is an assistant professor at Budapest University of Technology and Economics where he received his MSc degree and then completed his PhD degree. His professional field is the modelling of granular materials with the use of discrete element method (DEM). His e-mail address is: tamas.kornel@gt3.bme.hu and his web-page can be found at http://gt3.bme.hu/tamaskornel.

Backbone Strategy for Unconstrained Continuous Optimization

Michael Feldmeier
University of Edinburgh
Email: m.feldmeier@sms.ed.ac.uk

Thomas Husslein
OptWare GmbH
Email: thomas.husslein@optware.de

KEYWORDS

Backbones; Genetic Algorithms; Global Optimization

ABSTRACT

Backbones in optimization problems are structures within the decision variables that are common to all global optima. Identifying those backbones in a deterministic manner is at least as hard as solving the original problem to optimality because all optimal solutions to the problem have to be known. A number of different algorithms have been proposed which use heuristically determined backbones to speed up discrete combinatorial optimization algorithms by eliminating these backbones and thus reducing the dimensionality of the optimization problem to be solved in each step. In this paper we extend the concept of backbones to real-valued optimization. We propose a definition of such backbones and introduce means to identify them and determine their value by the use of a genetic algorithm. We compare the performance of the resulting algorithm with an ordinary optimization procedure on a widely used nonlinear and unconstrained optimization benchmark. We observe that our backbone strategy is superior in terms of both convergence speed and quality of the resulting solutions. Limitations of this first approach and ideas how to resolve them in future work are considered.

Previous Work on Backbones

Concepts of backbone-driven optimization techniques were first mentioned by Lin and Kernighan in 1973 [6]. The predecessor of the backbone algorithm, which they called 'Reduction', was mentioned as a refinement for their famous TSP-heuristic (Lin-Kernighan-heuristic, LKH). Reduction is a parallel and iterative process: Run multiple LKH instances in parallel and compare intermediate, solutions for links appearing in all those solutions. Those prevalent links are likely part of the optimal or a close-to-optimal solutions and are not allowed to be broken in the further optimization process. Reduction speeds up the optimization process by eliminating situations which need not be evaluated by LKH. Lin and Kernighan also stated that Reduction is a means to 'direct[ing] the search among otherwise indistinguishable cases' by preventing solutions which are most likely not optimal.

To the best knowledge of the authors not much attention was given to this topic until Schneider et al. developed an algorithm called 'Searching for Backbones' (SfB) in 1996 [9] which produces heuristic TSP-solutions primary driven by the extensive search for Backbones. SfB is an iterative algorithm as well. A number of instances of an arbitrary optimization algorithm are run in parallel (the authors used Simulated Annealing and Threshold Accepting algorithms). The solutions get analysed for coinciding links similar to the Reduction-algorithm, but instead of simply forbidding to break those links, the TSP instance gets recoded: each partial tour appearing in all of the solutions gets contracted to a single link with the respective length. This process is then repeated on this smaller problem, until the Backbone encompasses the whole tour. SfB was quite successful in tackling TSP-benchmarks.

After the publication of SfB there was generally more research done on Backbones: in 1999 Monasson et al. associated the number of backbones in (2+p) - SAT instances with their computational difficulty [8], confirming the importance and influence of Backbones in optimization. Their work is based on the connection of Backbones with the phase transition of the SAT problem, in particular that large backbones in the transition phase make problem instances hard to solve. Further research in this direction was conducted by Singer et al. 2000 [10], Walsh and Slaney 2001[11], Zhang 2001 [13] and others.

Successful Backbone-based optimization techniques to solve the SAT problem and its variants were developed by Dubois and Dequen 2001 [1] Menai and Batouche 2005 [7] and Zeng et al. 2012 [12] for instance.

Research on the theoretical properties of the Backbone of the TSP has been done by Kilby et al. 2005 [5], concluding that Backbone-based optimization heuristics are pretty good in practice, but can generally not give any guaranteed approximation to the optimal value.

Besides Schneider et al.'s SfB, other Backbone-based optimization techniques for the TSP have been developed, e.g. Fischer and Merz 2007 [2]. Jäger et al. 2013 used Backbones to solve very large TSP instances and set new world records [3].

Preliminaries

Genetic Algorithm

Genetic Algorithms are heuristic optimization procedures, based on the phenomena of natural selection and survival of the fittest. A population of solutions to an optimization problem gets evolved using Selection, Crossover, Mutation and Elitism. Usually the algo-

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

rithm produces good solutions, even though it is not guaranteed to be optimal. The standard implementation of an genetic algorithm can be seen is figure 1.

```
1:  Generate a randomly generated initial population.
2:  Compute the fitness of each individual.
3:  Copy the fittest individuals to a new population.
4:  while |new population| < | old population| do
5:      Select two individuals (parents).
6:      Generate new individual by applying crossover.
7:      Perform mutation on this offspring
8:      Add it to the new population.
9:  end while
10: Apply a stop criterion. If not satisfied, go to step 2.
11: Output individual with best fitness value
```

Fig. 1. The Genetic Algorithm

Backbone Strategy

A real-valued backbone is an assignment of values to the decision variables which is common in all global optima of a given optimization problem. Finding the backbones in a deterministic manner requires finding all these optima and is thus not feasible. Therefore, we propose a heuristic approach. Our algorithm involves producing a set of reasonably good solutions by use of a heuristic optimization procedure. We used a Genetic Algorithm in our approach. The different solutions created by the Genetic Algorithm get compared. Finding some of the decision variables to have very similar values in all solutions we observed it as unlikely that those values would change in further optimization. Hence we do not spend further computational power on their improvement. This can be accomplished by either disallowing changes to them or remove the according dimension from both the search space and the solutions. In the latter case, the cost function has to be adapted in order to correctly consider the backbone values. We repeat this process until the backbone covers all dimensions. Since the analytical form of the objective function might not be available, the backbones are dissolved before evaluating a solution. Note that the speedup of our algorithm does not come from accelerating the evaluation of the objective function but from speeding up the optimization procedure used by limiting the number of variables. As the dimensionality of the search space is continuously reduced, the underlying heuristic (GA) has to deal with a smaller problem, which can be solved more efficiently. The goal of our approach is to achieve better objective function values by limiting the search space to areas with promising solutions. The success of this, however, cannot be guaranteed, as the Backbone Strategy is a heuristic algorithm in itself.

Determining Backbones

Due to the nature of continuous numbers, numerical imprecision and the use of heuristic algorithms, it is highly unlikely for two real-valued variables to have the exact same value in all available solutions. For our applications, it is sufficient for the variables to be very close to each other in order to be considered as equal in the backbone determination process. We introduce a control parameter ϵ to define a criterion for detecting backbones:

```
1: while backbone does not cover all variables do
2:     Generate solutions using a heuristic optimization algorithm
3:     Analyse the solutions for coinciding variables
4:     Find values for the new backbones
5:     Remove new backbones from both the solutions and the problem
6: end while
```

Fig. 2. The algorithm for the Backbone Strategy. Note that step two involves running multiple instances of an optimization algorithm, which can be computed in parallel.

Definition: Given a set of solutions $S = \{x^1, x^2, \cdots, x^n\}$, with $x^i \in \mathbb{R}^>$ and deviation ϵ, variable x_j is regarded as a Backbone, iff $\forall x^k, x^l \in S : |x^k_j - x^l_j| \leq \epsilon$.

This definition gives a clear statement how to find backbones in real valued Problems. The ϵ parameter controls the behaviour of the algorithm and needs to be adjusted to the specific problem. At its best, the parameter is set in a way that a Backbone is only found if the partial solution is in the same local optimum in all available solutions. A high value will allow very different partial solutions, possibly located in different local optima, to be considered a Backbone. A low value will restrict the search for Backbones too much; partial solutions might not be recognized as Backbones, even if they are close to each other in the search space, resulting in slow or no convergence of the algorithm. The number of solutions n also works in a similar manner: the more solutions are involved, the less likely it will be for backbones to be found. This is a common trade-of between computational power requirements and quality of the solutions.

Determining Backbone Value

After the algorithm has identified variables which shall become Backbones, a value needs to be assigned to each of them, replacing the variable in the evaluation of the objective function. Since a set of values is available - one for each available solution - this value is not immediately obvious. For instance, the mean or the median of those values can be chosen.

Experimental Setup

Objective Function

To demonstrate the power of our algorithm we use Schwefel's function as benchmark. It was chosen as it is well researched and thus ensures comparability with other optimization procedures. As the function is highly non-linear, it is a challenging benchmark to solve. It is additively separable, i.e. it is of the form

$f(x) = \sum_{i=1}^{n} g_i(x_i), g : \mathbb{R} \to \mathbb{R}, f : \mathbb{R}^{\ltimes} \to \mathbb{R}$, and thus a good candidate to demonstrate the Backbone Strategy as fixing individual decision variables does not influence the optimization in the remaining dimensions. It scales to any dimension n easily.

$$\begin{aligned} \text{minimize } F(x) &= V + \sum_{i=1}^{n} -x_i \cdot sin(\sqrt{|x_i|}) \\ \text{with } V &= n \cdot 418.982 \end{aligned}$$

The search space usually encompasses the hypercube with centre 0 and range $[-500, 500]$ in every dimension. Its global optimum x^* can be found at $x_i = 420.98 \ \forall i$ with a value of $F(x^*) = 0$.

Parameter choices

Our goal was to compare the performance of the backbone genetic algorithm with an implementation of a standard genetic algorithm. In our studies we evaluated a wide range of parameters. The performance on Schwefel's function in various dimensions ($D = 10, 30, 50, 100, 150, 200$) was evaluated in combination with different numbers of instances run in parallel ($N = 2, 3, 4, 5$) and a set of different values for the deviation ($\epsilon = 1, 5, 10, 20$). The standard genetic algorithm parameters were the same within all experiments, i.e. crossover rate of 0.8, mutation rate of 0.05, and elitism of 5%. Each scenario is run 5 times and the average of the resulting fitness value is reported.

Experimental Results

In figure 3 the mechanics of the backbone strategy is demonstrated. While the standard genetic algorithm (blue line in main plot) suffers of degeneracy and does not improve further after just 500 iterations, the backbone variant (red line) restarts over and over again with fewer variables than before, increasing the backbone and eventually reaching the optimal solution. After each restart, the objective is worse than before, but improves to a better value. This is demonstrated in the small plot where the development of the objective value of a single instance is showed. In a way, the Backbone Strategy can be understood as a means to efficiently restart other heuristics. In table I the results of each scenario are plotted. The low dimensional case with 10 variables was solved to optimality by every choice of parameters. However, already 30 dimensions were sufficient to have scenarios with 2 or 3 instances fail occasionally. This trend continues: The more dimensions are involved, the more apparent it becomes that a higher number of instances leads to better results. The Epsilon value does not influence the results in any significant way, but there also is a trend recognizable: lower ϵ-values yield better results. However, the combination of low Epsilon with a higher number of dimension introduced problems: the algorithm did not converge in reasonable time. Comparing this to the results of the standard genetic algorithm in table II, it becomes evident that the backbone version yields the same or better results in every dimension, even if the standard algorithm is run for a very long time. This is due to the known problem of degeneracy, which is prevented by the use of the backbone approach. Even for the very high dimensional case of 200 dimensions, the Backbone variant was able to get very close to the optimal solution, at an average value of 525, while the standard genetic algorithm failed miserably with a value over 20000.

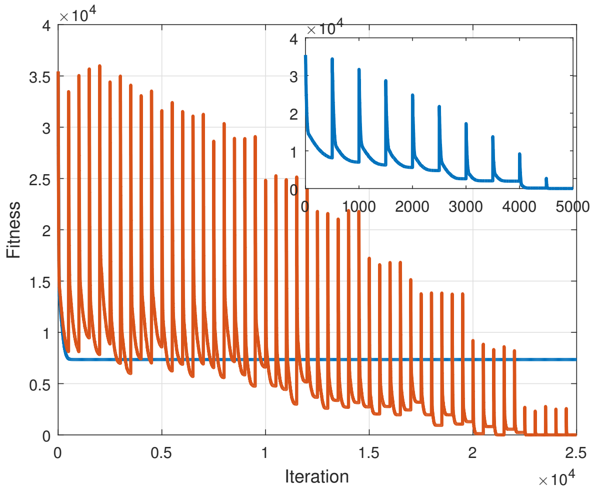


Fig. 3. Main plot: single instance of a standard genetic algorithm in blue, 5 instances of a backbone variant genetic algorithm in red. The objective values of each of the 5 instances are plotted one after another. Subplot: development of objective function value of a single instance of the backbone version.

In experiments in the literature, Schwefel's function is usually used with relatively few dimensions. A common dimension is $D = 30$ (see e.g. [4]). Resembling only the second smallest problem size in our experiments, these experiments yielded optimal or very close to optimal solutions in all experiments with $N > 2$, outperforming a number of algorithms in presented in literature, e.g. the Particle Swarm Optimization and an implementation of the Artificial Bee Colony (PSO and ABC1 in [4]). However, we firmly believe the strength of our algorithm lies in higher dimensions, so additional comparative analysis besides our implementation of a standard genetic algorithm with established optimization techniques will have to be carried out in the future.

LIMITATIONS

A genetic algorithm was used in all experiments so far. As genetic algorithms are inherently parallel, its use was clearly limited by the very high number of solutions that need to be dealt with. A high number of operations, such as fitness function evaluations, need to be executed in each iteration and so the Backbone approach is computationally expensive - a problem which it was supposed to solve in the first place.

The experiments were limited to a unconstrained benchmark - Schwefel's function. In theory, the Backbone Strategy's ability to solve constrained optimization problems depends on the underlying heuristic: If

D	N	Epsilon			
		1	5	10	20
10	2	0	0	0	0
	3	0	0	0	0
	4	0	0	0	0
	5	0	0	0	0
30	2	95	32	75	53
	3	0	10	0	10
	4	0	0	0	0
	5	0	0	0	0
50	2	297	224	256	299
	3	32	53	53	64
	4	0	0	10	21
	5	0	0	0	10
100	2	1077	1668	1292	1287
	3	318	531	319	424
	4	68	112	114	103
	5	22	56	26	25
150	2	3105	3150	3606	3184
	3		1140	1093	1219
	4			683	595
	5				215
200	2	5905	6549	6052	6237
	3		2332	2545	2269
	4			1269	1239
	5				525

TABLE I: Average fitness values over multiple runs for each scenario with dimensions D, N instances and the ϵ deviation used in the backbone determination process. Missing values indicate that at least one of the instances did not converge in reasonable time. Values closer to zero indicate high-quality solutions, with zero being the optimum.

	Dimension					
Iter.	10	30	50	100	150	200
500	0	549	2216	7916	18093	29642
1000	0	595	1981	7943	14506	24201
10000	10	651	1832	8348	15192	23320

TABLE II: Average resulting fitness value for the standard genetic algorithm for various dimensions and number of iterations. The algorithm fails to solve the problem to optimality at only 30 dimensions.

the heuristic guarantees feasible solutions, then the solution of the Backbone approach is guaranteed to be feasible as well. However, additional experiments need to be executed in order to determine whether the Backbone Strategy offers computational benefits when dealing with constrained optimization.

CONCLUSIONS

In this paper we introduced the Backbone Strategy as a means to restart heuristics to prevent degeneracy of their solution, resulting in better objective function values at the cost of more computational power. As the genetic algorithm is inherently parallel, the limitation of its use were the very high number of solutions that need to be dealt with. Further study will apply the Backbone Strategy to other search heuristics, such as Simulated Annealing or Firefly Algorithms. They usually work with only few solutions at a time and are therefore excellent candidates to apply the Backbone Strategy. This way optimization problems with many more dimensions, as they arise in applications in data science, finance, and engineering, should be tackled efficiently and globally.

REFERENCES

[1] Olivier Dubois and Gilles Dequen. A backbone-search heuristic for efficient solving of hard 3-sat formulae. In *IJCAI*, volume 1, pages 248–253, 2001.

[2] Thomas Fischer and Peter Merz. Reducing the size of traveling salesman problem instances by fixing edges. In *European Conference on Evolutionary Computation in Combinatorial Optimization*, pages 72–83. Springer, 2007.

[3] Gerold Jäger, Changxing Dong, Boris Goldengorin, Paul Molitor, and Dirk Richter. A backbone based tsp heuristic for large instances. *Journal of Heuristics*, 20(1):107–124, 2014.

[4] Dervis Karaboga and Bahriye Basturk. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm. *Journal of global optimization*, 39(3):459–471, 2007.

[5] Philip Kilby, John Slaney, Toby Walsh, et al. The backbone of the travelling salesperson. In *IJCAI*, pages 175–180, 2005.

[6] Shen Lin and Brian W Kernighan. An effective heuristic algorithm for the traveling-salesman problem. *Operations research*, 21(2):498–516, 1973.

[7] Mohamed El Bachir Menaï and Mohamed Batouche. A backbone-based co-evolutionary heuristic for partial max-sat. In *International Conference on Artificial Evolution (Evolution Artificielle)*, pages 155–166. Springer, 2005.

[8] Rémi Monasson, Riccardo Zecchina, Scott Kirkpatrick, Bart Selman, and Lidror Troyansky. Determining computational complexity from characteristic phase transitions. *Nature*, 400(6740):133–137, 1999.

[9] Johannes Schneider, Christine Froschhammer, Ingo Morgenstern, Thomas Husslein, and Johannes Maria Singer. Searching for backbonesan efficient parallel algorithm for the traveling salesman problem. *Computer Physics Communications*, 96(2):173–188, 1996.

[10] Josh Singer, Ian P Gent, and Alan Smaill. Backbone fragility and the local search cost peak. *J. Artif. Intell. Res.(JAIR)*, 12:235–270, 2000.

[11] Toby Walsh and John Slaney. Backbones in optimization and approximation. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI-01)*, 2001.

[12] Guoqiang Zeng, Yongzai Lu, Yuxing Dai, Zheng-

guang Wu, Weijie Mao, Zhengjiang Zhang, and Chongwei Zheng. Backbone guided extremal optimization for the hard maximum satisfiability problem. *Int. J. Innovative Comput. Inf. Control*, 8(12):8355–8366, 2012.

[13] Weixiong Zhang. Phase transitions and backbones of 3-sat and maximum 3-sat. In *International Conference on Principles and Practice of Constraint Programming*, pages 153–167. Springer, 2001.

Michael Feldmeier was born in Wörth a.d. Donau in Germany and studied Computer Science at the Technical University of Applied Sciences Regensburg. Currently he is pursuing a Master's degree in Operational Research at the University of Edinburgh. His research interests include heuristic optimization, column generation, and interior point methods.

Thomas Husslein was born in Munich, Germany and went to University of Regensburg, where he studied computational physics and obtained his degree in 1993. He obtained his PhD in physics from the University of Regensburg in 1996. Afterwards he did his Postdoc at the IBM Research Center Yorktown Heights and the University of Pennsylvania. In 1999 the founded OptWare in order to transfer mathematical methods into application which he is leading ever since. Since 2012 he has a teaching assignment for Operations Research at the University of Applied Sciences OTH Regensburg.

GENERATION ALGORITHMS OF FAST GENERALIZED HOUGH TRANSFORM

Egor I. Ershov
Moscow Institute of Physics and Technology
141700, 9, Institutskiy per., Dolgoprudny,
Moscow Region, Russia;
Institute for Information
Transmission Problems, RAS
127994, 19, Bolshoy Karetny per.,
Moscow, Russia
E-mail: ershov@iitp.ru

Evgeny A. Shvets
Moscow Institute of Physics and Technology
141700, 9, Institutskiy per., Dolgoprudny,
Moscow Region, Russia;
Institute for Information
Transmission Problems, RAS
127994, 19, Bolshoy Karetny per.,
Moscow, Russia
E-mail: vortexd77@gmail.com

Timur M. Khanipov
Institute for Information
Transmission Problems, RAS
127994, 19, Bolshoy Karetny per.,
Moscow, Russia
E-mail: timur.khanipov@iitp.ru

Dmitry P. Nikolaev
Institute for Information
Transmission Problems, RAS
127994, 19, Bolshoy Karetny per.,
Moscow, Russia
E-mail: dimonstr@iitp.ru

KEYWORDS

Hough transform, fast Hough transform, generalized Hough transform, greedy algorithm, graph matching

ABSTRACT

In this paper we investigate the problem of finding the minimal operations number for the generalized Hough transform computation (GHT). We demonstrate that this problem is equivalent to the addition chain problem and is therefore NP-complete. Three greedy methods for generating GHT computation algorithms are proposed and their performance is compared against the fast Hough transform (FHT) for different discrete straight line pattern types. The additional result of this work is the experimental proof of the FHT non-optimality.

INTRODUCTION

The Hough transform was patented in 1962 by an american scientists Paul Hough for detecting straight tracks in a bubble chamber (Hough. and Arbor 1962) on a photograph taken during an experiment. In 10 years a research was conducted, studying the possibility of using the Hough Transform (HT) for detecting analytical (Duda and Hart 1972) and arbitrary (Merlin and Farber 1975) lines and patterns.

However the suggested algorithms were too demanding on computational power which initiated research about possible reduction in HT computation time and memory resources. In 1972 a principle was suggested to reduce the Hough space dimension for ellipses. This idea used information about possible ellipse center position according to gradient, computed for a given point (Kimme et al. 1975), and can obviously be expanded for any analytical curve.

Later in 1981 D. Ballard published a generalization for all line shapes (patterns) (Ballard 1981) which are computable on a grayscale image, calling it a generalized Hough transform (GHT). After a year in (Davis 1982) an hierarchical GHT computation method for image scale pyramid was suggested which reduced computation time but was still insufficient for real-time systems. In 2003 Ulrich suggested to use hierarchical mapping and R-tables computation (Ulrich et al. 2003), what, as authors say, allows to use this algorithm in real time t create image feature space and solve pattern recognition problems.

An alternative GHT branch was pattern choice randomization during computation (Xu et al. 1990; Xu and Oja 1993). This approach allows to drastically reduce the amount of necessary operations and memory used but does not guarantee an optimal result for a given image. In more detail HT and GHT research and development questions are reviewed in (Illingworth and Kittler 1988; Brown 1992; Mukhopadhyay and Chaudhuri 2015).

Variety of such research prove actuality of fast Hough transform for arbitrary pattern creation problem. Indeed feature extraction is a basic technique in unmanned transport visual algorithms (Konovalenko et al. 2015; Karpenko et al. 2015a,b). Moreover, HT can be useful as an alternative to the methods of pattern recognition (Kuznetsova et al. 2015).

Computing GHT might also be considered as a certain variation of the addition chain problem (Garey and Johnson 1979): for any given pattern (straight line, circle, etc) find such a family of subsets of the pixels set which are to be summed, that the total number of summation operations is minimal. It is worth noting that any binary operation (subtraction, maximum, minimum) might serve as the key

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

operation and not only the addition, which substantially expands the GHT computable image feature set. Hence, we have a problem of developing an algorithms generator parametrized by the target pattern and operation types. It should be noted that this approach is in some sense a generalization of the Ulrich's (Ulrich et al. 2003) approach except that the researcher's effort of creating algorithmic constructions accelerating exhaust search should be automated by means of the algorithm generator to be created. We will call the obtained all image-pattern summation algorithm the fast generalized Hough transform (FGHT).

The addition chain problem is NP-complete (Garey and Johnson 1979). In Pippenger's 1980 paper a generalized algorithm for computing the monomials set is suggested (Pippenger 1980), which was further analyzed in more detail in the article (Bernstein ????). Despite this problem's NP-completeness the question of building its approximations with polynomial complexity remains open. For measuring performance of the proposed methods we will compare the constructed algorithm with the fast Hough transform.

PROBLEM STATEMENT

We now formulate the addition chain problem as in (Garey and Johnson 1979). Let C be a family of subsets of a finite set A and J is a positive integer. Is there a sequence

$$S = z_1 \leftarrow x_1 \cup y_1, ..., z_i \leftarrow x_i \cup y_i, i \leq J \qquad (1)$$

of union operators, where each x_i and y_i is either a for some $a \in A$, or z_k for some $k < i$, and for all i, $1 \leq i \leq j$, x_i and y_i do not intersect and for each subset $c \in C$ exists such z_i, $1 \leq i \leq j$, that the corresponding set is equal to c.

We can draw an analogy between the addition chain problem and the problem of constructing the FGHT algorithm. Each element of the A set is an image pixel. A set of patterns for summation (subtraction, maximum or minimum) is given by the C set. In essence we need to find a minimal value J which provides the positive answer for the problem question and produce the corresponding summation sequence.

This problem is NP-complete and further in this paper we show its approximate solution.

GENERALIZED FAST HOUGH TRANSFORM

Consider A elements enumeration with natural numbers, i.e. each c subset can be presented as a vector of natural numbers.

The key idea of our approach is accounting and reusing sums of elements combinations which simultaneously appear in several c subsets. As a replacement we will choose such elements pair p which has the maximal number of occurrences in all patterns n_p (obviously, that is a greedy algorithm). This substitution repeats until each c subset contains one element. Since each such operation reduces the total number of elements in all patterns, this condition has to occur at some point. The case when at a certain step we have several pairs with maximal n_p will be analyzed later.

To implement the algorithm we have to maintain the current n_p values for every pair at each step. For that purpose we will use a hash-table T with increasingly ordered elements with pair p serving as a key and n_p as a value. Assume that elements of each pattern c are also initially ordered by their numbers in increasing order. Before executing the algorithm we fill in T iterating over all patterns c. We will now describe the procedure of fast T value update after each pair substitution. Suppose the new element's e_n number is greater than the maximal one by 1. After substitution we perform pattern content correction during which we consider only the patterns where substitution has taken place. It can be determined by the maximal element value: a substitution occurred if and only if it is equal to the new element's number. For new elements of every pattern where substitution took place, the hash-table is edited: we remove element pairs of kind {pattern element - removed p pair element} and add a pair {pattern element - new element}. This way of accounting for n_p for all pairs p allows to reduce computation time.

Let us now consider a case when at a certain step we have more than one pair with maximal n_p^*, comprising a set M.

We offer three distinct methods of choosing a substitution pair from M: random choise algorithm, the greedy algorithm and maximum matching search. According to the first one algorithm we randomly choose element of M at each step. We perform computational experiments to investigate properties of this algorithm.

Greedy algorithm

As already mentioned above, the proposed FGHT algorithm is greedy and performs search and substitution for the pair with maximal n_p. Note that maximal n_p cannot change after substitution. We will then choose such $p \in M$ that after substitution we get the biggest number of pairs occurring n_p^* times. In other words, we look for such p which will invalidate minimal number of pairs from M.

To implement such method we should determine how many pairs would become invalidated after p substitution. We make a pass through M using the T table and comparing pairs from M to find equal elements e. If two pairs do not have a common element then it is guaranteed that replacing one of them will not invalidate the other. If they do have a common element, then invalidation will happen. Pseudo-code of this method is depicted in algorithm 1 border.

For the case when several p^* have been found which invalidate equal minimal number of pairs $p \in M$, we construct a set $M^* \in M$ of such pairs. We then again choose such p^* from M^* which invalidate a minimal number of pairs in this set. We repeat the procedure recursively until only one pair is left. If the set size is not reduced during the next iteration, we take a random pair in the result set M^{fin}.

Algorithm 1 Greedy selection algorithm

```
Input: $T, M$;
Output: $p$;
$n_{max} = 0$;
$p_{max} = 0$;
for $p \in M$ do
    n = ComputeInvalidateNum(p,M)
    if $n > n_{max}$ then  $n_{max} = n$;  $p_{max} = p$;
    end if
end for
return p;
________________________________________
procedure COMPUTEINVALIDATENUM(p,M)
    $inv\_num = 0$;
    for $q \in M \setminus p$ do
        if q and p have equal element number then
            $inv\_num = inv\_num + 1$;
        end if
    end for
    return $inv\_num$;
end procedure
```

Graph matching algorithm

We will take the maximal (by number of elements) subset of $p \in M$ not having common pairs and substitute all those pairs at once. Since they do not have common elements, the order of substitutions does not matter.

To use this method one should choose a set of maximal number of pairs which do not have common elements. This problem is similar to finding maximal matching in a graph. Let us build a graph where elements of pairs $p \in M$ are vertices and edges connect elements contained in one of the maximal pairs. Then the maximal edge matching of the graph is exactly the set of edges pairs p of maximal size, such that two distinct pairs do not contain a common element. To find the matching we use the Lemon graph third party library which implements the Blossom matching search algorithm with $O(V^2E)$ complexity (V and E are vertices and edges number resp.).

COMPUTATIONAL EXPERIMENT

To verify the generation algorithm correctness and to estimate the complexity of the generated FGHT we will solve the problem of computing the Hough transform.

In the first experiment we generated FGHT for dyadic patterns (a discrete straight line type used in the fast Hough transform algorithm (FHT) (Nikolaev et al. 2008)) and compare FGHT instructions number with the fast Hough transform method (Götz 1995). In the second experiment we measure complexity of the generated algorithms for Bresenham's lines and compare it with FHT.

In the scope of FHT it is common to consider straight image lines as either primary-horizontal (PH) or primary-vertical (PV), as shown in fig. 1 belonging to one of 4 groups (quadrants):

– PH with right slope $\alpha \in [-\frac{\pi}{4}, 0]$
– PH with left slope $\alpha \in [0, \frac{\pi}{4}]$
– PV with right slope $\alpha \in [\frac{\pi}{4}, \frac{\pi}{2}]$
– PV with left slope $\alpha \in [\frac{\pi}{2}, \frac{3\pi}{4}]$

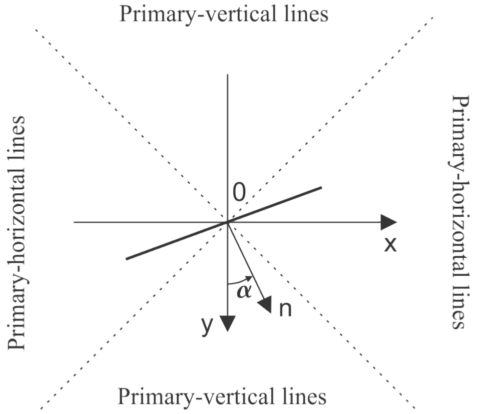


Figure 1: Discrete image straight line types

During computational experiments FGHT generation was performed for all dyadic patterns (Dyadic 4Q), for a half (PH and PV) (Dyadic 2Q) and for a single quadrant (Dyadic 1Q). Apparently, the choice of a particular subgroup for the two latter cases does not matter because of symmetry.

The computational results for three algorithm modifications are given in table 1. Table contains the number of computational operations to compute corresponding Hough transform. One can see that random M element choice version of the FGHT generation algorithm supersedes FHT on operations number per quadrant in all cases. The table shows that greedy M element choice allows to reach a smaller number of operations per quadrant for the Dyadic 2Q and Dyadic 4Q cases, this is explained by appearance of common instructions for different quadrants. The table shows that using the M element choice using maximal matching search does not increase performance even for the Dyadic 2Q and Dyadic 4Q cases.

Random algorithm				
N	FHT 1Q	Dyadic 1Q	Dyadic 2Q	Dyadic 4Q
4	32	36	30	24
8	192	243	223	178
16	1024	1296	1293	1141
32	5120	7330	6773	6399
Greedy algorithm				
N	FHT 1Q	Dyadic 1Q	Dyadic 2Q	Dyadic 4Q
4	32	40	26	24
8	192	208	184	162
16	1024	1024	952	940
32	5120	5120	4784	5083
Maximal matching search algorithm				
N	FHT 1Q	Dyadic 1Q	Dyadic 2Q	Dyadic 4Q
4	32	32	28	23
8	192	244	217	191
16	1024	1457	1048	1151
32	5120	7640	5720	6523

Table 1: Operations number per quadrant for FHT and FGHT generated using random choise, greedy choise and maximal matching search.

During the second experiment FGHT was generated for discrete patterns constructed using the Bresenham algorithm (Butler et al. 1991).

In fig. 2 it is shown how instructions number depends on square image size N for various modifications. The FHT $N^2 \log N$ complexity is given for comparison for dyadic lines of a single quadrant. FGHT instructions number is given relatively to the number of quadrants.

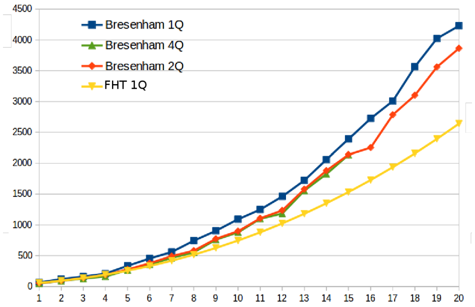


Figure 2: Dependency between instructions number and image size N for FGHT. The Bresenham 1Q curve represents FGHT generated for a single quadrant, Bresenham 2Q - for PH (halved), Bresenham 4Q for all straight lines (divided by 4), FHT 1Q - FHT operations number per single quadrant.

In fig. 2 one can see that FGHT efficiency is slightly less than that of FHT.

It is interesting to notice that computing two quadrants is faster per quadrant than computing just a single quadrant. It means that during two quadrants computation certain common instructions are used, i.e. two quadrants have a greater number of common pairs.

It is not necessarily true, however, that these common instructions are the only reason of greater performance. The gain may also be caused by the fact that when computing two quadrants the first pairs chosen for substitution integrally lead to a more effective algorithm even for a single quadrant. We plan to investigate this question in future works.

The experimental results show that the proposed greedy algorithms do not always demonstrate optimal FGHT performance. This is the case, for example, for images of sizes 4 and 8 when FHT uses less operations.

However, the proposed algorithms do make it possible to generate FGHT which is faster than FHT what apparently means that FHT is not optimal at least for dyadic patterns with $N = 8, 16, 32$.

Table 1 and fig. 2 demonstrate that it is possible to generate FGHT with better performance for dyadic pattern than for the Bresenham algorithm pattern.

We measured generation algorithms execution times on a single CPU core Intel(R) Core(TM) i7-4790 CPU. All three generation algorithms had execution time less than a second for square images with side $N = 4, 8$. However already for $N = 16$ generation time of two algorithms (with randomized and greedy choice) reaches 15 seconds, while the maximal matching algorithm generates FGHT in 1.2 seconds.

For $N = 32$ the randomized, greedy and maximal matching choice algorithms complete computation in 47 minutes, 35 minutes and 4 seconds resp.

These results show that for FGHT generation one may choose between minimizing operations number with greedy choice algorithm showing the best result and the generation time itself where the maximal matching search algorithm wins.

CONCLUSION

In this paper we demonstrated the connection between the FGHT generation and the additive chain problem. We proposed three FGHT generation algorithms for arbitrary patterns and measured their performance in comparison with FHT. It was showed that in certain cases the suggested algorithms outperform FHT by operations number, hence proving that the latter is not always optimal. However, such image sizes exist when FHT has better performance which suggests that the FGHT we constructed is not always optimal either. Based on the obtained results we can state that the FGHT generation approach we suggested has potential and we plan to continue its development.

ACKNOWLEDGEMENTS

We thank Simon Karpenko for advice and inspiration at the beginning of the work. The research was supported by the Russian Science Foundation grant (project No. 14-50-00150).

REFERENCES

Ballard, Dana H. 1981. "Generalizing the hough transform to detect arbitrary shapes." *Pattern recognition*, 13(2):111–122.

Bernstein, Daniel J. ???? "Pippenger exponentiation algorithm." *http://cr.yp.to/papers/pippenger.pdf*.

Brown, Lisa Gottesfeld. 1992. "A survey of image registration techniques." *ACM computing surveys (CSUR)*, 24(4):325–376.

Butler, Nicholas D; Adrian C Gay; and Jack E Bresenham. 1991. "Line generation in a display system." US Patent 4,996,653.

Davis, Larry S. 1982. "Hierarchical generalized hough transforms and line-segment based generalized hough transforms." *Pattern Recognition*, 15(4):277–285.

Duda, Richard O and Peter E Hart. 1972. "Use of the hough transformation to detect lines and curves in pictures." *Communications of the ACM*, 15(1):11–15.

Garey, Michael R and David S Johnson. 1979. "Computers and intractability: a guide to the theory of np-completeness. 1979." *San Francisco, LA: Freeman*, 58.

Götz, H. J. Druckmüller, W. A. 1995. "A fast digital radon transform—an efficient means for evaluating the hough transform." *Pattern Recognition*, 28.12:1985–1992.

Hough., Paul V.C. and Ann Arbor. 1962. "Method and means for recognizing complex patterns."

Illingworth, John and Josef Kittler. 1988. "A survey of the hough transform." *Computer vision, graphics, and image processing*, 44(1):87–116.

Karpenko, Simon; Ivan Konovalenko; Alexander Miller; Boris Miller; and Dmitry Nikolaev. 2015a. “Uav control on the basis of 3d landmark bearing-only observations.” *Sensors*, 15(12):29802–29820.

Karpenko, Simon; Ivan Konovalenko; Alexander Miller; Boris Miller; and Dmitry Nikolaev. 2015b. “Visual navigation of the uavs on the basis of 3d natural landmarks.” pages 98751I–987510I.

Kimme, Carolyn; Dana Ballard; and Jack Sklansky. 1975. “Finding circles by an array of accumulators.” *Communications of the ACM*, 18(2):120–122.

Konovalenko, I; A Miller; B Miller; and D Nikolaev. 2015. “Uav navigation on the basis of the feature points detection on underlying surface.” pages 499–505.

Kuznetsova, E; E Shvets; and D Nikolaev. 2015. “Viola-jones based hybrid framework for real-time object detection in multispectral images.” pages 987501N–987506N.

Merlin, Philip M. and David J. Farber. 1975. “A parallel mechanism for detecting curves in pictures.” *IEEE Transactions on Computers*, 100(1):96–98.

Mukhopadhyay, Priyanka and Bidyut B Chaudhuri. 2015. “A survey of hough transform.” *Pattern Recognition*, 48(3):993–1010.

Nikolaev, D.; S. Karpenko; I. Nikolaev; and P. Nikolayev. 2008. “Hough transform: underestimated tool in the computer vision field.” *Proceedings of the 22th European Conference on Modelling and Simulation*, pages 238–246.

Pippenger, Nicholas. 1980. “On the evaluation of powers and monomials.” *SIAM Journal on Computing*, 9(2):230–250.

Ulrich, Markus; Carsten Steger; and Albert Baumgartner. 2003. “Real-time object recognition using a modified generalized hough transform.” *Pattern Recognition*, 36(11):2557–2570.

Xu, Lei and Erkki Oja. 1993. “Randomized hough transform (rht): basic mechanisms, algorithms, and computational complexities.” *CVGIP: Image understanding*, 57(2):131–154.

Xu, Lei; Erkki Oja; and Pekka Kultanen. 1990. “A new curve detection method: randomized hough transform (rht).” *Pattern recognition letters*, 11(5):331–338.

AUTHOR BIOGRAPHIES

EGOR ERSHOV was born in Moscow, Russia. He studied engineer science and mathematics, obtained his Master degree in 2014 from Moscow Institutee of Physics and Technology. Now he is a Ph.D. student. Since 2014 he is working in Vision Systems Lab at the Institute for Information Transmission Problems RAS. His research activities are in the areas of computer vision. His e-mail address is e.i.ershov@gmail.com.

TIMUR KHANIPOV was born in St. Petersburg, Russia. He studied mathematics at the Moscow State University, graduated in 2008. Since 2010 he works as a researcher at the RAS Institute for Information Transmission Problems. Timur’s research activities are in the areas of technical vision, industrial automation and recognition systems. His e-mail address is timur.khanipov@iitp.ru.

EVGENY SHVETS was born in Chelyabinsk in 1990. He studied mathematics and obtained his Master degree in 2013 in Moscow Institute for Physics and Technology. Since 2014 he is a research scientist at the Institute for Information Transmission Problems, RAS. His research activities are in the areas of robot localization and mapping and computer vision. His e-mail address is vortexd77@gmail.com

DMITRY NIKOLAEV was born in Moscow, Russia. He studied physics, obtained his Master degree in 2000 and Ph.D. degree in 2004 from Moscow State University. Since 2007 he is a head of the Vision Systems Lab. at the Institute for Information Transmission Problems RAS. His research activities are in the areas of computer vision with primary application to color image understanding. His e-mail address is dimonstr@iitp.ru.

High Performance Modelling and Simulation

Modelling and Simulation of Data Intensive Systems - Special Session

Computer Intensive vs. Heuristic Methods in Automated Design of Elevator Systems

Leopoldo Annunziata, Marco Menapace, Armando Tacchella

KEYWORDS

Computer-Intensive Modeling and Simulation, Artificial Intelligence, Computer Automated Design of Physical Systems.

ABSTRACT

Automated design of systems may require modeling and simulating potential solutions in order to search for feasible ones. This process often involves a trade-off between heuristics and computer-intensive approaches. Since neither of the two methods guarantees to always succeed, each problem domain requires a dedicated evaluation. In this paper, the domain of computer-automated design (CautoD) for elevator systems is studied with the goal of providing experimental evidence about which approach is best in which circumstances, and to serve as guidance for automated modeling of elevator systems.

INTRODUCTION

Automated design of physical system — see, e.g., [ZWPG03] — is the process whereby a project of some implement is carried out by computer programs which partially substitute the work of engineers and technicians. At the highest level of automation, the process requires a designer to enter configuration parameters, guidelines and physical constraints only, and the burden of generating feasible designs will rest on computer programs. Since physical systems involve the combination of several different elements to perform their stated function, the problem often becomes combinatorial in nature. In particular, when confronting a huge number of alternative designs, the question arises as to whether heuristics or computer-intensive methods should be leveraged. While heuristics tend to be less demanding in terms of computing power, they might also fail to explore potentially fruitful directions; computer-intensive methods, on the other hand, do explore more solutions than heuristics to achieve accurate results at the expense of computing power. It is well known that heuristics — see, e.g., [Pea84] — cannot guarantee desirable properties such as optimality with respect to some cost function, or completeness with respect to some set of solutions. However, computer-intensive methods often fail to deliver because of an excessive request of computational resources. Since neither of the two methods guarantees to always succeed, each problem domain requires a dedicated evaluation.

In this paper, computer-automated design (CautoD) of elevator systems is considered. Computer-automated design (CautoD) differs from "classical" computer-aided design (CAD) in that it is oriented to replace some of the designer's capabilities and not just to support a traditional work-flow with computer graphics and storage capabilities. While CautoD programs may integrate CAD functionalities, their purpose goes far beyond the replacement of traditional drawing instruments and most often involves the use of advanced techniques from artificial intelligence. As mentioned in [BOP$^+$16], the first scientific report of CautoD techniques is the paper by Kamentsky and Liu [KL63], who created a computer program for designing character-recognition logic circuits satisfying given hardware constraints. In mechanical design — see, e.g., [RS12] — the term usually refers to tools and techniques that mitigate the effort in exploring alternative solutions for structural implements, and this is the flavor of CautoD that will be considered hereafter.

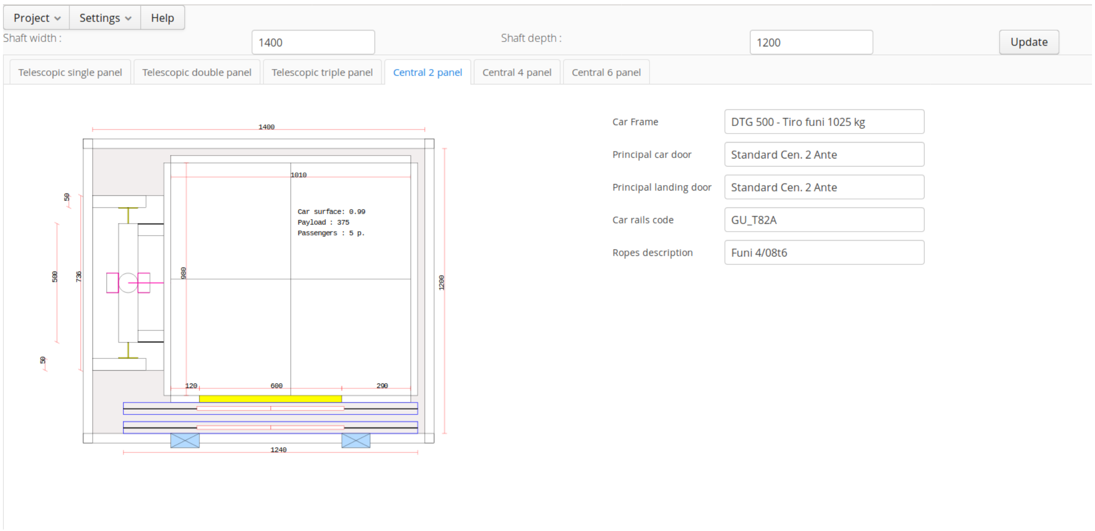
Elevators are complex implements whose design requires the combination of several standard components which must be fitted to custom spatial and usage requirements. Since human designers cannot simulate all possible viable models, they leverage "good design practices", i.e., heuristics, that usually yield reasonable engineering solutions. On the converse, while a program might thoroughly simulate the space of alternative elevator designs, the process is not guaranteed to be computationally feasible. The goal of this paper is to provide experimental evidence to evaluate which approach is best in which circumstances, to serve as guidance for automated modeling of elevator systems. As the experimental results clearly show, the impact of heuristics in pruning the search space of feasible solutions can be dramatic, enabling CautoD programs to achieve a number of good alternative solutions with very little computational effort. Still, a computer-intensive strategy that filters feasible designs using *a-posteriori* reasoning, can find designs that are disregarded by heuristics, but that could still provide useful solution in specific niche conditions.

BACKGROUND AND MOTIVATIONS

With nearly 5 million installations in EU-27 countries as of 2012, elevators are complex automation systems which nevertheless are part of daily routine for hundreds of thousands of non-technical users. It is a fact that most installations are to be found in residential and tertiary buildings, with industrial sites accounting for a mere 4% of the total market share — see [DAHP$^+$12] for more details. In Italy alone, as of 2015, the number of operational elevators was approaching 1 million, with a total market value of about 1.3 Beuro whereof 366 Meuro are coming from new installa-

Leopoldo Annunziata is an independent professional and mechanical engineering consultant for lift builders and contractors. E-mail: l.annunziata@studio-annunziata.it — Marco Menapace and Armando Tacchella are with "Dipartimento di Informatica, Bioingegneria, Robotica e Ingegneria dei Sistemi" (DIBRIS), University of Genoa, Viale Causa 13, 16145 Genoa, Italy. E-mail: marco.menapace@edu.unige.it, armando.tacchella@unige.it. The corresponding author is Armando Tacchella.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)



Shaft width : 1400

Shaft depth : 1200

Car surface: 0.99
Payload : 375
Passengers : 5 p.

Car Frame: DTG 500 - Tiro funi 1025 kg

Principal car door: Standard Cen. 2 Ante

Principal landing door: Standard Cen. 2 Ante

Car rails code: GU_T82A

Ropes description: Funi 4/08t6

Fig. 1: Main view of LIFTCREATE under the guideline which maximizes door size given shaft size. The designer can change shaft width and depth in the top bar, and then ask LIFTCREATE to compute feasible solutions with the "Update" button. Alternative designs appear as tabs, each featuring different types of doors in this case. The plan view showed in the picture is about a hydraulic elevator — piston and car frame are visible on the left side of the car — and a central sliding door with two panels — both car ande landing doors are shown on the front side of the car.

tions[1]. With such volumes, intended users and an expected operational life of decades, it is not surprising that elevators are subject to stringent normative requirements which discipline their design, construction, initial test and maintenance. As products, elevators are considered to be in their "maturity" stage, but the approach to their design, construction and maintenance still shows a lot of room for improvements. In particular, while elevator design is mostly a manual process based on "classical" productivity tools such as CAD programs, spreadsheets and word processors, a push towards innovation comes from the trade-off between relatively small profit margins, and the need to perform an accurate design in order to comply to the above mentioned normative framework. In this direction, automating the design of elevators, while it cannot (and should not) replace certified professional designers, may support them to reduce design time and cost without sacrificing the overall quality of the project.

Currently, only two publicly available products are endowed with some CautoD functionality targeted to elevator design: LIFTDESIGN[2] from DigiPara® and ASCENSORI[3] from ApplicativiCAD. Both applications offer libraries of commercial off-the-shelf components wherewith 2D elevator drawings (plan and vertical views) are generated trying to accomodate physical constraints, designers' choices, and customers' requirements. While LIFTDESIGN can also generate 3D models, it consists of "predefined elevator parameters, component structure and elevator logic" which makes the creation of customized solutions rather difficult. Furthermore, LIFTDESIGN does not provide guidance to the designer amidst alternative implementations, but it just provides warning and error messages when drawing generation is attempted in the presence of conflicting parameters. ASCENSORI provides more support for customization and more design automation than LIFTDESIGN, in that it guides the user through various steps of the design by trying to ban alternatives that will almost surely lead to unfeasible designs. The main issue with ASCENSORI is that it relies on a rather contrived and acronym-laden graphical interface which, together with some maturity issues, severely affects usage by all but the most experienced designers. Unfortunately, neither of the two applications is available for research purposes, so they cannot be used as a platform for comparative evaluations and implementation of new CautoD techniques to extend existing functionalities.

Providing designers with an easy-to-use, yet flexible tool with full-fledged CautoD functionality for elevators is the main aim of the AILIFT suite[4] Currently in its early prototypical stage, AILIFT is expected to group three different but synergistic applications, namely LIFTCREATE to generate structural designs, LIFTREPORT to generate accompanying documentation for installation and certification of the elevator, and LIFTPLAN to generate detailed parts count. The core CautoD functionalities are implemented by LIFTCREATE which takes the designer from the very first measurements and requirements, e.g., shaft size and payload, to a complete project which guarantees feasibility within a specific normative framework. To achieve this, LIFTCREATE works in two steps. In the first step, the user is asked to enter relevant parameters characterizing the project, and an overall "design philosophy" to be implemented. For instance, if the size of the elevator's shaft is known and fixed in advance, LIFTCREATE can generate solutions which maximize payload, door size, or car size. A design philosophy is just a set of guidelines which, e.g., prioritize door size over other elements, still keeping into ac-

[1] Mediacom report on the Italian elevator market (2016). Available from: http://www.anacam.it/.

[2] https://www.digipara.com/products/liftdesigner/.

[3] http://www.applicativicad.it/ascensori.php.

[4] http://www.ailift.it.

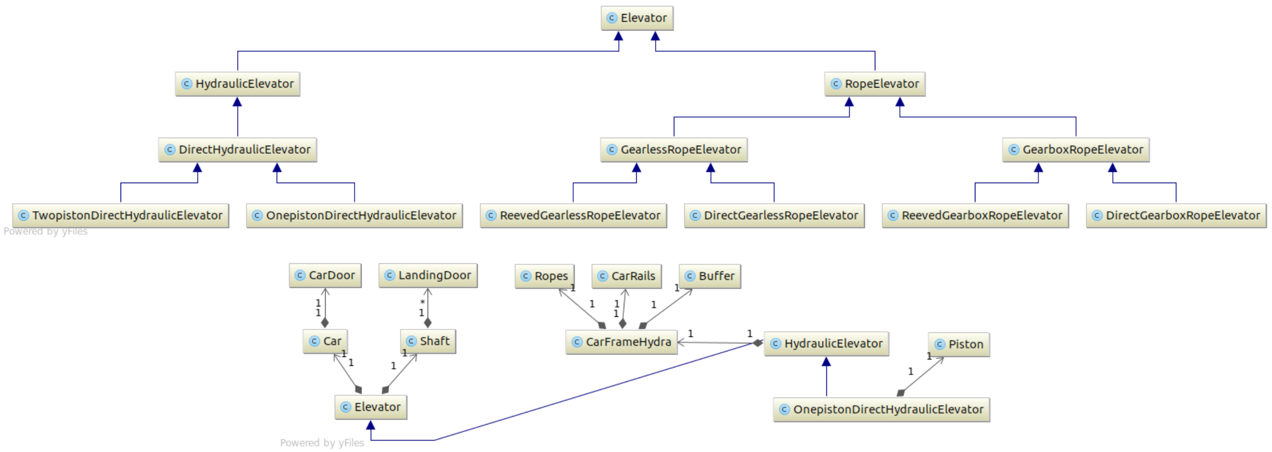


Fig. 2: Taxonomy of elevator types handled by LiftCreate (top) and components of OnePistonHydraulicElevator (bottom). Rectangles represent entities, IS-A relations are denoted by solid arrows, and HAS-A relations are denoted by diamond-based arrows.

count hard constraints, e.g., payload and car size should not fall below some threshold. In the second phase, LiftCreate retrieves components from a database of parts and explores the space of potential solutions. This is exactly where the tradeoff between heuristic and computer-intensive methods becomes relevant, since the solution space grows as $O(p^s)$, where p is the number of different parts to be fitted and s is the maximum number of alternatives for each part. Since the growth of such space is exponential in s, and the number of alternatives for each part can be large — consider, e.g., different suppliers, different mechanisms, and different builds — it is easy to see that the space of alternative designs is subject to a "curse-of-dimensionality" issue. In principle LiftCreate should be able to output optimal solutions, i.e., designs that maximize desiderata within the given constraints. However, for the second phase to be computationally efficient, by default LiftCreate aggressively prunes the search space of alternative designs using heuristics. This implies that the set of designs presented to the user in the end could be incomplete and single designs could be sub-optimal. If enough computational resources are available, LiftCreate can afford to simulate more designs among which the optimal ones can be picked and presented to the user. However, since this implies less aggressive search space pruning, the computer-intensive approach is to be evaluated carefully to avoid blow-up in computation time.

CASE STUDY

As shown in the taxonomy of Figure 2 (top), elevators can be differentiated in two broad categories, namely traction — also called rope in the following — and hydraulic elevators. In traction elevators, the car is suspended by ropes that are moved via an electrically driven sheave. The opposite end of the ropes is connected to a counterweight. Depending on whether the sheave is driven directly by the electric motor or whether a gearbox is used, these elevators are further differentiated into geared and gearless traction systems. According to [DAHP+12], geared traction elevators are the most common "legacy" elevator type in Europe, constituting more than two thirds of the European elevator stock. Gearless traction elevators are a comparatively young technology and only constitute about 8% of the total elevator stock. The remaining elevators operate on hydraulics, i.e., they rely on one or more pistons to move the car. Energy is usually provided to the hydraulic fluid by an electrically driven pump, and typically no counterweight is needed to compensate for the weight of the car. Hydraulic elevators (HEs) are often used in low-rise applications and are widely used in new installations in some European countries, including Italy: Their low initial costs, compact footprint and ease of installation makes them the most viable choice for retrofitting old residential buildings, and a cost-effective solution for new ones alike. The choice of HEs as a case study is thus motivated by their popularity, and by the fact that, in spite of their relative low part count, their structure presents already most of the challenges that are to be found in other elevator types.

The components of HEs considered by LiftCreate CautoD procedures are shown in Figure 2 using an UML class diagram to outline the corresponding part-whole hierarchy. Notice that, in order to manage the space of potential designs components cannot be solely available as drawing elements, like in classical CAD solutions, but they must be handled as first class data inside LiftCreate logic. In particular, OnePistonDirectHydraulicElevator is both a leaf entity in the taxonomy shown in Figure 2 (top), and also the root node of corresponding part-whole hierarchy in Figure 2 (bottom). Looking at the hierarchy, the structure of HEs with one piston direct drive can be easily learned, the only peculiar aspect being that these implements feature only one piston (Piston). The remaining components are common to HydraulicElevator or Elevator. In particular, the car frame (CarFrameHydra), i.e., the mechanical assembly connecting the car with the piston, is specific of hydraulic elevators. Albeit not physically part of the car frame, the entities CarRails, i.e., the rails along which the car is constrained to move, Buffer, i.e., the dumping device placed at the bottom of the elevator shaft, and Ropes, are logically part of it since their type and size must be inferred from or melded with the type and size of the car frame. Common to all elevator types, the entities Shaft and Car are both logically part of the Elevator entity, but only Car is also a physical component, together with its sub-component CarDoor. In the case of Shaft, while landing doors (LandingDoor) are not physically part of the shaft, they are attached to it and their size and type must be inferred from or melded with car doors. The relationships encoded in such part-whole hierarchy are instrumental to LIFTCREATE when it comes to handle drawing, storage and retrieval of designs, but also to reason about the various trade-offs of a design when searching in the space of potential solutions, as described in the next section.

AUTOMATED DESIGN METHODOLOGIES

For the sake of clarity, in the ensuing discussion about LIFTCREATE CautoD procedures for hydraulic elevators it is assumed that only one supplier and build are available for car frames — including all logically-attached components, i.e., car rails, buffers and ropes — and for doors. This is not a severe limitation, as often designers and elevator installers will have their preferred pool of suppliers and builds for car frames and doors, opting for different ones only when the setup requires solutions which are manufactured only by specific suppliers. The CautoD procedure operates according to some predefined parameters:

- Reductions, i.e., distances from car to shaft on those sides of the car which are free from doors and car frame.
- Car wall thicknesses (different values for each car wall).
- Maximum car frame overhang (distance from the central axis of the car and piston).
- Choice of reduced or standard landing door frames.
- Door size tolerances with respect to other components, e.g., car frame.

Finally, it is assumed that the car will have only one door on the front, and that the car frame is to be placed either on the left side or at the back of the car. The case in which the car frame is placed to the right is simmetrical to the one considered.

Independently from whether LIFTCREATE uses heuristics or computer-intensive methods to guide the designer amidst alternative choices, the CautoD procedure scheme is the following:

1. Shaft size (width and depth) is input by the designer; no other configuration parameters are necessary since it is assumed that there is only one door on the front side of the car.
2. All available car frames are considered in ascending payload order; each selected car frame is placed either on the left or at the back of the car, aligned to its center.
3. Taking into account the selected car frame size, car wall thicknesses and reductions, the current internal width of the car is computed.
4. Door selection depends on whether heuristics or computer-intensive techniques are used (see below)
5. For each selected car frame and door, the weight of the car — doors included — and its payload are computed; given also the maximum overhang, it is possible to validate the selected car frame: if adequate, the current solution is saved into a list of feasible designs; otherwise, the solution is discarded and the procedure goes back to step (3).

To complete point (4) in the procedure scheme above, one could resort to either heuristics or computer intensive methods. In the former case, the following steps are taken:

a. The "internal cabin door" parameter — ICD in the following — is computed starting from the value computed in step (3) above, considering the size of landing door frames.
b. In order to select car and landing doors of feasible size, the ICD parameter, the shaft width, and the door size tolerances are considered to perform checks depending on the door types, i.e., sliding and folding; for the sake of brevity, details are omitted, but it is important to notice that such checks involve some non-trivial reasoning about door placement.
c. The doors selected at step (b), together with the selected car frame are part of the evaluation carried out in step (5) of the CautoD procedure scheme.

If computer-intensive methods are opted for, the following steps are taken:

a. Door sizes larger than the shaft are filtered away.
b. For each combination of door size and selected car frame, the parameter "residual car space" — RCD in the following — is computed; the parameter amounts to the difference between the shaft width and the total space allocated for the door; RCD is computed for each door type, since the space available for door placement is clearly a function of the current combination of door, car-frame, reductions and car wall thicknesses.
c. RCD is divided up into intervals of equal length — 5mm in the current implementation — so that a set of projects is generated, each with a different door placement; some of these projects will not be feasible, because door placement will not be coherent with the overall constraints, but these will be filtered at the end of the CautoD procedure.

While heuristics generate projects that are guaranteed to be feasible, the computer-intensive approach requires post processing to filter out remaining unfeasible projects. There are five checks that serve this purpose, namely:

- the door should not be placed outside the shaft;
- the door opening should be contained in the car;
- the car frame should be contained in the shaft;
- there should be no interferences between car doors and car frame;
- the landing door frame, once aligned to the car door, should not be outside the shaft;

If at least one of the checks above fails, the corresponding design is discarded.

EXPERIMENTAL EVALUATION

The experimental evaluation is carried out considering eleven hydraulic elevator case studies — CS in the following. These include both configurations for which there exists feasible solutions, and configuration for which there are none. Among configurations for which feasible designs exist, both typical and "borderline" cases are considered. In more details:

- CS#1 features a shaft size which is too small to have feasible solutions;

- CS#2 is the minimum shaft size to have exactly one feasible solution: clearly the solution found has to be the same across heuristics and computer-intensive methods;
- CS#3-7 represent “typical” shaft sizes found in residential buildings;
- CS#8-10 feature unconventionally large shaft sizes;
- CS#11 features a shaft size which is too large to have feasible solutions.

In all the cases above, feasibility is constrained by the working hypothesis outlined before and by the available set of components. For instance, in CS#11 there are no feasible designs because in the component library there are no car frames available that can handle the resulting maximum payload. The experimental results herewith presented are obtained using LIFTCREATE prototype implemented in Java 8 and based on the SPRING object-persistence framework [5] using Vaadin[6] to generate and display GUIs. To handle components data and generated projects, a local instance of Mysql server 5.7 is adopted. All the simulations are executed on a machine with an Intel i7 5th generation 8 core CPU, featuring 8GB of RAM and running Ubuntu Linux 16.10.

TABLE I: Heuristics vs. computer-intensive methods.

CS	W	D	Heuristics		Computer-Intensive		
			CPU	OK	CPU	GEN	OK
1	900	830	235	0	1504	50	0
2	910	830	147	1	1488	120	1
3	1200	1000	670	122	3260	18972	1777
4	1200	1200	596	174	3736	18078	4181
5	1400	1000	1314	265	4122	40930	5939
6	1400	1200	943	291	3367	36440	11355
7	1600	1000	1606	406	3913	66717	9035
8	1600	1200	1349	470	2801	63854	24935
9	1800	1200	1614	516	2576	73046	28920
10	2000	1200	3628	360	1838	58888	22131
11	2000	1400	4236	0	1266	0	0

Table I presents experimental data about the comparison between heuristics and computer-intensive methods on the selected case studies. In the Table, **CS** is the unique case study id, **W** and **D** are the corresponding shaft width and depth, respectively; both for heuristics and computer-intensive methods, **CPU** is the amount of time (in milliseconds) required to generate solutions, and **OK** is the number of feasible solutions found; **GEN** is the number of generated projects which, in the case of computer-intensive methods, does not readily correspond to feasible solutions, i.e., those that pass the checks mentioned at the end of the previous section. From Table I it can be observed that heuristics and computer-intensive methods do not present substantial differences when it comes to over-constrained configurations. In particular, for CS#1 and CS#2, the only noticeable element is that heuristics are faster than computer-intensive methods, as they can prune many unfeasible designs in the early stages of search — in CS#2 there is a 10× factor between the two. As far as “typical” configurations are considered, the picture changes. The gap in performances is never greater than a 6× factor, and computer-intensive methods are generating a strict superset of the projects generated by heuristics. The difference set is populated by solutions that, albeit feasible, do not correspond to straightforward designer choices, whose spirit is embedded into heuristics methods. Nevertheless, many such designs do have practical value. For instance, since computer-intensive methods explore many alternative door placements, they find solutions which often end up being preferred by implementors because they allow an easier fitting of cables or other implements, whereas customers may prefer them, e.g., because of aesthetic reasons. Finally, as for configurations which admit many alternative solutions, it can be observed that both heuristics and computer-intensive methods struggle with an ever-increasing search space. In some cases, e.g., CS#10, *a-posteriori* pruning techniques implemented in the computer-intensive approach end up being more efficient that heuristics.

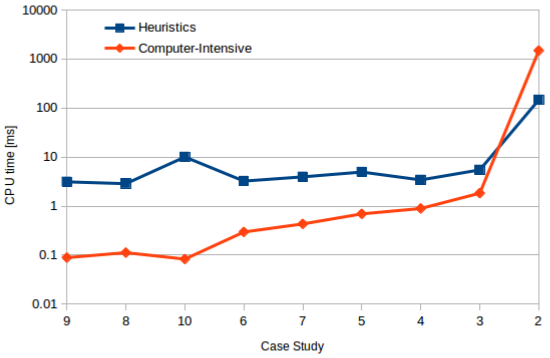


Fig. 3: Average time to compute a single solution; case studies appear along the x-axis, ordered descendingly according to their estimated complexity; y-axis report CPU times in milliseconds, on a logarithmic scale; the plot report the performances of heuristic (squares) vs. computer-intensive (diamonds) methods.

In figure 3 an alternative view of the results shown in Table I is shown. In the Figure, the average time to compute a single solution for heuristics and computer-intensive methods is plotted. For each case study, the average time per solution is just the ratio between the total time and the number of feasible projects — columns CPU and OK in Table I. However, case studies are sorted along the abscissa of the plot considering an ascending order of configuration complexity, whose approximate indication can be obtained considering the number of solutions generated by the computer-intensive approach — column GEN in Table I. The principle behind this choice is that, the more solutions are evaluated by computer-intensive techniques, the less constrained the original configuration is, and the less complex the overall problem is. Notice that CS#1, CS#2 and CS#11 are excluded from the plot in Figure 3 since the corresponding data would not make any sense. What can be observed from the plot is that, when the complexity of the problem is low, computer-intensive methods may have an edge over heuristics: it takes more than 10× time to compute a single solution using heuristics, on average. The gap becomes

[5] `https://spring.io/`.
[6] `https://vaadin.com/home`.

smaller and smaller while the complexity increases. This is to be expected, because computer-intensive methods will have to reject more and more solutions, increasing the overhead to reach feasible solutions. For the most complex design, namely CS#2, the cost of computing 120 solutions in order to reach just one, is overwhelming with respect to what can be achieved using heuristics. This reveals that, all other things being equal, computer intensive methods have an advantage over heuristics when the problem complexity is relatively low, whereas contrived configurations might benefit the most from the pruning power of heuristics.

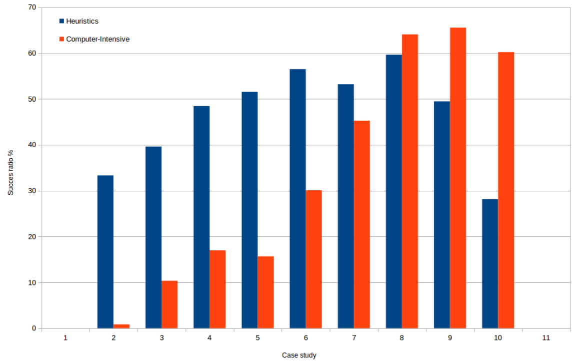


Fig. 4: Success rate (percentage); each bar represents the success rate of either heuristic (blue) or computer intensive (orange) methods.

One last perspective about the results is shown in Figure 4. Here the success ratio of heuristics and computer-intensive methods is computed as a percentage value, considering the number of feasible designs and the number of attempts made to generate them. While in the case of computer-intensive methods this is simply the ratio between columns OK and GEN in Table I, in the case of heuristics the value is computed considering the number of solutions that are "early-pruned", i.e., those which heuristics do not attempt to extend into a full-fledged solution. For all the case studies, except CS#1 and CS#11 which do not admit feasible solutions, Figure 4 tells that the success ratio of heuristics falls below 30% only in one case, i.e., CS#10, whereas computer-intensive methods are more configuration-dependent. In particular, the success ratios of "easy" configurations — namely CS#8-10 — are much higher than relativealy "hard" configurations, e.g., CS#2, but also CS#3-5. This confirms the evidence exctracted from the plot in Figure 3, i.e., that heuristics have an edge over computer-intensive methods on "hard" configurations, whereas "easy" configurations can be within the reach of computer-intensive methods.

RELATED WORK

Searching the literature for contributions related to CautoD, one of the most recent contribution to be found is Bye et al. paper [BOP$^+$16] on automated design . Here the authors focus on the design of offshore cranes and, more specifically on the optimization of parameters related to the design. The spirit of the contribution is very similar to the one given herein, but the domain and the techniques considered are quite different. The same applies also to [RPL$^+$16], wherein the exterior lighting is subject of the CautoD procedure.

Other recent related works include [WRWS15], where a cloud-based environment is proposed to support design and manufacturing. This contribution is somewhet related in the sense that leveraging a cloud deployment is an expected development in order to achieve improvements in the design process, e.g., by pooling solutions and allowing designers to browse solutions by problem similarity. In [CRS$^+$13], the problem of knwoledge representation is considered in automated designed is considered. Since AILIFT uses an ontology of components to represent the elevator structure as well as the relations between the design componets, it falls in the taxonomy of potential representations outlined in that paper. While [CRS$^+$13] considers many other alternatives to (formal) ontologies, the investigation of potential improvements in this direction is left for future research.

CONCLUSIONS

Summing up, this paper compares heuristics and computer-intensive methods in the automated design of hydraulic elevators. The main lessons learned from our experimental evaluation are the following:

- Heuristics are comparatively faster than computer-intensive methods on contrived configurations and generate a manageable number of solutions among which the preferred one can be picked even manually; however, they do miss "good" practical solutions that can be found by computer intensive methods; this makes heuristics a good default choice in a CautoD procedure whose main purpose is to provide feasibility assessement and budget estimates, run by personnel in non-technical businness units, e.g., sales department.
- Computer-intensive methods are able to find feasible solutions that are missed by heuristics with an increase in CPU time which is not overkilling in most cases; while the number of generated solutions cannot be managed manually, suitable helper procedures can be implemented to allow designers to navigate among the solutions in order to perform design fine tuning for niche configurations.

Considering future extensions and, more specifically, elevators with more than one door, it can be observed that the number of feasible solutions given n feasible single door placements grows as $O(n^2)$. In this case the interplay between heuristics and computer-intensive methods can address the problem of generating enough configurations to be analyzed, yet prune rapidly those that would not lead to feasible designs and would therefore clutter the efficiency of the overall process. These issues are even more compelling when considering a number of different componenent suppliers. Indeed, while the experimental analisys herewith presented is limited to only one set of components from one supplier, common practice involves mixing and matching components from several suppliers in order to meet structural as well as economical constraints. The implementation of the full AILIFT suite, including a complete database of components and suppliers, will enable furthering the analysis along these lines.

REFERENCES

[BOP+16] Robin T. Bye, Ottar L. Osen, Birger Skogeng Pedersen, Ibrahim A. Hameed, and Hans Georg Schaathun. A software framework for intelligent computer-automated product design. In *30th European Conference on Modelling and Simulation, ECMS 2016, Regensburg, Germany, May 31 - June 3, 2016, Proceedings.*, pages 534–543, 2016.

[CRS+13] Senthil K Chandrasegaran, Karthik Ramani, Ram D Sriram, Imré Horváth, Alain Bernard, Ramy F Harik, and Wei Gao. The evolution, challenges, and future of knowledge representation in product design systems. *Computer-aided design*, 45(2):204–228, 2013.

[DAHP+12] Aníbal De Almeida, Simon Hirzel, Carlos Patrão, João Fong, and Elisabeth Dütschke. Energy-efficient elevators and escalators in europe: An analysis of energy efficiency potentials and policy measures. *Energy and Buildings*, 47:151–158, 2012.

[KL63] Louis A. Kamentsky and Chao-Ning Liu. Computer-automated design of multifont print recognition logic. *IBM Journal of Research and Development*, 7(1):2–13, 1963.

[Pea84] Judea Pearl. Heuristics: intelligent search strategies for computer problem solving. 1984.

[RPL+16] Hugo Rocha, Igor S Peretta, Gerson Flávio M Lima, Leonardo G Marques, and Keiji Yamanaka. Exterior lighting computer-automated design based on multi-criteria parallel evolutionary algorithm: optimized designs for illumination quality and energy efficiency. *Expert Systems With Applications*, 45:208–222, 2016.

[RS12] R Venkata Rao and Vimal J Savsani. *Mechanical design optimization using advanced optimization techniques*. Springer Science & Business Media, 2012.

[WRWS15] Dazhong Wu, David W Rosen, Lihui Wang, and Dirk Schaefer. Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. *Computer-Aided Design*, 59:1–14, 2015.

[ZWPG03] F Zorriassatine, C Wykes, R Parkin, and N Gindy. A survey of virtual prototyping techniques for mechanical product development. *Proceedings of the institution of mechanical engineers, Part B: Journal of engineering manufacture*, 217(4):513–530, 2003.

LEOPOLDO ANNUNZIATA graduated from the University of Genoa, Italy with a "Laurea" (MSc equivalent) in Mechanical Engineering in 1995. He is an independent professional and mechanical engineering consultant working for elevator industries and building contractors. He designed about 500 working elevators in more than 10 years of activity in the field, and he is actively involved in the development of AILIFT-LIFTCREATE prototype since its early development phases.

MARCO MENAPACE graduated from the University of Genoa, Italy with a "Laurea" (MSc equivalent) in Computer Engineering in March 2016. He is now a Research Engineer at the Department of Informatics, Bioengineering, Robotics and System Engineering. He is currently the main developer of the AILIFT-LIFTCREATE prototype with an expertise in Java-based OOD/OOP and techniques for cloud and web-based application deployment.

ARMANDO TACCHELLA graduated from the University of Genoa, Italy with a "Laurea" (MSc equivalent) in Computer Engineering and a PhD in Computer Science and Engineering. Dr. Tacchella was a research associate at Rice University in Houston (US) and then a research fellow at University of Genoa, where he became associate professor of information processing systems in 2005. His research interests are in the field of artificial intelligence and formal methods applied to system engineering.

EXTENSION OF BANK APPLICATION SCORING MODEL WITH BIG DATA ANALYSIS

Dr. László Madar
Institute for Training and Consulting in Banking
H1011 Szalag u 19., Budapest, Hungary
E-mail: lmadar@bankarkepzo.hu

KEYWORDS
Big Data, Social Scoring, Scoring Model,

ABSTRACT

Application scoring models of credit institutions has been subject to research since the 1960s. Micro and SME lending has also been improving, however current application scoring models for smaller firms have still lower power statistics as models for private individuals or larger firms. This portfolio has not so strong financials and have less bureau collected behavioral information that makes individual assessment of these firms a hard job for credit institutions. This paper presents an actual example on how external unstructured information can be collected, assessed and used in order to increase performance of this portfolio segment.

INTRODUCTION

During the 1960s, several important advancements were made in the area of credit risk assessment. It was the time when rating and scoring systems for credit institutions were first written in a way that these institutions could use them for differentiation good and bad clients in a procedural way (e.g. Altman (1968), Orgler (1970)). Institutions were given a toolset to improve lending process and if data was collected and analyzed correctly, improve the quality of their credit portfolio. Extent of this improvement was different for each portfolio segment credit institutions have, as different amount of reliable data sources were available.

Application scoring systems, where the credit institution first meets its clients always bear the most risk in term of credit risk, as for already approved clients behavioral models can provide exact repayment probabilities at an individual basis, resulting in very high discriminatory power figures. Application information is usually pooled for a more or less homogeneous pools of clients, enabling to treat them as a portfolio. However, if institutions wish to be more precise, they can assess also clients on an individual basis.

In case of private individuals, institutions do not only assess their socio-demographic characteristics, they rather include their individual behavior. By using credit bureau data on how that individual client pays other loans, bills, how much is the debt burden, the credit risk of that individual can be assessed far more accurately than using only socio-demographic variables. This additional accuracy comes from the inclusion of more data that can be bound to the individual behavior.

In case of firms, inclusion of individual payment information is far more difficult. In case of established firms with existing loans bureau information can be accessed and included in the analysis. Also, for large corporation it is worth the bank to assess the corporation on an individual basis. In case of smaller firms, in the micro and SME segment, this gets trickier. Loan sizes are small, individual treatment of clients is expensive, financials are not as solid as is case of larger firms. However, if and institution can lock to itself prospering firms at an early stage, it might get before their competitors. Competitors can win over these companies only with significantly better conditions to make these companies to switch banks. The initiative and profit will remain at credit institutions who at an early stage can differentiate their clients better.

In this paper I introduce a logic that helps institutions to obtain an individual-basis information about SME clients from various data sources that can be applied in the process of lending and credit assessment. This can be used as an augmentation of the current business process as it processes additional information to that is already available. My method is that we collect unstructured information about clients available on the World Wide Web from various sources and process them in a way that it will predict the firm's future creditworthiness, even in cases where the firm have not applied for a loan before. This is the part when we introduce the now trending big data analysis. In our interpretation, big data analysis helps us obtain information from unstructured data. As we include information on an individual basis, the discrimination of clients will be better as just looking at their mere financials.

Collected unstructured data is processed using text mining methods, and a targeted, simplified semantic analysis to determine the context of the piece of unstructured information collected. The complete process enhances the predictive power of the model as described in detail. This logic was performed using Hungarian language that belongs to a rather complicated family of languages that can be hard to crack with semantics. However, using the targeted approach described, this procedure can be used with any languages on the world.

This presented approach can be used by credit institutions that are willing to invest in unstructured information collection and analysis to improve their lending process for micro and SME corporations.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

MODEL EXTENSION LOGIC

This section presents the approach I followed by specifying, building and testing our approach.

The basic micro and SME model used in this paper was a model I built previously, described in the PhD dissertation in Madar (2015). In this reference work, an SME/micro rating model was developed using financial information of small-medium firms. The developed model was a crisis-proof, stable rating system, where the rating logic itself did not cause cyclicality or movements in the capital requirement, the ordering of the model was stable during crisis years.

In this paper I will use the ROC-Curve and its index number, the AUC measure to assess discriminatory powers. Further information of this measure can be seen in BIS (2005). There might be other measures to be considered, however, all indices show roughly the same changes, as it is a real-word data.

Our overall starting AUC measure that we will use to show the discriminatory power was 0.718 (with a 95% confidence interval of 0.707-0.729), which is an average value for these types of firms, using only financial information to assess their creditworthiness. This model was built using stepwise logistic regression, containing six variables as described in the dissertation.

This basis model was extended by an additional module that contained the result of a big data analysis. The two models were combined with an ensemble approach to achieve optimal combination.

The unstructured data we collected came from various sources. My main source for unstructured data was the social media sites of SME/micro corporations – if they existed. Unlike private profiles of individuals that are well protected by privacy filters, feeds on pages of firms and corporations can be accessed easily, likes and comments, social media activity can be traced. From social media sites, Facebook was processed providing a view of the corporation. For further extension, other social media sites and

To deriving data from these sources, we depended heavily on the algorithms and tools provided by Russel (2015). Data queries used for the current work were based on processes and samples provided by this book.

Lack of unstructured data sources is also an information that adds to the assessment of the firm. It displays either that the firm is too small to have a social media responsible (micro companies with 1-2 employee) or they do not see that as a source of customer acquisition.

Overall, those firms that have web presence is mostly a good thing. Web articles of SME are mostly have a strong positive impact or a strong negative impact. A smaller company is news only when something big is happening around it – getting some investment, boosting sales, having a successful market entry and rocketing sales or when they are negatively impacted, have financial difficulties, have a legal dispute or have to close business just to name a few. Social media sites provide a more sophisticated picture of the corporations. Most SME companies have some social media presence, it is regularly a Facebook presence in Hungary. Activity on the Facebook page displays the SME's commitment to marketing, public relations and social management of clients. Generally SMEs with products for the public are available on the social media, however in some industries it is not so widespread (such as building industry, chemistry, etc., where most of the firms have a limited number of potential customers or customer recruitment is far more effective using non-electronic ways).

This information was analyzed using text mining approach. Word tokens were built from the articles, posts and comments and tokens were aggregated to obtain tokens of similar meaning. Most common tokens (tokens with the most number of counts) were classified manually into three categories: positive, negative and neutral meaning – i.e. a vocabulary was built reflecting the relation of a word to the state of the company in terms of creditworthiness. A positive token would indicate an increase, development, funding of the firm, a negative token would indicate a loss, drop or shutdown. This was the most tedious part of the model building, as in Hungarian language plenty of words can be stemmed from the same root. Unflagged tokens and many common word tokens were treated as neutral, having no impact on the meaning of the article, post or comment collected.

Using this approach, the text of each collected item were classified and could play part of a second tier analysis, where number of posts/likes/comments/articles and its positive/negative classification were assessed. Final assessment of the extension module came from a logistic regression, assessing all information came from the approach described above.

The logistic regression model using financial variables and the logistic regression model using the above analysis were combined using an ensemble logic.

Final model was assessed using the standard AUC measure that showed significant improvement of the discriminatory power. The power of the model improved well outside the 95% AUC confidence level. Based on the effort and sophistication of additional data collection, our results might be further improved.

DATA

In this section, data sources for all model components are described.

Financial data of firms

Financial data of firms

In Hungary all of the data for corporation are public. However, compiling a database from these data takes time, as it can be queried one firm at a time. Data of the complete Hungarian firm register can be queried from the firm service the Ministry of Justice online (http://www.e-cegjegyzek.hu/). This contains the basic information about the corporation, firm type, establishment date, owners, etc. Financial information for each financial year is available also on the homepage of Ministry of Justice (http://e-beszamolo.im.gov.hu/oldal/kezdolap). Data is available mostly in PDF format, however for firms providing their financials in an electronic format, data is available in table format online. Negative information (used to define the default of counterparties) can be accessed on the homepage of the Hungarian Firm Registry Court under liquidations, bankruptcy (http://www.cegkozlony.hu/gazdasagi_ugyszak), and additional negative information is available on the homepage of the tax authority for those with tax payment arrears, enforcement proceedings, suspended tax numbers (http://www.nav.gov.hu/nav/adatbazisok/adoslista).

Although this information is freely available, composing a large enough database takes time and effort, so Hungarian credit institutions rather delegate this task for specialized data provider companies that collect and manage these types of information.

We used the databased compiled for Madar (2015), to further improve with data described in the next section.

Unstructured data for firms

In case of social media sites, data extraction from Facebook followed the logic decribed here. Firms could be in most cases identified by running a search query through the API and select the page with the most number of likes/comments/posts. In case of Facebook that is the most widespread social app in Hungary, a Facebook search query had to be run first and the page IDs and total number of fans had to be queried using the Facebook graph API.

Data from sites were available at structured JSON format that could be deserialized (decompiled) into ordinary SQL database table content to be used in modelling. The complete page feed could be accessed generally by getting first all post id-s belonging to the page and then querying all posts and then querying all comments to these posts. For example in Facebook we used the Graph API to access this information (an access token (denoted as <TOKEN> from now on) is needed for it. I created a web application (available at http:///www.bankarkepzo.hu:15555) to get the queries with its own type of application token. At first step, only page id and the related posts were queried (for example https://graph.facebook.com/ 100878286637188/feed?limit=1000&fields=id,message, created_time&access_token=<TOKEN>). There are a lot of posts, we limited the first 1000 occurrences. However, at the end of the obtained JSON data a link is provided by Facebook for querying the next 1000 occurrences, and so on, until the data block becomes empty. At the second step, we queried 1) the post text themselves and its 2) likes and 3) comments using three different query (e.g. for one post: https://graph.facebook.com/100878286637188_1031983890193285?access_token=<TOKEN>;likes:https://graph.facebook.com/100878286637188_1031983890193285/likes?limit=1000&access_token=<TOKEN>;comments:https://graph.facebook.com/100878286637188_1031983890193285/comments?limit=1000&access_token=<TOKEN>). Text from post and comments are stored in unstructured format, number of likes, number of comments, etc. are available for all feed items, showing a time series of information about the page.

Unstructured data storage was simple, text information from all various sources were stored in the same result table. Text was saved in a long character object and flagged, where it was coming from (web or any of the social media sites), what type of text it was (article, post, comment) and which corporation it belonged to and what weight it got, how relevant was the match of the company name.

Structured data was stored in a timely basis, for each day a summary was generated about the number of contents found (posts, news, likes and comments).

All these compiled dataset enabled us to mine a large amount of data, and extend the scoring model using information from these unstructured data source.

The following table provides an overview of some database totals of the data collected:

Table 1: Collected data sample size

Data source	Type	Count
Facebook	Post	141 688
Facebook	Comment	2 408 190
Facebook	Like	5 298 308

The table above shows that for each data source a reasonable amount of information could have been collected.

However, for about 80% of all Hungarian micro and SME firms, not a single entry was available. This could be a sign of weakness of the approach, however, a number of these firms are dormant, serve as a privately held company controlled by a few people, having no active side bank contact (only marginal interest expense can be observed).

DATA

To address the information hidden within the unstructured data, the collected information had to be processed by using text analytic approach. To uncover some meaning of the text, we had to introduce a quasi-natural language processing logic that could interpret the general context of the article/comment.

Processing of free-form text can be done using two different approaches: rule-based text mining, where statistics are generated from the text for modelling, and linguistics-based text mining, where – using a simple or complex dictionary – some interpretation of the text is made before the modelling starts. As simple rule-based mining can be misleading due to the fact that free-form text tends to use different words and/or phrasing for the same concept. Linguistic methods have an advantage of mapping synonyms or even similar concepts to a common reduced dictionary, and using this common set of words, modelling can be more precise as using the natural form of the free text.

Both approach starts with the tokenization of text. By creating text tokens (i.e. 1-2 word elements without punctuations), the text is formalized and can be processed by text mining approaches. Normally, words are selected as tokens from the text and reduced to their roots. Our source language, Hungarian, is an agglutinative language, i.e. it uses affixes (attached to the word) to change the meaning of the word root, and show tense, plural, etc. Tokenization itself therefore needed a further step to derive etymons, the root meaning of the word that could be classified further. This was made using a statistical approach: the typical affixes were searched within the word (either from the beginning or from the end), and by purging them, if a meaningful word remained that was found within the existing token list, the affix could be deleted, and the token could be changed to the purged token. A token mapping was created that enabled us to classify only the etymons, and not a definitely larger variety of word tokens.

As there is no Hungarian text mining dictionary available at the time, or own mapping had to be created using the derived tokens. We only categorized the most used text tokens, as they were relevant in the business language. The text mining needed a dictionary, and we created a very simplified dictionary of terms. Adjectives, nouns and verbs were classified having either positive or negative meaning. We kept our text context target simple so that the model we build would be more robust. This also enabled us to avoid a creation of a complete linguistic dictionary.

Using this approach, all of the words were classified into a positive, negative or neutral context. Economic articles were the best fit for text mining as they use a type of language that is more or less free from negations (e.g. by stating "not increasing" instead of "decreasing"), sarcasm, jokes, bad grammar and other characteristics of the human language that obfuscate the true meaning of the words. On the other hand, comments, tweets, and sometimes posts are using a language that are sometimes very hard even to tokenize, because they use sometimes abbreviations, have grammatical errors, accent shortage (typical in Hungarian language), sarcastic phrases. This lead to our approach of classification where tokenization and classification of simplified tokens were conducted on web articles, and extended to other collected unstructured data. This resulted that the proportion of positive/negative terms found in non-article sources were lower than in case of articles, but only comments/tweets with proper phrasings were taken into account (however, we had no secret weapon against sarcasm, if well-formulated, it can blur the results of the analysis).

After the assessment of text entries were made, the database was extended a count of negative and positive items within the text record. It was now time to generate explanatory variables for scoring.

As articles/posts/feed have a temporal distribution, beside the general positive/negative state of the client it was also important how the perception about this firm has changed over time. Therefore we built temporal variables measuring the change of comments and their trends in positive and negative relation. Numbers were not absolute as firms might get an increased social/web activity, therefore the change in the ratio of positive and negative comments were measures (from 100% of all text, how many was positive or negative in a given timeframe and with how many percentage points did it increase/decrease).

Generating explanatory variables for defaulted firms was not easy. Normally, social media activity and even the social media pages are closed when a firm declares bankruptcy and data that was available prior to the proceedings is no more available. However, web articles here are also insightful, and in addition to that, social media pages of news portals and news aggregators still have references to defaulted firms. This might lead to a positive bias for firms that are still alive, as they have more positive content over the social media sites as defaulted ones.

Our final list of analyzed variables from the free form text and social media data are the following:

- Overall positive proportion of text
- Overall negative proportion of text
- Overall net proportion of text
- Change in positive proportion over last 3/3 months
- Change in negative proportion over last 3/3 months
- Change in number of likes/+1s over last 3/3 months

It shall be emphasized that we did not only search for correlation within the unstructured data, I tried to build predictive variables that can foresee the future of the firm at a predefined period of time. Similar to behavioral scoring developments at credit institutions, where the payment arrears are the best performing variables, having also a short term predictive power, we expected the same with the collected individual unstructured data. If the firm gets into trouble, customer satisfaction falls right before bankruptcy, and this can provide an additional information to predict the firm's default status. However, predictive power might be also as short lived as payment arrears, as the firm can go bankrupt in a very short amount of time, therefore – also as in case of behavioral scoring – these variables might be used only as a warning signal in credit monitoring, rather than showing the general riskiness of a client.

MODEL CREATION

After the variables have been created, they were processed similarly as any other variable in the logistic regression described in Madar (2015). To avoid the drop in the discriminatory power due to non-linear effects, variables were categorized along their general level of riskiness and a WOE (Weight of Evidence) transformation was performed, to calculate the log odds for each category created. This allowed also the use of logistic regression in the model.

Single factors had low discriminatory power overall, due to the fact that not many firms had observable web activity from the selected sample. Even firms with larger revenue were affected in less web-intensive sectors (e.g. in Hungary building industry, blue-collar services), most of them lacking social media presence and no articles were written about them. The most powerful single factor was the overall negative proportion and change in negative proportion over the last 3 months. These two can predict the defaults well – if only whose corporations are selected that have social media presence, single factor power is in the 30-40% GINI region, depending on categorization. This seems good, however, in relation to the total population the GINI is in the 5-10% region, adding only a bit to our knowledge about the clients.

After the single factor several models were run, bankruptcy as the target variable. Modelling was made using IBM Modeler and models were compared using standard scoring comparison methods. Beside logistic regression, decision tree and neural network models were put together to analyze that the variables are selected and treated appropriately. All models were built on the same dataset, and so, a separate model was created beside the one used as our starting point.

The results of the model was less than expected, the final text model using logistic regression provided a GINI of 14.3%, however, about 80% of the clients were classified in the "unknown" category having no web presence.

Using the ensemble logic of the IBM Modeler, a joint model was created using both logistic regression models, the financial and the unstructured big data source, and the resulting model was also tested. Test showed that discriminatory power increased in a small extent, it is outside the 95% significance limit, so we can say that the extension model brought in additional knowledge of these firms. It provides more insight, however discriminatory power is not so much elevated for a complete bank portfolio. However, if we again reduce our scope to those SME and micro firms that have web presence, the rise of the discriminatory power is in all terms significant, especially negative articles and web posts about firms predict in most cases their downfall.

As a result of the model extension, simple models can be created from big data, and using the approach described above, even further variables might be built to get a more detailed view of the portfolio. The final model has an extended statistical performance, and therefore provides and additional information source for credit institution to profile their corporate clients according to their expected riskiness and provide credit to those companies that have on the one hand better financials, on the other hand have better web perception, good online satisfaction.

This model extension adds to the financials as it is most likely measures the service quality and customer relations quality of a company. Limits of these approach are seen in that not all firms are active online, and they are in the grey zone of this extended analysis. However, the trend is clear, the proportion of clients with online presence is growing year by year..

CONCLUSION

Extension of a financial model with big data analysis is not a simple task. Data has to be accessed, queried and stored in a semi-structured format so that data mining (text mining) is possible on them. Text mining itself has its own challenges as languages might be tokenized with a different efficiency.

The paper showed a process that can be followed to collect and analyze unstructured data about firms, and use this data to assess the quality of the company using their web media content, articles, posts and comments about that firm. This information is out there and can be processed by interested third parties, such as credit institution to get an insight about the firms' online perception.

Power of the model is somewhat low, however we shall note that having a correlation is not always equal to having a prediction (i.e. a random variable n is perfectly correlated with 1-n, however we cannot predict their future state). Big data model in this case seems to extend our knowledge a bit, by introducing new variables that have some relation to the future performance of the company.

We see that there is more room in this topic to create better performing models. There are four points were performance could be increased. First, we can focus on a simpler language and country with deeper web penetration. This would enable to classify more clients from the grey zone. Second, a more detailed token classification can be made, instead of positive/negative, a more differentiated target can be set (e.g 'decreasing' and 'catastrophic' can be separated within the negative section). Third, more explanatory variables can be built that might add to the power of the model. Fourth, the time factor is important, if we collect data in an ongoing manner, we can also collect information about pages where history is not available or pages that ceased to exist due to their bankruptcy.

M. A. Russell, "Mining the Social Web, 2nd Edition". O'Reilly Media, 2013.

L. Madar, "Effects of scoring systems to the results of economic capital models in the institutional capital calculation". PhD Dissertation, University of Kaposvár, 2015.

E. I. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", The Journal of Finance Volume 23, Issue 4, September 1968, pp 589–609

Y. E. Orgler, "A credit scoring modell for commercial loans", Journal of Money, Credit and Banking , Vol. 2, No. 4, Nov., 1970, pp. 435-445

LÁSZLÓ MADAR was born in Budapest, Hungary and attended the Corvinus University of Budapest. He completed his PhD at University of Kaposvár in the field of economics. He is the partner consultant at the Institute for Training and Consulting in Banking (Bankárképző) and a part-time lecturer at Corvinus University of Budapest, Faculty of Finance. He researches big data and social scoring since 2015. His e-mail address is: lmadar@bankarkepzo.hu

IMPROVING MESSAGE DELIVERY IN VEHICULAR AD-HOC NETWORKS

Nnamdi Anyameluhor
Evtim Peytchev
Javad Akhlaghinia
Department of Computer Science
Nottingham Trent University
Clifton Lane, NG11 8NS, Nottingham, United Kingdom
E-mail: Nnamdi.Anyameluhor, Evtim.Peytchev, Javad.Akhlaghinia {@ntu.ac.uk}

KEYWORDS
Framework, Model, Vehicular Ad-hoc NETworks, VANET, Wireless Communication, Network Simulation.

ABSTRACT

Road traffic information has been a primary source for traffic management, user services and other systems that enable road congestion prevention and control; however, road traffic information has not been adequately exploited as a part of the design for wireless communication network. In vehicular ad-hoc networks, the protocol design must consider the dynamic nature of the topology and the probability of available alternate routes for wireless routing. The information provided by either the road itself or the activities on the road can help to void common issues such as broadcast storms, hidden node problems and lost data caused by an increase in road traffic density or sparse road traffic respectively. Routing protocols must therefore be able to dynamically adjust to the current road traffic information. In order to improve message delivery in vehicular ad-hoc networks, this project proposes a collaborative process of utilizing real time road traffic information and route knowledge to enhance routing decisions in order to maximize packet delivery ratio and reduce delays in transmission.

INTRODUCTION

Intelligent Transportation Systems has become an important field because of the growth of wireless networking technologies and devices as well as the increase of transportation vehicles. However, vehicular ad-hoc network is still an emerging technology aided by number of research and issued standards, which still requires some further work to attain the level at which it could provide its all promises. Vehicular ad-hoc networks are wireless networks that can be formed with or without infrastructure, comprised of vehicles with wireless capabilities in a self-organizing manner. Traffic management, road safety aid, entertainment, and user targeted services (Morsink et al. 2002, Wischhof et al. 2003) are among the uses of vehicular ad-hoc networks.

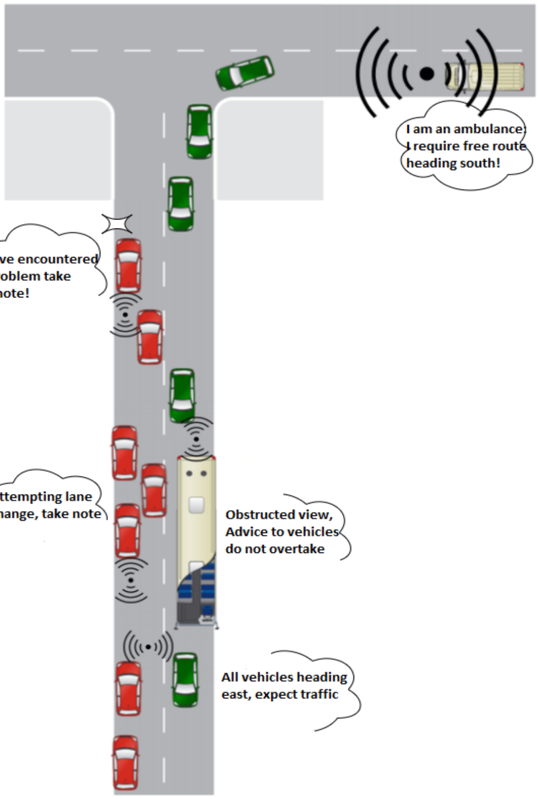

I am an ambulance, I require free route heading south!
I have encountered a problem take note!
Attempting lane change, take note
Obstructed view, Advice to vehicles do not overtake
All vehicles heading east, expect traffic

Fig 1: VANET in operation

Message delivery in vehicular ad-hoc networks is affected by factors such as the speed and mobility of the vehicles, i.e. topological fluctuation, as well as the number of participating vehicles in an area, i.e. vehicular density. These factors impose serious and unique challenges on researchers in order to design encompassing solutions capable of handling various possible scenarios. Models to be designed must consider the unique characteristics of vehicular networks as well as issues that may arise, such as broadcast storms where transmission attempts surpass the available bandwidth due to a high vehicular density. The result of a broadcast storm will be network congestion and added delays caused by packet collisions. On the other hand, very low vehicular density results in sparse networks that might not be suitable enough to guarantee message delivery

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

from one point to another. In this paper, a Framework for Improving Message Delivery in VANETs (FIMDEV) is proposed and discussed. The aim of the framework and model is to assist message delivery from one location to another by using current road traffic information as it relates to the wireless network condition of that area. Among the aims of the framework includes low delays and some level of guarantee regardless of the routing protocol used in the dissemination process. FIMDEV, in conjunction with directed propagation will ensure messages avoid congested routes on their way to their destination. It also employs the use of furthest node reachable as well as the use of established traffic routes where applicable. The rest of this paper is organized as follows; section II discusses some work relevant in this area, section III describes the framework, section IV discusses the performance through evaluation and results and section V concludes the paper and provides possible future work.

II RELATED WORK

Message dissemination in vehicular ad-hoc networks is conducted either through the spread of one-hop messages, e.g. the ones spread periodically by neighbouring vehicles to share relevant updates such as travel data or cluster formation data and are not sent to other vehicles further than a hop away. Message dissemination is also handled through multi-hops, in this case messages such as the spread of safety related messages or value added services may need to transverse beyond the area of origin and will therefore need to travel beyond one hop radius in the wireless network. The IEEE 1609.4 standard based on the 802.11p specifies multichannel operations at the 5.9 GHz band. This band is divided into Control Channels, and Service Channels ready for safety and non-safety applications. One-hop safety messages using this standard are generated periodically at a typical rate of 10 Hz in VANETs to provide updated information about traffic conditions while multi-hop transmissions rely on smart routing protocols for efficient message delivery. J Park and Y Lim proposed RSMB, Reliable and Swift Message Broadcast Method for vehicular ad-hoc networks. In this model, the authors use duplicates of the same message to increase the probability of that message transmitting to the next relay on its way to its final destination. The next relay was selected in such a way as to transverse the network quickly. Although the duplication of messages can increase the chances of vehicles' reception and it is done selectively, it is unclear how scalable the model is under real scenarios. This model showed good results in its evaluation. Fogue et al. in (Fouge et al., 2012) showed that the use of roadmaps in an innovative way that could help in improving message delivery. The protocol called enhanced Message Dissemination for Roadmaps, eMDR builds upon a prior protocol called enhanced Street Broadcast Reduction (Martinez F. J. et al., 2010) which also used information from maps and GPS to alert vehicles within the vehicular network. The eSBR protocol works by choosing the farthest vehicle from the sender on the map, this method works well when all routes are not congested and the aim is to spread the message as far as possible. In the case of eMDR extra controls are put in place to avoid retransmissions by preventing vehicles from broadcasts, allowing only the node closest to the center of a junction to retransmit any message. Multi-hop vehicular broadcast, MHVB (Osafune et al., 2006) was proposed as a flooding solution for vehicular ad-hoc networks to efficiently disseminate safety application information, such as the positions and the velocities of the vehicles within a small geographical area of up to 300m while achieving maximum delays of up to 0.5s. It is presented as a flooding protocol with two main algorithms which are; (i) a Congestion Detection algorithm which reduces unnecessary packets due to vehicular congested traffic by limiting packet spread and (ii) Backfire algorithm which efficiently disseminates the messages through the network by selecting the right receiver node based on the distance from the original node. Adaptive Traffic Beacon (ATB) by Sommer C. et al. (2011), presented a message dissemination protocol which uses two main metrics to adjust the rate at which beacons are transmitted. The metrics are; channel quality and kind of message to be sent. From results shown, the protocol compares well to pure flooding broadcasts, albeit done at a slower rate. It has a knowledge database with important information shared through the use of previously described adaptive beaconing technique in order to maintain a congestion free network. In Cross Layer Broadcast Protocol, CLBP (Bi et al., 2010) the same principles of using knowledge of the channel condition, geographic position, and speed of the car to improve message dissemination in vehicular ad-hoc networks. However, in this protocol, they employ the use of Broadcast Request to Send and Broadcast Clear To Send techniques which may add to the time taken to setup a connection and perform transmissions. This protocol may fail if used in a more rapidly evolving mobility scenario which the authors did not test. Some researchers (Burns et al, Burgess et al, Y Li and E Peytchev) have suggested the use of vehicles such as buses with more permanent routine as agents of message dissemination with the argument that buses are fairly frequent and therefore a reliable way to share data in vehicular network. This works very well for areas with a frequent bus network. Trajectory based forwarding (Tian et al, Niculescu and B. Nath) show that by understanding the road through which the vehicles travel, delivery of messages can be improved upon. Here messages are forwarded greedily along a pre-defined trajectory.

III FRAMEWORK DESCRIPTION

Framework for Improved Message Delivery (FIMDEV) is a dynamic framework that works by comparing current road traffic information against stored values of that road under similar road traffic conditions. In the proposed framework, the vehicles are supplied with information about the current road traffic data by using which they make decisions. Research on how to gather such data have been conducted by (Gamati et al. 2011). In this research, however, we have assumed that current road traffic information is provided by the city's transport authorities which is collected from inductor loops along various points of the road. The figure below shows the structure of the framework.

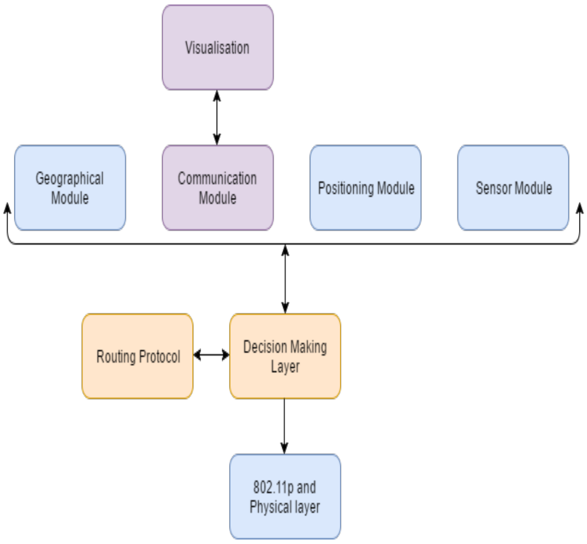


Figure 2 Proposed Framework

The framework consists of various modules representing different data sources such as geographic, communication, position and sensor data. While geographic data can be retrieved from digital maps, positioning will be done by using information from any global navigation satellite system available to the vehicle. Gathering of sensor data has been discussed previously (Gamatti et al. 2011).

Sensor Module

Data from speed sensors for example will play a vital role in ensuring accuracy of information (Tripp-Barba et al., 2014). For this research project, the sensor data is important in order to maintain validation or corroborate the traffic conditions reported by other vehicles; sensor data at this stage are simulated information such as the speed of the vehicle.

Positioning Module

The position of the vehicles along the route is a very important factor as it allows algorithms to differentiate between reachable and unreachable nodes. As will be seen below, if a path is to be determined between two points then those points' locations must be identified in relation to the area under consideration. In (Sun et al., 2000) the authors proposed and implemented a protocol called TRADE which uses the information obtained from Global Positioning System in order to categorise the neighbouring vehicles, by doing this, the protocol is able to accurately detect the furthest reachable node within the neighbouring list. Though the use of positioning is widely accepted, the use of Global Positioning System as the default choice is a bit of a stretch because of the level of accuracy available for use in the public domain. In terms of experimental simulation, pure dependence on a god-like accuracy of positions is faulty, as this will ultimately prove impossible to achieve in real world scenarios. In this research project, the proposal is that the most suitable Global Navigation Satellite System to use will be the Galileo system that offers higher accuracy to within a meter (Steigenberger, 2017).

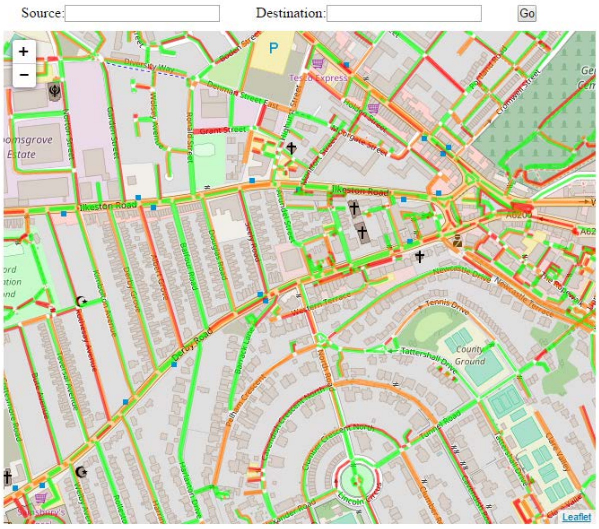


Figure 3: Visualisation of the communication map

Communication Module

This module contains information about the wireless network condition of the area in question, it is a wireless network condition map of the area as visualised in figure 3. It is a collection of data obtained by running experiments offline under various conditions and recording the data. In order to achieve the aim of this research, it is crucial to determine the wireless communication condition of the network of each road with various combinations of traffic elements. That is, capturing the level of wireless network condition when the state of vehicles in this scenario are altered by changing the number of participants and their average speeds. These data can then be used as a reference when comparing against current wireless network communication levels based on current traffic data. These factors have been chosen after careful consideration of the factors that have considerable effect on the potential outcome of any VANET communication.

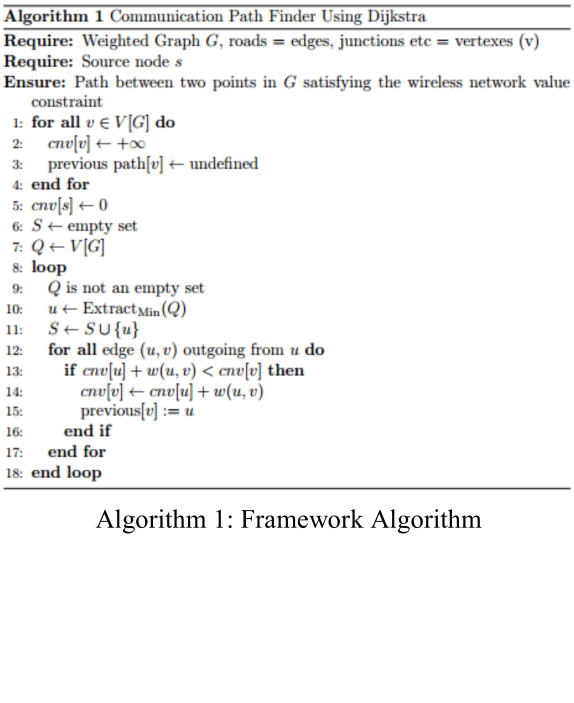
The experiments conducted consists of a series of simulations on each of the City's roads under consideration, the metrics considered as outcomes include the packet delivery ratio and average end to end delay. This method of analysing and storing values is inspired by application of similar methods used in SatNavs where each road and landmark is digitised for future use. If each road's wireless network situation is known beforehand then it allows for some adjustment to how messages are moved about those roads.

Packet delivery ratio is the ratio of the number of packets received by the destination to the number of packets sent, it is a uniquely important factor as it represents how efficient a wireless network is; high values indicates a good network (Shobana M., Karthik S., 2013).

In order to classify each road as good average or bad, we varied number of vehicles and speeds for each road simulated and recorded the results of those experiments. Next the results are identified as:
Good, where packet delivery ratio is ≥ .75 and ≤ 1
Average, where packet delivery ratio is ≥ 0.5 and ≤ 0.74
Poor, where packet delivery ratio is ≥ 0 and ≤ 0.49

Together these values are called communication network values (cnv) which are then stored for each road against its corresponding number of vehicles and speed in a database

Geographical Module
Several researches have shown digital geographic map of areas to be useful tools in implementing vehicular network designs. The map of an area can provide road information critical in directing messages along the right path, or used in implementing conditional routing at junctions etc. For example, the authors in (Xiang et al., 2013) proposed a map based protocol called GeoSVR aimed at solving local maximum and sparse connectivity problems in VANETs thereby increasing packet delivery ratio. It works by locating the nodes on a digital map and calculating the optimum path by using a shortest path greedy algorithm. Similarly, for FIMDEV to function, sections of the area under consideration will be represented as a weighted graph, with junctions, intersections and roundabouts represented as vertices (nodes) in the graph while the roads represent the edges of the graph. These edges will all have weights which represent the current wireless communication condition values of each road which will be compared to stored values for each road as described in the previous section. Paths between two locations can then be found using Dijkstra's algorithm as follows;



```
Algorithm 1 Communication Path Finder Using Dijkstra
Require: Weighted Graph $G$, roads = edges, junctions etc = vertexes (v)
Require: Source node $s$
Ensure: Path between two points in $G$ satisfying the wireless network value constraint
for all $v \in V[G]$ do
    $cnv[v] \leftarrow +\infty$
    previous path$[v] \leftarrow$ undefined
end for
$cnv[s] \leftarrow 0$
$S \leftarrow$ empty set
$Q \leftarrow V[G]$
loop
    $Q$ is not an empty set
    $u \leftarrow \text{Extract}_{\text{Min}}(Q)$
    $S \leftarrow S \cup \{u\}$
    for all edge $(u, v)$ outgoing from $u$ do
        if $cnv[u] + w(u, v) < cnv[v]$ then
            $cnv[v] \leftarrow cnv[u] + w(u, v)$
            previous$[v] := u$
        end if
    end for
end loop
```

Algorithm 1: Framework Algorithm

i. All nodes are initially set to infinity with the exception of the starting position which is given a value of zero. That is, the communication network values between the starting point and every other point is regarded as infinity.
ii. All nodes are regarded as temporary with the exception of the starting node, in order to indicate what nodes have been "visited".
iii. The starting node begins the process by being marked as active.
iv. Calculation of the ability to reach all neighbour nodes from the active node by summing up its value with the weights of the edges.
v. If such a calculated path of a node is smaller than the current one, update the value and set the current node as the previous node.
vi. Next the node with the minimal temporary value is marked as active.
vii. Steps 4 to 6 are repeated till all nodes examined.

At the end of the algorithm, a path that has the best communication network ability between the source and intended destination is found and this information is forwarded to the decision layer in order to forward the intended message.

Decision Layer
The decision making is one based on trajectory forwarding (Niculescu and Nath, 2003) which is a method that directs messages closer to its destination by selecting a node in a specific direction. It is a forwarding strategy based on iterations of the algorithm on each node while considering each node as the source node until the message arrives its destination. Trajectory Based Forwarding is favourable in this case because it uses similar data already found in FIMDEV, hence it is easier to implement.

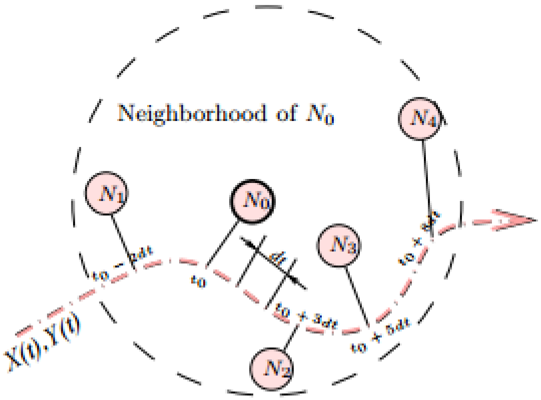


Figure 4: Trajectory Forwarding Strategy (Niculescu and Nath, 2003)

In trajectory based forwarding messages follow a trajectory established at the source, but each successive node takes a greedy decision to infer the next hop based on local position information from a Global Navigation Satellite System such as GPS. If the trajectory is expressed as a coordinate X(t), Y (t) then the equation for routing along a line with slope α passing through the source with coordinates x_1, y_1 would be described by $X(t) = x_1 + t \cos(\alpha)$; $Y(t) = y_1 + t \sin(\alpha)$ respectively. Where α, x_1, y_1 are constants, and the t describes euclidean distance traveled along the line. More information on this can be found in (Finn, 1987) (Niculescu and Nath, 2003).

IV RESULTS AND EVALUATION

The proposed framework was tested on Network Simulator, NS-3 and evaluated in comparison to Ad-hoc On Demand Vector Routing (AODV) protocol (Perkins et al. 2003). All experiments were done following IEEE 802.11p standard. The scenario tested was an Open Street Map (OSM) extract of the Nottingham City Centre that was prepared using mobility patterns from Simulation of Urban Mobility (SUMO). The results shown here for each evaluation represents a mean of 20 executions.

Parameter	**Value**
Simulation Scenario	Nottingham City Center
Frequency Band	5.9 GHz
MAC Protocol	802.11p
Node Density	Varied
Node Speed	Varied
Interface Type	Queue
Range	250m
Propagation Model	Two-Ray Ground

Table 1: Simulation Parameters

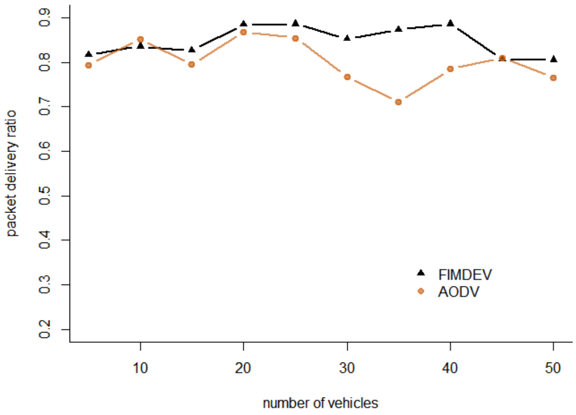


Figure 4 PDR results at speed 10m/h

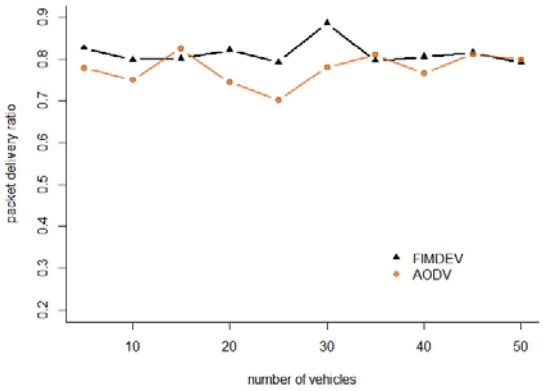


Figure 5 PDR results at speed 20m/h

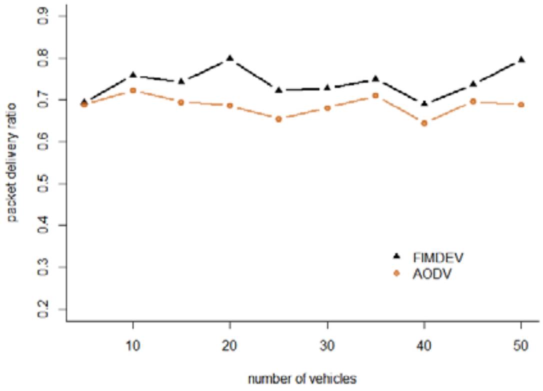


Figure 6 PDR results at speed 30m/h

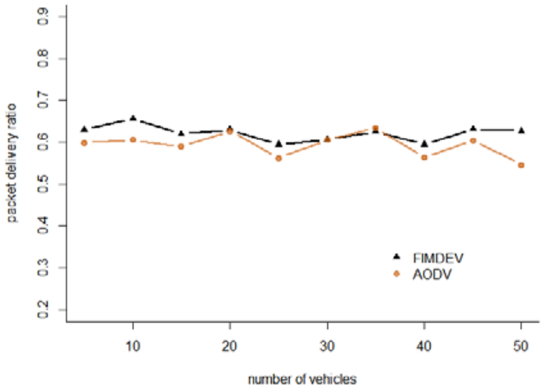


Figure 7 PDR results at speed 40m/h

The graphs show FIMDEV provides improvement in terms of the packet delivery ratio at various speeds and number of vehicles. The speeds are representative of the average speeds within the Nottingham city center. The strength of FIMDEV is its ability of nodes to quickly utilize current road traffic condition as a yardstick to determine what its premeasured wireless communication condition is and therefore find alternate routes where possible. Nodes wanting to transmit messages compare the current traffic condition (i.e. number of vehicles on the route and speed) against the corresponding values in the stored database. Two areas where FIMDEV performs better than AODV is when the vehicles travel at high speeds and situations where there are a lot of vehicles in the network as previous experiments have shown that high speeds and high vehicle congestion inhibits efficient message dissemination (Hafeez et al. 2013).

V CONCLUSION

The goal of the framework is to provide a system through which vehicles can make more intelligent decisions about how to transmit messages in a vehicular ad-hoc network. It can be noted from the evaluation of the work that the model compares favourably against pure flooding and AODV. The explanation for this is that the vehicles should not only apply suppression for reducing multiple broadcasts when the network is perceived to be congested through a comparison of the current traffic condition and recorded corresponding wireless network values in the communication database. Vehicles can therefore use selective path forwarding to push messages towards less congested routes by using a modified greedy path algorithm while applying directed forwarding strategy to ensure messages follow the specified path. The aim of the research can be said to have been attained, being to describe a novel way of using road traffic information to make intelligent decisions related to message distribution in a vehicular ad-hoc network.

Storing values in a database is acceptable for small areas but it is not entirely scalable, that is, as the size of the network increases the amount of data to search through increases.

Hence in future work, we will consider using strategies like Farthest-in-Future Optimal Caching to store the communication data to reduce the amount of obsolete data the algorithm needs to go through.

REFERENCES

Bi Y., L.X. Cai, X. Shen, and H. Zhao. 2010. “A cross-layer broadcast protocol for multi-hop emergency message dissemination in inter-vehicle communication,” in Proceedings of the IEEE International Conference on Communications (ICC ’10), pp.1–5, Cape Town, South Africa

Burns B., O. Brock, and B. Levine. 2005. “MV Routing and Capacity Building in Disruption Tolerant Networks,” in InfoCom.

Burgess J., B. Gallagher, D. Jensen, and B. N. Levine. 2006. “Maxprop: Routing for vehicle-based disruption-tolerant networking,” in InfoCom.

Crovella, L. Feeney, D. Rubenstein, and Raghavan S. 2010. Eds., vol. 6091 of Lecture Notes in Computer Science, pp. 265–276, Springer, Berlin, Germany.

Finn G. 1987. “Routing and addressing problems in large metropolitan-scale internetworks.” in Technical Report ISI Research Report ISI/RR-87-180, University of Southern California.

Fogue M., Garrido P., Martinez, F.J., Cano J., Calafate, C. T. and Manzoni, P. 2012. “Evaluating the impact of a novel message dissemination scheme for vehicular networks using real maps”, Transportation Research Part C: Emerging Technologies, vol. 25, pp. 61–80.

Fogue M., Garrido P., Martinez, F.J., Cano J., Calafate, C. T. and Manzoni. 2010. “Evaluating the impact of a novel warning message dissemination scheme for VANETs using real city maps” in NETWORKING 2010: 9th International IFIPTC 6 Networking Conference.

Hafeez, K.A., Zhao, L., Ma, B. and Mark, J.W., 2013. "Performance analysis and enhancement of the DSRC for VANET's safety applications." in IEEE Transactions on Vehicular Technology, 62(7), pp.3069-3083.

Morsink, P., Cseh, C., Gietelink, O. and Miglietta, M. 2002. “Design of an application for communication-based longitudinal control in the CarTALK project.” in IT Solutions for Safety and Security in Intelligent Transport (e-Safety).

Niculescu, D. and Nath, B. 2003. “Trajectory Based Forwarding and Its Applications.” in MobiCom'03.

Perkins, C., Belding-Royer, E., Das, S. et al. 2003. “Rfc 3561-ad hoc on-demand distance vector (aodv) routing”, Internet RFCs pp. 1–38.

Shobana M. and Karthik S. 2013. “A Performance Analysis and Comparison of various Routing Protocols in MANET” in proc. International Conference on Pattern Recognition, Informatics and Mobile Engineering (PRIME), pp 391-392.

Sommer C., Tonguz O.K., and Dressler F. 2011. “Traffic information systems: efficient message dissemination via adaptive beaconing” in IEEE Communications Magazine, vol. 49, no. 5, pp. 173–179.

Steigenberger, Peter. 2017. “Accuracy of Current and Future Satellite Navigation Systems”. Habilitation, Technische Universität München.

Sun, M.T., Feng, W.C., Lai, T.H., Yamada, K., Okada, H. and Fujimura, K. 2000. “GPS-based message broadcast for adaptive inter-vehicle communications” In Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd (Vol. 6, pp. 2685-2692).

Tian J., Han L., Rothermel K., and Cseh C. 2003. “Spatially Aware Packet Routing for Mobile Ad Hoc Inter-Vehicle Radio Networks,” in Proc. of the IEEE 6th Intl. Conf. on Intelligent Transportation Systems (ITSC).

Tripp-Barba, C., Urquiza-Aguiar, L., Igartua, M.A., Rebollo-Monedero, D., de la Cruz Llopis, L.J., Mezher, A.M. and Aguilar-Calderón, J.A., 2014. “A multimetric, map-aware routing protocol for VANETs in urban areas” in Sensors, 14(2), pp.2199-2224.

Wischhof, L., Ebner, A., Rohling, H., Lott, M. and Hafmann, R., 2003. "Adaptive Broadcast for Travel and Traffic Information Distribution Based on Inter-Vehicle Communication" in IEEE IV'2003.

Yong Xiang, Zheng Liu, Ruilin Liu, Weizhen Sun, Wei Wang. 2013. "GeoSVR: A map-based stateless VANET routing" in Ad Hoc Networks, Volume 11, Issue 7, Pages 2125-2135, ISSN 1570-8705.

Ziliaskopoulos, A.K. and Zhang, J. 2003. A Zero Public Infrastructure Vehicle Based Traffic Information System. in TRB 2003 Annual Meeting.

AUTHOR BIOGRAPHIES

DR. EVTIM PEYTCHEV is Reader in Wireless, Mobile and Pervasive Computing in the school of Science and Technology at Nottingham Trent University, UK. He is leading the Intelligent Simulation, Modelling and Networking Research Group, which consists of 5 lecturers, 3 Research Fellows and 6 research students. He is the Module Leader for Systems Software; and Wireless and Mobile Communications. He is a participant in various European projects such as MODUM and he also teaches modules Software Design and Implementation; Mobile Networking; Enterprise Computing; and Computer Architecture at the university.

Dr. JAVAD AKHLAGHINIA is an experienced research scientist in intelligent transportation systems, embedded technologies, machine learning and intelligent environments. He has been working as a research fellow in EC funded European projects such as REMOURBAN at Nottingham Trent University, MODUM and ARUM at the University of Manchester.

NNAMDI ANYAMELUHOR is a post graduate researcher in the field of vehicular networks and wireless communications at the Nottingham Trent University with a first degree in Electrical and Electronic Engineering, MSc in Cybernetics and Communication Engineering from the Nottingham Trent University, United Kingdom.

SUPPORTING PENSION PRE-CALCULATION WITH DYNAMIC MICROSIMULATION TECHNOLOGIES

Dávid Burka and László Mohácsi
Department of Computer Sciences
Corvinus University of Budapest
1093, Budapest, Fővám square 13-15
e-mail: david.burka@uni-corvinus.hu

József Csicsman and Benjámin Soós
Faculty of Sciences and Informatics
Department of Computer Optimalisation
University of Szeged
6720, Szeged, Árpád square 2.
e-mail: csicsman@calculus.hu

KEYWORDS

Microsimulation, Demography, Pension system, Ageing population

ABSTRACT

Population ageing induces many challenges in the pension system of developed countries. It is necessary to support the decision-making processes regarding these challenges by forecasting different future scenarios. Long-term forecasts are required to understand the development process of the population and the pension system. The microsimulation approach has many benefits over other forecasting methods, though it requires high level of programing skills and significant computing capacity. Moreover, a long-term demographic microsimulation must be dynamic and it should preferably also include the relations between individuals. In this paper, we will introduce two different microsimulation based solutions for the above-mentioned forecasting tasks. The first one is a complex model – aiming to forecast the Hungarian population – built in SAS, that can highlight the advantages of the microsimulation approach. The second solution is a Simulation Framework (written in C#), that aims to drastically reduce the difficulties regarding microsimulation using the findings of the SAS model. Our goal is to introduce our systems in the hope of future collaboration with economists and demographers.

INTRODUCTION

In the last decade, all OECD countries induced some kind of pension reform (OECD, 2015). This is a good indicator of how urgent it is to manage the challenges regarding the changing structure of the population. Pension systems can be categorized whether they are based on funded or unfunded plans. Some hybrid systems exist, but even those usually prioritize one of the plans.

A funded or pre-funded plan means that the contributions are invested in a fund to cover the benefits of the individual after the retirement. Unfunded – or so called pay-as-you-go – plans on the other hand cover the pension costs of the retired with the contributions of the active population. The two approaches must face different kinds of challenges.

In a funded system, the adequate living conditions of the elderly are not properly secured. People, who outlive their life expectancy will exhaust their fund even if it had high yield, and they might end up in poverty. In an unfunded system if the total contributions do not reach the amount of the total pension costs, the government must supplement the shortage from taxes. In recent years, the main reason for the shortage is the ageing of the population and the shift of the active worker per pensioner ratio. The pay-as-you-go systems can care for their elderly through the whole retirement since the funds will not run out. However, the monthly benefits are fixed at the time of retirement, and if high life expectancy is paired with high inflation, the pension payments might not be able to secure appropriate living conditions.

The adequacy and the financial sustainability of a pension system are conflicting interests, and both plans must aim to balance these interests. Funded plans are more sensitive regarding adequacy while unfunded plans often have to face the problems of financial sustainability. The systems introduced in this paper aim to support forecasts for both types, however our own models are built for Hungarian examples, thus we will mainly focus on the financial sustainability of pay-as-you-go systems.

AGEING POPULATION

The main reason behind the uncertainty regarding the unfunded pension systems is the ageing of the population. The life expectancies continuously increase while fertility rates decrease, thus the ratio of the active population and the pensioners worsens every year in most of the developed countries. Since unfunded plans cover their expenses via the contributions of the labor market, the shift of the ratio creates an ever-increasing burden on the society. There are three main factors that influence the structure of the population: the increasing life expectancies, the decreasing fertility rates, and the international migration.

The technological advancement has improved medicine and granted acceptable living conditions for many

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

people, thus increasing the overall life expectancies at birth. This resulted in the increase of the worldwide ratio of the elderly (60+) population from 9.9 percent to 12.3 percent between 2000 and 2015 (United Nations, 2015).

From an evolutionary point of view the more children the better the survival chances of the gene pool. However, as the society prospers, the fertility rates tend to decrease. The social survival of the children becomes a more realistic concern than the biological survival. The death of a child is unlikely in a modern society so parents tend to focus on other worries regarding their offspring. Good studies, acceptable career path, and even finding a partner are important to become an honored member of the society. Ensuring this requires resources – like time and money – per child, so a smaller number of children is preferred, thus the fertility rates decrease.

International migration can have a significant effect on the structure of the population. For example, in Central- and Eastern European countries – where the pension plans are typically unfunded – the migration of the young workforce to the economically more prosperous regions enhances the negative effects of population ageing (Cerami, 2011).

These global trends are not likely to change soon, so it is important to prepare for the effects of the ongoing processes in the structure of the population. The changes will affect the society just as much as the economy (Kulik, 2014), thus it will be impossible to simply forecast every single effect that might influence the financial sustainability of the pension system. Rather, it would be more beneficial to model the cause and effect of as many scenarios as possible. However, this approach requires models that go beyond the simulation of death and birth, and include the relations between individuals.

RELATIONS

The relations and overall circumstances of an individual have a strong influence on his or her life path, and these effects can be confirmed through the analysis of statistical data. Married people, who lived together for a long time, often die in a short interval after each other, and married men tend to live longer than bachelors. The education of a child is highly correlated to the education level of the parents. Moreover, the income of the whole household has a more significant influence on ones living conditions than the individual income.

Examining and implementing the above-mentioned connections in a model can highlight such complex, wide-branching effects in economic indicators, that would be impossible to notice through simple trend-based forecasting methods. This way the number of estimated parameters will be lower, and – if we assume that the implemented connections are correct – the error of the forecasted indicators can also be reduced. However, such a detailed model requires far more resources to realize than traditional forecasting solutions.

FORECASTING METHODOLOGY

Demographic forecasting methods can be categorized on a macro-micro scale based on how detailed they are. The simplest forecasting methods – like fitting a trend or regression – can estimate the future development of indicators based on historical data. The cohort-component method groups the population based on the properties of the individual, and does not differentiate between the individuals. These properties are usually age and gender, but more complex models can include for example education, ethnicity, or residency. The groups are represented with the number of individuals in them, and the group numbers are changed at every iteration step based on a transition probability matrix. Agent based models extend this functionality by using multiple smaller transition matrices and different rule sets – if state-transition is not appropriate – to iterate the properties of agents. The agents can represent groups or even individuals, thus microsimulation is the most detailed, special version of the agent based models. It follows the life path of every single individual throughout the whole simulation, and changes their properties in every iteration step according to transition probabilities, rules, or interactions with other individuals.

Microsimulation – being one of the extremes – has many advantages over other methods. It allows the implementation of complex logics like the family relationships between individuals and it results in such a detailed estimated dataset, that would otherwise only be obtainable by repeated data collections. Using rules and fragmented transition matrices, a microsimulation model can be easily extended. However, as we move on the macro-micro scale toward the micro side, the models become increasingly more complex. A simple regression can be done in most of the statistical software packages (i.e.: R, SPSS) and a cohort-component model can be realized in a spreadsheet, but implementing a microsimulation model requires high level of software development skills and significant computing capacity.

These difficulties are the reason, why the microsimulation methodology has not become as wide-spread as the cohort-component method in the field of demographic forecasting, even though it was introduced more than half a century ago (Orcutt, 1957). Microsimulation started to appear more frequently in the last few decades thanks to the advancement in technology (Merz, 1991). However, most of the implementations are ad hoc solutions since even if the computing capacity is given, the microsimulation based systems still require programming skills and many work hours. This resulted in the implementation and publication of many systems that are difficult to reuse and understand by other researchers. It is usually more beneficial to develop a new solution for every new problem, especially if the target countries differ, since it would take too much effort to utilize someone else's work.

Our main goal is to address this problem. We created two different solutions, that both aim to show the capabilities and benefits of a microsimulation based demographic forecasting model, while being modular and easy to use for anyone interested.

DATA STRUCTURE

The Hungarian Central Statistical Office (HCSO) uses a microsimulation based forecasting system since the 90ies (Csicsman, 2012). We also focused on building models to forecast the Hungarian population and improve the existing solutions, thus we used datasets provided mainly by the HCSO. We classified our data sources into two different categories: attributes (or properties) and parameters.

Attributes are data values representing the individuals. These values get changed throughout the iteration steps according to the parameters or the attributes of others. The initial attributes should be based on a wide spread data collection like a census or a micro census. Often there exists no appropriately detailed dataset, so multiple datasets must be merged through statistical matching. We based our models on the results of the Hungarian micro census of 2005.

Every transition matrix or rule in our system is based on historical data. These parameters are the core of the model, and allow the interaction between the individuals. Some parameters barely change with time. For example, the probability of a newborn being male has been relatively constant for the past half century. However, most of the time the parameters change throughout the years. The life expectancies increase while the fertility rates decline. All time varying parameters must be forecasted before the start of the simulation. Fitting a trend or creating a regression to estimate the future values of these parameters can be adequate, but sometimes the parameters depend on multiple factors. For example, in most models, the mortality rates vary based on the age and gender of an individual, thus the estimation requires the forecasting of a whole matrix. A good practice is to use singular vector decomposition to obtain a time varying vector component, that can be forecasted by traditional methods (Burka, 2016).

We used similar, automated techniques to forecast the parameters in our systems. In the following chapters, we will describe both our SAS based model and our Simulation Framework in detail, to give an insight on what solutions we used and how these systems can be beneficial for other researchers, economists, or demographers.

SAS BASED MODEL

The motivation behind this project was to apply dynamic microsimulation methods on real life problems. We decided to implement our thoughts and ideas into a fast and flexible SAS program. SAS (Statistical Analysis Software) is a software suite developed by SAS Institute, that contains many advanced statistical features (e.g.: data management, predictive analysis). All the tools needed for mining important information from raw data - like data quality management or advanced analytics – are built in the software by default. Most of the microsimulation models that allow a wide range of configuration options for the user are built on their own, unique language. The Hungarian MIDAS_HU (Dekkers, 2015) or the Slovenian LIPRO (Majcen, 2011) programming tools are great examples. One of the greatest benefits of SAS is that it is a widely used statistical software, thus a great portion of those who are interested in demographic forecasts are already familiar with its syntax. In this section, we would like to introduce what we have accomplished in the SAS approach so far, and which direction we are going to focus on in the future.

Beyond the essential information such as personal and household data, there are lots of additional details about a given household's shopping habits, financial position, or school records of the individuals, which can be used in a model. However, we only included the most important attributes in order to create a general-purpose system. This means, that the software provides basic functions of demographic modules, such as death, birth, marriage, divorce, and besides these the user must include additional properties about studied topic. It is important to note, that the development of a new module requires the knowledge of the SAS Base and SAS Macro language, and of course the configuration of the base program.

Our system's most important feature is modularity. This means, that all the basic and user written modules can be added or removed by the user. An appropriate use of this feature can improve performance, and make it possible to split the research into smaller segments. We used the built-in procedures of SAS to forecast the parameters of the modules based on historical data between 1995 and 2015. The Proc Forecast toolbox can be used for exponential smoothing, applying Winters method, or thanks to the modularity of the system, we can fit the needs for different forecasting solutions, depending on the current task.

Thus far, we have mentioned fictional population that can be manipulated by demographic modules over time. Every year is an iteration step, and within every step the modules change the properties of the individuals. However, there are certain tasks that cannot be managed on the entity level. These tasks are going to be handled by sampling procedures, that take a set of individuals out from the base population, execute the given action on the selected group, then puts the manipulated portion of data back into the original dataset. The matching algorithm used for managing marriages can be a good example for this methodology. We select a portion of women grouped by age intervals, and we pair them with an equal portion of men with similar attributes (relevant matching conditions can be set to strict or loose). Finally, we remove the couples from their previous households, and assign a new household ID for them, thus creating a brand-new household for the married couple.

There are two different approaches to sampling: Unrestricted random sampling (URS) allows replacement in the selection, while Simple random sampling (SRS) is a selection with equal probability and does not allow replacement. In case of marriages URS would mean, that as the algorithm loops through the potential partners, one individual could get assigned to multiple partners. In this case, a breaking rule could decide who the individual gets partnered with. However, this would ruin the equal probability of distribution. The SRS approach solves this problem by removing the already selected individuals from the list of potential partners, but the additional queries will increase the runtime significantly.

During the simulation of a given demographic event, we can define conditions for every scenario that should be handled. For example, the death of different members in a household should be treated with a different approach, depending on the individual's role in the family. We can manage these set of rules like as a module. It is easy to modify, attach or detach them, thus we can set different levels of elaboration. However, this can influence runtime significantly.

SIMULATION FRAMEWORK

One of the main reasons, why microsimulation based models did not became wide-spread is, that high level of software development skills is necessary to implement them. Modular solutions, like our SAS based model could help researchers with basic programming skills and/or SAS knowledge to implement their own model. Our goal was to further reduce the necessary skill level while also managing the other disadvantages of a microsimulation based model. To achieve this, we started the development of a Simulation Framework based on three major requirements: speed, flexibility, and ease of use.

A demographic microsimulation requires a lot of computing capacity. For example, in the case of Hungary, the population consists of approximately 10 million entities. The aim is to create long term forecasts of 50 years, and multiple attributes must be recalculated at every iteration step. Considering, that managing the relationships further increases the complexity of the algorithm, the runtime of a simulation can easily reach multiple hours. Parallelizing of the simulation can reduce the runtimes significantly (Mohácsi, 2014). It is important to note though, that a parallel code can be far more complex than a single core solution, thus its realization requires high level of software development skills.

A fast framework, without a wide range of configuration options would quickly lose its purpose. Our goal is to reduce the limitations of the system as much as possible, and allow the users to create any model, that fits their needs. The framework should be used to compare multiple – even hundreds of – future scenarios.

Most published microsimulation based models can be accessed, and with enough time and development skills they can be modified for one's own needs. However, the resources spent on learning the system are taken away from understanding and implementing the model itself. It is important for a development environment to be intuitive. The user should be able to learn the handling of the available tools in a short time, thus reducing the resource costs of incorporating the results of others into the user's own model.

It is clear, that the above-mentioned properties are conflicting requirements. The optimization of the code makes it hard to understand, and the more configuration options are available, the harder it becomes to speed up the simulation. Our main goal was to find an appropriate balance between speed, flexibility, and ease of use, while developing our Simulation Framework. We had to keep a scripting feature, otherwise it would have been impossible to avoid limitations, thus we decided to implement a layer on top of the main code. The relatively simpler algorithms of the simulation steps – that implement the state transitions and rules described before – can be built from simple blocks in an intuitive graphical user interface. The optimized, parallel algorithms, that control the whole simulation, are hidden behind the scenes, and should not be accessed by the user. The blocks represent simple functions, variables, or matrices, but often have relatively complex algorithms behind them.

The configuration of the framework happens with Excel sheets that follow some simple rules, and forms that aid with the final settings based on the content of the sheets. The dynamic code is built in two layers. First the blocks available in the graphical programming interface – that allow the configuration and implementation of the model – are created based on the initially imported datasets. The second layer uses the structure built in this interface to generate the code for the simulation and the necessary queries to only save the indicator values of interest instead of the enormous dataset. Finally, the finished code runs on all the available processor cores. This approach significantly decreases the runtime since it saves time on directly writing down the different dynamic variables instead of searching for them in queries.

The main algorithm is relatively straight forward except of the thread safe random generator. This object generates a random seed for every single individual at the start of an iteration, and the individuals use these seeds in their own random generator to control their "decisions". The extra random number generators slightly worsen the performance, but in exchange we can ensure that the random number generation is deterministic, thus we can reproduce our results any time.

The relationships module is the most restricted part of our solution. It operates on the household level, since data is usually available for households instead of relationships. So, the individuals are divided into households at the start of the simulation. The structure of a household can change four ways: a new member can be born, a member can die, a member can leave or a member can join. Birth and death is handled by the default modules. If an individual leaves the household we automatically create a new single member household for him or her, and if two individuals decide to get into a relationship we move them into a brand-new household instead of joining their households together. This approach allows us to handle these two events in a single interface – like the simulation step interface – since in both cases the individuals join a new household.

If an individual leaves a household or gets into a relationship, they get a tag that represents the reason they joined the household for. Leaving is relatively simple, the individual needs a tag for the reason of leaving (i.e.: divorce, growing up or leaving family home), but getting into a relationship is more complicated. We wanted to allow any kind of relationship in our model, so the user can set up relationship types in the settings menu. A type (i.e.: marriage, life companion, flat mate) can have restrictions towards the available partners. For example, in our models we only allow women to choose partners and we do not allow same sex marriage so we exclude women completely from the available partners. The saying that "opposites attract" is statistically not true, most people choose partners with similar properties to themselves. (Of course, this only includes properties that can appear in a database.) In our framework, the user can select properties that are relevant in the choice of a partner regarding the given relationship type, and in the simulation the available partners will be divided into an *n* dimensional matrix of groups where *n* is the number of selected relevant properties. An individual will choose a partner based on a probability distribution that defines the weighted Euclidean distance in the *n* dimensional space between the individual and the future partner. The distribution function and the weights are set by the user. Usually a distribution will prefer a partner who has similar properties, namely one who is closer to the individual. Unlike the main simulation step, the relationship algorithm is not parallelized as of now, thus it increases runtimes significantly.

ASSESSMENT

In this section, we would like to provide some insight on how our solutions perform. However, the comparison of two different solutions can only be effective if both implemented the same base model. This means that comparison requires the exact recreation of a model. In most cases papers in the topic do not supply access to the dataset, since it is not relevant to recalculate the results, thus the models cannot be recreated for comparison in terms of runtime.

For this reason, we can only effectively compare our own two solutions. It is important to note though, that our goal is to introduce our solutions and the detailed description of the models is outside the scope of this paper. Moreover, we intend to use the two solutions to implement fundamentally different models that are still in the development phase. Thus, we implemented the same simplified model with both of our solutions.

We forecasted the population from 2004 to 2054 and compared the results to the summarized population numbers between 2005 and 2015, that are available in the public online database of the HCSO. The initial population is based on the micro census of 2005. The census is made in the middle of the year, thus death and birth numbers of 2005 would be incomplete, so we start the simulations from 2004 with the appropriate part of the data. We used the Lee-Carter method (Lee, 1992) to forecast the mortality rates and we modified the method for the fertility rates. Both parameter forecasts were based on the historic values between 1995 and 2004. Figure 1 shows the results of these forecasts. The SAS model underestimated the number of deaths and the Simulation Framework overestimated the small increase in the birthrates. However, both solutions managed to forecast the size of the population for 10 years with a maximum error of less than 1.5 percent. The difference between the result of the two theoretically identical models appears because of the different random generation of our solutions. This further supports our claim, that a fast forecasting is necessary to allow the comparison of multiple random seeds and configurations in a short time.

A pure birth and death model would not properly introduce the possibilities of our solutions, thus we included other modules in our tests. We implemented a simple pension and income module based on the model of Péter Mihályi (Mihályi, 2016). The incomes are fixed as the overall weighted average gross income in the micro census of 2005. The age when some individual starts working is 21 years and the retirement age limit is 65 years. Based on the original model the overall percentage of contributions from the gross income is 20 percent. Assuming, that an individual is retired for half as long as he or she has worked, and that the system works well – so on average everyone gets the same amount of pension till their death back as they contributed – the amount of pension is set as 40 percent of the gross income. We also included a marriage and divorce module. According to the statistical data on average 1 percent of marriages end with a divorce every year and 2 percent of singles marry annually. We chose to implement the mate choice so that only women can choose partners, and their choice is only dependent on age, so with the highest probability they choose a man of the same age and as the age difference grows the probabilities decrease exponentially. With these additions, we aimed to simulate a more complex model to present runtimes, but to simplify implementation we excluded the influence of these properties on the mortality and fertility rates, thus the additions did not change the results seen in figure 1.

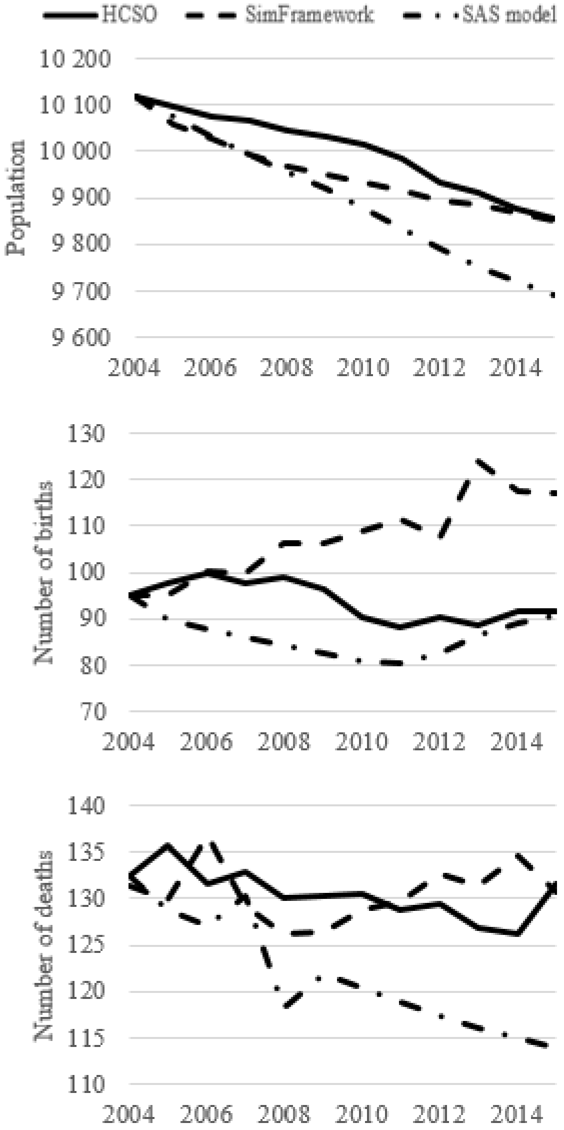


Figure 1: Simulation Results (in thousands) Compared to the Actual Values Between 2004 and 2015.

Table 1 summarizes the runtimes of the above described extended models. We showed the results with and without the relationship module, since it is the most resource heavy part of our solutions. The simulations were tested on a personal computer (Intel Core i7-3632QM CPU @ 2.20 GHz). The table shows the runtimes of the Simulation Framework for both sequential runs and parallelized runs with 4 cores. The speed up is basically linear and the relationship runtimes get simply added, since those are unaffected by parallelization.

Table 1: Runtime Comparisons of 50 Year Simulations

Model	Standard	With Relationships
SAS model	35:27	1:42:14
SimFramework (1 core)	40:18	49:06
SimFramework (4 core)	10:44	19:16

It is clear, that the Simulation Framework outperforms the SAS model, but it is important to note that the latter project is in a less developed stage. The SAS model is also yet to be implemented as a parallel solution to improve runtimes. Moreover, we implemented exponential smoothing to forecast the parameters, thus in the future, we would like to implement more sophisticated algorithms. The Simulation Framework is also in the middle of rework since many UI features (other than the block programming interface) are outdated. However, the most important future task is to improve the parallelization of the household module. We are continuously fixing bugs, improving the documentation and still actively developing the framework. The latest version to date can be found on GitHub [dburka001/SimulationFramework]. Since the SAS model is a business project, its code is not yet published, but can be discussed with developers via e-mail.

CONCLUSION

The continuous ageing of the society results in structural changes in the population of every major country. The changes create new challenges that the governments must face. Ensuring the adequacy and financial sustainability of the pension system is among the most critical challenges. The forecasting of the population is necessary to aid the decision-making process regarding the sustainability of the pension system. However, the development of demographic processes is slow, so only long-term forecasts can highlight the ongoing changes.

We showed, that the most appropriate method to support the decision-making process regarding the issues we raised is a complex dynamic microsimulation solution. We discussed the advantages of the chosen approach, and also analyzed the difficulties that prevented microsimulation to become a wide spread method for demographic forecasting. We proposed two solutions that aim to offset the disadvantages of microsimulation. We hope that in the future our systems can facilitate the use of dynamic microsimulation techniques in the field of demographic forecasting and allow us to collaborate with other researchers, economists, and demographers.

REFERENCES

Burka, D. 2016. “Supporting the Hungarian Demographic Pre-calculations with Microsimulation Methods”. *SEFBIS Journal*, 10, 13-23.

Cerami, A. 2011. “Ageing and the politics of pension reforms in Central Europe, South-Eastern Europe and the Baltic States”. *International Journal of Social Welfare*, 20(4), 331-343.

Csicsman, J., László, A. 2012. “Microsimulation Service System”. *Hungarian Electronic Journal of Sciences*

Dekkers, G., Desmet, R., Rézmovits, Á., Sundberg, O., Tóth, K. 2015. “On using dynamic microsimulation models to assess the consequences of the AWG projections and hypotheses on pension adequacy: Simulation results for Belgium, Sweden and Hungary”. *Federal Planning Bureau – Central Administration of National Pension Insurance*

Kulik, C. T., Ryan, S., Harper, S., & George, G. 2014. “Aging populations and management”. *Academy of Management Journal*, 57(4), 929-935.

Lee, R. D., Carter, L. R. 1992. “Modeling and Forecasting U.S. Mortality.” *Journal of the American Statistical Association*, 87(419), 659-671.

Majcen, B., Cok, M., Sambt, J., Kump, N. 2011. “Development of pension microsimulation model”. *Institute for Economic Research, Slovenia*

Merz, J. 1991. “Microsimulation – a survey of principles, developments and applications”. *International Journal of Forecasting*, 7(1), 77-104.

Mihályi, P., Vincze L. 2016. “A „Nők–Férfiak 40” nyugdíjkoncepció pénzügyi következményeinek szemléltetése a felosztó-kirovó rendszerben.” *Gazdaság és Pénzügy*, 3(1), 3–24.

Mohácsi, L. 2014. “Gazdasági alkalmazások párhuzamos architektúrákon = Business Computing and Parallel Architectures” *Doctoral dissertation, Corvinus University of Budapest*.

OECD 2015. “Pensions at a Glance 2015: OECD and G20 indicators”. *OECD Publishing*, Paris.

Orcutt, G. H. 1957. “A new type of socio-economic system”. *The review of economics and statistics*, 116-123.

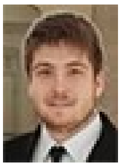
United Nations, Department of Economic and Social Affairs, Population Division 2015. “World Population Ageing 2015“ (ST/ESA/SER.A/390).

AUTHOR BIOGRAPHS

DÁVID BURKA was born in Nagykanizsa, Hungary. He acquired his degree in Business Informatics at the Hungarian Corvinus University of Budapest and is currently working on his PhD in the same field and institute. He obtained his second master’s degree as an Info Bionics Engineer at Pázmány Péter Catholic University of Budapest. His research focuses on simulation and modelling in the fields of demography, social choices, and neurobiology. He currently teaches at Corvinus University and works at the Sleep Oscillation Research Group of the Research Center for Natural Sciences.

JÓZSEF CSICSMAN was born in Budapest, Hungary and attended the University of Szeged, where he studied mathematics and obtained his degree in 1976. He worked for 20 years for the Hungarian Central Statistical Office before founding his private company, Calculus in 1996. From 2001 he was active in Budapest Technical University where he is now leading a Microsimulation research group in the field of simulation for Demographic and Economic changes, and different kind of Statistical Matching applications. He gives lectures on Practical usage of Data Mining and Statistics at BUTE and University of Szeged. His e-mail address is: csicsman@calculus.hu and his webpage can be found at http://calculus.hu.

LÁSZLÓ MOHÁCSI was born in Budapest, Hungary. He attended the University of Óbuda, where he obtained his BSc degree in Electrical Engineering. He acquired a MSc degree in Informatics at the Eötvös Loránd University of Budapest. He obtained another MSc degree in Biomedical Engineering in a collaborative program of the Semmelweis University and Budapest University of Technology and Economics, Hungary. He got his PhD in the ICT Doctoral School of the Corvinus University of Budapest, where he is currently working as a senior lecturer at the Department of Computer Science. His research focuses on the economic application of parallel computing and economic modelling. His e-mail address is: laszlo.mohacsi@uni-corvinus.hu.



BENJÁMIN SOÓS was born in Cegléd, Hungary. He is a student at the University of Szeged, and he is preparing for getting his BSc degree in Business Information Technology. Within the confines of a course about Statistical Softwares, he learned the use of the SAS language, and got acquainted with the tutor, József Csicsman. This relationship led to a thesis work with the title of “Demographic Forecasting with Dynamic Microsimulation”, and to a job at Új Calculus Ltd. As a subcontractor, he currently works at CIB Bank as a Junior SAS Developer. He has the enthusiasm for learning more about dynamic microsimulation techniques and developing a SAS based applications. His e-mail address is: soosbenji@gmail.com.

DATA FUSION IN CLOUD COMPUTING:BIG DATA APPROACH

Piotr Szuster
Cracow University of Technology
Warszawska 24, 31-155 Cracow, Poland
AGH University of Science and Technology
al. Mickiewicza 30, 30-059 Cracow, Poland
E-mail: pszuster@pk.edu.pl

Jose M. Molina
University Carlos III of Madrid
Avda. Gregorio Peces-Barba, 22
28280 Colmenarejo (Madrid) SPAIN

E-mail: molina@ia.uc3m.es

Jesús García-Herrero
University Carlos III of Madrid
Avda. Gregorio Peces-Barba, 22
28280 Colmenarejo (Madrid) SPAIN

E-mail: jgherrer@inf.uc3m.es

Joanna Kołodziej
Cracow University of Technology
Warszawska 24, 31-155 Cracow,
Poland
E-mail: jokolodziej@pk.edu.pl

KEYWORDS
Computational Cloud, Data Fusion, Big Data, Data Integration

ABSTRACT

Data Fusion System should maximize throughput during the complex fusion analytics in order to extract the critical information from a huge amount of data. Therefore, there is an immediate need to leverage efficient and dynamically scalable data processing infrastructure for the analysis of the Big Data streams from multiple sources in a timely and scalable manner. This is necessary for establishing the accurate Situation Awareness (SA) during the real time processes or real time simulations.

INTRODUCTION

Timely acquisition and processing of data from different sources and extraction of accurate information play an important role in many realistic scenarios (such as emergency situations, evacuation systems, crowd management, remote health monitoring, etc.). The growing ubiquity of on-site sensors, social media and mobile devices increases the number of potential sources of outbound traffic, which ultimately results in generation of huge amount of data. This data avalange phenomenon is being described as a new grand challenge in computing: The *Big Data (BD)* problem[1]. BD problems are usually defined as “the practice of collecting complex data sets so large that it becomes difficult to analyse and interpret manually or using on-hand data management applications” (e.g., Microsoft Excel and Relational Database Systems). *Data Fusio*n in solving the BD problems is the major process, where the information is received from the multiples sources. In the case of sensor sources, this problem is known as *sensor fusion*. Sensor fusion problems have been widely analyzed in Air Traffic Control scenarios with multiple sensors and multiple targets. In this domain, many works focused on the specification of the filter structures based on Kalman filter definition of parameters for filter structures [1], definition of management in distributed fusion architectures [2] and the application of Artificial Intelligent Techniques to the fusion problems [3]. Currently, the capabilities of the available data sources, such as sensors, may be relatively big in order to collect big setes of data files. Multi-sensor integration is the essential aspect of modern sensor networks designed forsensor fusion..

Data acquisition systems, such as air surveillance systems, alert systems based on mobile TV-cameras, evolve towards complex information systems, being capable of providing the operator with a great amount of data obtained through a net of spatially distributed heterogeneous sensors. Modern perception systems are composed of several sensors (laser sensors, infrared sensors, acoustic sensors, image sensors, GPS, ambient sensors, etc.). Sensors provide data of each element of the environment in their coverage area. This data is used by the control logic after being merged at the fusion centre. The design of a perception system in an environment have for solving two related problems. The first one is data fusion problem [4], which refers to the optimal way of combining the detections, local tracks and attributes received from the net of sensors before the information could be used by the control logic. The second one is the coordinated sensor-task management [4, 5], in order to optimise each sensor-data acquisition process. Two different types of information can be considered: (i)

[1] http://cra.org/ccc/wp-content/uploads/sites/2/2015/05/bigdatawhitepaper.pdf

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

global or high level knowledge retrieved from the data fusion process; and (ii) local, internal-to-sensor information, such as the knowledge about the current state of sensor load.

The analysis of information received from multiples non-structured sources is another important phase of the general data fusion process. Over 20 million tweets posted during Hurricane Sandy (2012) lead to an instance of the BD problem. The statistics provided by the Pear Analytics[2] study reveal that almost 44% of the Twitter posts are spam and pointless, about 6% are personal or product advertising, while 3.6% are news and 37.6% are conversational posts. During the 2010 Haiti earthquake[3], text messaging via mobile phones and Twitter made headlines as being crucial for disaster response, but only some 100,000 messages were actually processed by government agencies due to lack of automated and scalable data processing infrastructure. Furthermore, many message reporting issues such as medical emergencies and water shortages had unclear or no location information, making response impossible. All those scenarios can be reffered as *Situation Awareness (SA)* - a classical aspect of the Joint Directors Laboratory model for data fusion. SA means that the main actors in the considered situation must be aware of everything, what does and may happen around in order to understand the impact of the neihbourhood on their decisions and actions. SA is strictly related to the Levels 2 and 3 (L2 and L3) of JDL model. *Situation Assessment (L2)* can be interpreted as the capacity to obtain a global view of the environment in order to describe the relationship between the entities tracking in the environment and to infer or describe their joint behavior. *Impact Assessment (L3)* is related to the estimation of the impact of a special situation. Processing and evaluating probability of occurrence of particular situations that are of special relevance are performed.That has to be done because those situations relate to some type of threatening, critical situation, or any other special world state.

Inadequate SA in complex realistic scenarios has been identified as one of the primary factors in human errors with grave consequences such as loss of infrastructureThere is a need for a complete ICT (information and communication technology) paradigm shift in order to support the development and delivery of Big Data applications, in a way that applications do not get overwhelmed by incoming data volume, data rate, data sources, and data types. Dealing with these huge amounts of data requires a new end-to-end data management and analysis paradigm in which methods need to be efficient not just as stove-pipe processes, but as part of a well-integrated system.

Cloud Computing (CC) assembles large networks of virtualised *Information and Communications Technology (ICT)* services such as hardware resources (such as CPU, storage, and network), software resources (such as databases, application servers, and web servers) and applications. In industry, those services are defined as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). CC services are hosted in large data centres (data farms) owned by companies such as Amazon[4], GoGrid[5], Rackspace[6] and Microsoft[7] . Computational clouds, with their special features, such as elasticity, thin-client user interface as well as data and computational servers that seem to be the well-designed infrastructures for collecting, hosting and processing the BD.

In this paper, we propose the *Data Fusion Module (DFM)* architectural model of low level data fusion suitable for use in conjunction with cloud computing systems. We describe the general model concept and module architectures. We consider two following scenarios: (i) DFM as internal part of the cloud and (ii) DFM external part respectively. We specified a simple realistic use scenario on which future research will be focused. The paper is organized as follows. Section 2 describes low level data fusion architectures. In section 3, we present our new model of low level data fusion. Section 4 describes proposed approaches of integration low level data fusion module with different cloud architectures. Section 5 shows practical application example. Sections 6 and 7 conclude the paper.

LEVEL 1 AND 2 OF JDL: DATA FUSION ARCHITECTURES

The specification of sensor fusion architecture and algorithms is necessary to achieve a coherent and accurate estimation of the environmental model. In this section, we will provide a brief characteristics of three main types of the fusion architectures: (i) centralized, (ii) fully decentralized and (iii) hybrid. In centralized fusion architectures, the only available information is located in "the fusion center". Therefore, the overall characteristics of such architecture can be in fact restricted to the definition of the system responsible for the organization of the information in the fusion center. Such system maintains only central information by using the sensor data. Distributed architecture maintains information for each sensor. That datais combined to the central information. Finally, hybrid architecture maintains the information for each sensor for local association purposes (such as distributed architecture), and manages the central information similarly to the centralized architectures. All the information in any type of fusion architectures is generated and maintained in the fusion centers (centralized or distributed). Real local information is considered only as a context information.

[2] http://38r0us9g9l1438rwf2z2tcsz.wpengine.netdna-cdn.com/wp-content/uploads/2009/08/Twitter-Study-August-2009.pdf

[3] https://faculty.ist.psu.edu/wu/papers/emerse-iscram2011.pdf

[4] https://aws.amazon.com/

[5] https://www.datapipe.com/gogrid/

[6] https://www.rackspace.com/

[7] https://cloud.microsoft.com/en-us/

The fusion system uses collateral information, also named context information that is related with sensor models, teterrain model, dynamical models of surface objects, etc. Fusion processes and collateral information is integrated using three types of architectures: centralized, distributes, hybrid.

Global process of fusion begins with transformation of the local data to a global reference point. It is usually based on transformation of coordinates to common format. After this process the system should integrate the new information of sensors with the system information (this is made in three steps: gating, association and filtering).

Centralized Architecture

The centralized architecture maintains a set of central information, $\{T_i\}$. The data received from sensors, $\{P^{Sk}_j\}$, is associated to the central information and, then, filtered to update the track. In this architecture, the processes (association and filtering) work directly on data and the additional information of the sensor format. Data format is used in the association process to reduce the computational load. Figure 1 shows the typical data fusion processes in the centralized architecture..

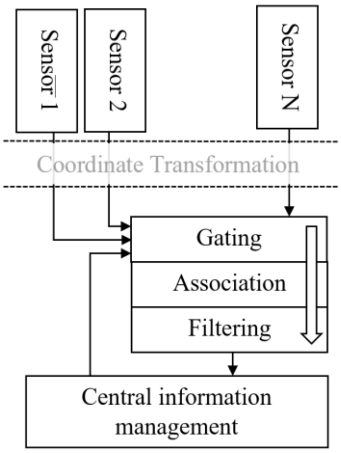


Fig. 1. Centralized architecture

The first step of the data fusion in centralized architectures is the coordinate transformation. In this phase, all data, $\{P^{Sk}_j\}$, (with coordinate values respect to sensor position, $\{S_k\}$), is transformed to unified system, C, $\{P^C_j\}$. In the second phase of data fusion, temporal and kinematic compatibilities are calculated for each pair of measure-central information, $\{P^C_j, T_i\}$. *Gating function[6]* allows to calculate the possibility of measure-to-central information association. The third step is the association process, where, a set of bidimensional matrix,matrixes (one matrix for each sensor) is defined. The rows in each matix are defined by central information and columns are defined by sensor measures. Each matrix position contains either (a) , the value of the distance between measure and central information if the association $\{P^C_j, T_i\}$ is possible, or, (b) zero, if the association is impossible. Munkres algorithm[7] calculates the measure-central information association that minimizes the total distance. Then, central information is updated with the measures associated in the fourth step, the filtering function. The management of central information (generating new central information, deleted central information without measures o fused similar new central information) is the final step.

The following major advantages of centralized architecture can be specified: (a) optimization of the position estimation for any sensor measure variance, and (b) minimalization of the effects of a delay between the time when a manoeuvre begins in the real scenario and detectionof target manoeuvring.

The main disadvantage of this centralized architecture is related to systemic errors among different sensors, due to the difficulty for estimating them and vulnerability in final output to residual non-estimated multi-sensor biases.

Distributed Architecture

Distributed architecture maintains central and local information. In this architecture, local information is maintained for each sensor and the transformation, gating, association and filtering functions are carried out locally to the sensor. The central information is the result of a fusion process over the local information that represents the same central information. This function is similar to the one carried out in the management of central information in centralized architecture. In this case, the fusion process works at local information level, instead of measure level in the centralized architecture. Figure 2 shows the main processes in the distributed architecture.

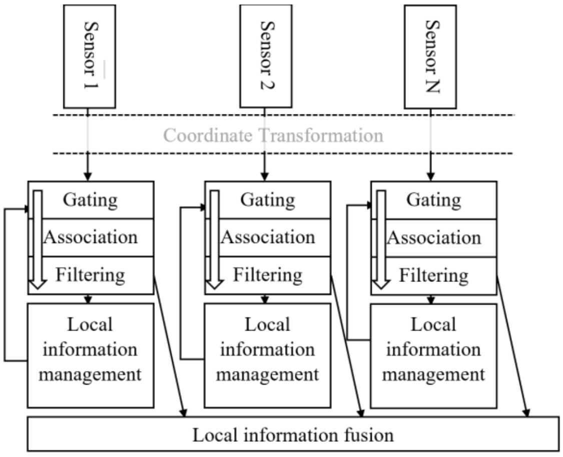


Fig. 2. Distributed architecture

The main advantage of distributed architecture is related with the distributed processing of data that allow the adaptation of the functions specifically to each type of sensor. In this way, we can group the data of a sensor instead of associate each individual measure as in centralized architecture. In our proposal, there are noadvantages for the computational load in the fusion centre as shown in the general scheme presented in section 2, although parallel execution of functions associated to each sensor is possible in a natural way. Other advantages are the ability of sensors calibration and to easier bias estimation than in the centralized architecture. The main disadvantage is the loss of precision in the fused estimators.

Hybrid Architecture

The hybrid architecture is a combination of the previous ones. The fusion system combines the capacity to fuse at measure level or central level. In our implementation, local information is maintained to carry out the association of measures (as in the distributed architecture). However, the central information actualization uses the data instead of local measurement results (as in centralized architecture). Figure 3 shows the processes in this architecture. The mains advantages of the hybrid approach is the high accuracy, systematic error robustness, and the possible distribution of computational load.The disadvantages may be , a higher computational load of the processing of measures: this process is repeated two times.

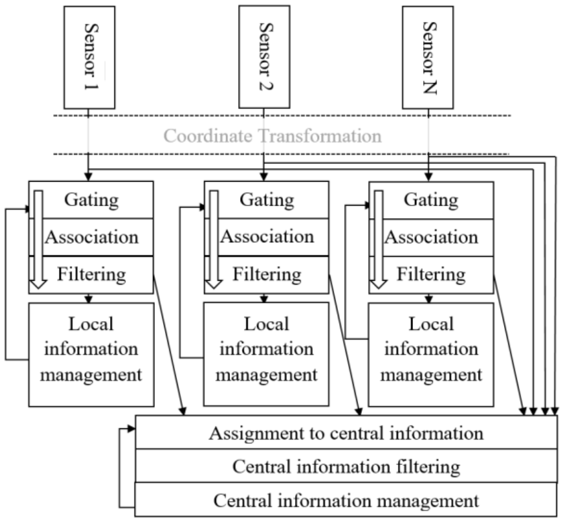

Sensor 1
Sensor 2
Sensor N
Coordinate Transformation
Gating
Association
Filtering
Local information management
Gating
Association
Filtering
Local information management
Gating
Association
Filtering
Local information management
Assignment to central information
Central information filtering
Central information management

Fig. 3. Hybrid architecture

THE CONCEPT OF LOW LEVEL DATA FUSION MODULE (DFM)

Data fusion module is a component of data processing and analytics system that performs data fusion process. The following requirements must be considered for making such module a component of computational cloud: a) it has to incorporate distributed fusion processes, b)it has to employ communication mechanism, c) its communication mechanism should allow to response for asynchronous events, d) it has to be able to use context. Based on the above requirements, we propose the *Data Fusion Module (DFM)* architecture for computational clouds.

Te most important element of our DFM module is a *fusion algorithm* of desired architecture (centralized, distributed or hybrid). The next central component of DFM is the environmental *model.*. That model will be constatntly updated in refinement process. Another important component of DFM is the *context adaptation* module. It transforms the context (if used) to the format suitable for common referrecing, data association, state estimation and fusion management. The other components are: the *preprocessing unit* (which in distributed model can be placed outside the DFM) and *communication layer*, which should use event-driven mechanism. Event driven-mechanismprovides asynchronous flow of information and steering commands. In distributed and hybrid architectures, each component which can work in the same time with the other modules, must have some communication mechanisms embedded

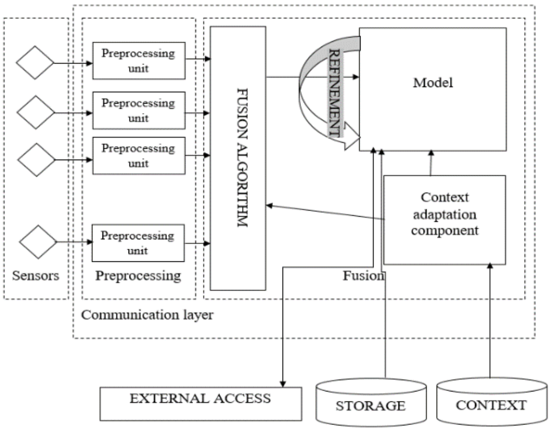


Fig. 4. General architecture of Low Level DFM

Proposed architecture of data fusion module allows DFM to be a part of system that incorporates cloud computing environments. It is achieved mainly due to inclusion asynchronous communication mechanism. It's architecture allows to use it in distributed systems of different kind.

Results from this module are stored in the database. The High Level Data Fusion Module (SA) extracts the knowledge from this database using Big Data techniques that are able to analyse the situation using Cloud capabilities.

INTEGRATION OF LOW LEVEL DFM WITH CLOUD INFRASTRUCTURE

The methods of integration of DFM with coud system depend on particular cloud infrastructure. Platform-as-a Service (PaaS) and Infrastructure-as-aService (IaaS) cloud layers allow us to define DFM as the internal cloud component. . Software-as-a-Sservice (SaaS) layer must be connected to DFM hosted on external machine. Differences of such approaches are clearly visible. In case of DFM embedded in the cloud, there is ability to use distributed data fusion in order to utilize clouds scalability and efficiency. When DFM is located outside the cloud, we use cloud capabilities for data post-processing, data management and storage. In both approaches we use cloud advantages such as flexibility, high availability and its security to incorporate data fusion process.

a) Low level DFM as an internal component of the cloud system

Let us consider DFM as an internal component of the cloud system. and the network with large number of sensors. Such sensors can generate and send the data into the cloud. that the generated data can have different format (depending on sensor). Therefore, the preprocessing must be provided in order to achieve homogeneity of data characteristics. Preprocessing must be provided as a parallel process for different data sets and files, in order to achieve the maximal efficiency. It has also necessity to trigger the whole fusion process by notifying fusion module about end of data processing and incoming data portion. Asynchronous communication mechanism based on event-driven paradigm has to be deployed between corresponding nodes in order to ensure flow control mechanism.

After data preprocessing, the real fusion process is activated. Processes performed by DFM should also be parralelised and distributed. Its internal design should be based on distributed or hybrid architecture. In that manner. event-driven control mechanism should be implemented also between its internal components.. When the fusion process is completed, the results should be stored in database. That database can be internal part of the cloud. Also the database can be placed outside. For storing large amount of data NoSQL databases[8] are preffered. Context which can be useful in fusion process can be stored in cloud or outside. Access to each element of the system should be controlled by specific module/service in order to maintain security level. That model can be applied in PaaS and IaaS cloud layers.

Internal low level DFM can provide outcoming data to high level data fusion module in order to achieve situation awareness or knowledge extraction. High level data fusion can be performed inside the cloud environment or by external component.

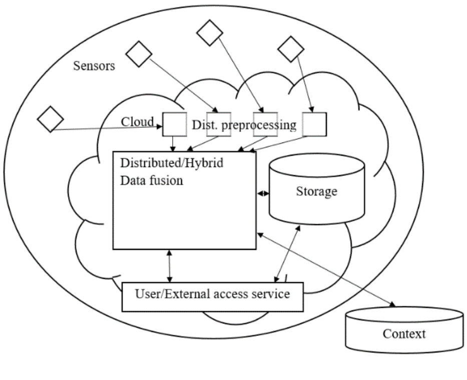


Fig. 5. Proposed architecture for system that incorporates Low level DFM as internal component of cloud system

b) Low level DFM as external component of cloud system

In the case of network of sensors sending data to the cloud, where DFM is not an integral part of the cloud, there is a need of deploying DFM on the external infrastructure (or simply server). In that case, the sensors send their information to the cloud or directly to DFM. In the first case, the cloud passes data to the DFM. DFM preprocessing module generates the unified format for heterogenous data and the fusion process is intialized. The results of the fusion process are sent to the cloud for postprocessing or storage. There is also a need to establish communication platform between sensors network and data fusion and also between cloud and DFM. In that case, the cloud system is used only to preprocess, postprocess or for data storage. Context which can be useful can be delivered from external source of data. Preffered type of architecture of DF algorithm is centralized architecture. Mainly due to putting DFM in external machine. That is case that use Software as a service cloud type.

Internal low level DFM can provide outcoming data to high level data fusion module in order to achieve situation awareness, knowledge extraction, threat recognition or prediction.

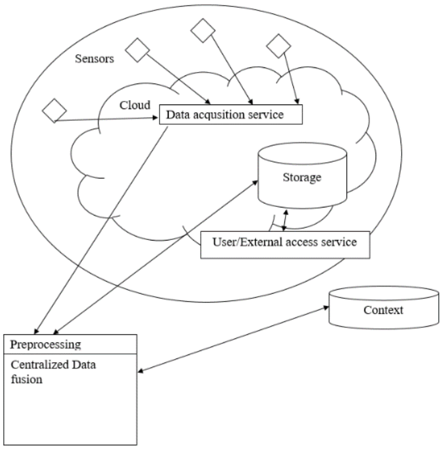


Fig. 6. Proposed architecture for system that incorporates Low level DFM as external component of cloud system

REALISTIC SCENARIO

Data fusion in conjuction with cloud computing could be applied in meteorology. Let us consider spatially distributed network of platforms of sensors in Poland (see map on Fig 7).

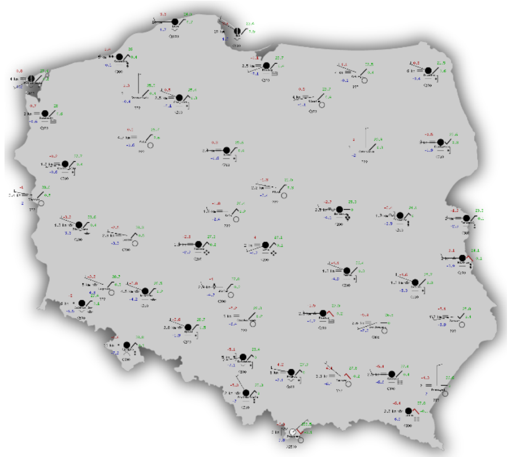


Fig. 7. Considered spatial distribution of sensor platforms in Poland

Each platform is composed from sensors which measure: temperature, pressure, relative humidity, precipitation, visibility, cloud cover, cloud base height, wind speed and its direction. Also information about sensors localization are sent. Measurements are made with one minute interval, and then they are sent do the cloud system trough the Internet.

Such sensor networks can be connected with DFM for the fusion process. The most challenging part of fusion process is to recognize the problems of tracking and data association.

Data generated by sensors can be supplemented with results of radiosonde observations after process of fixing [9].

In mentioned meteorogical scenario we have to deal with multiple tracking problems: tracking in time and spatial tracking problem. Let's consider the data generated by specific sensor of single weather station. That data can be defined as a time series of values. The tracking process is based on observing actual readings and comparing them to the predicted values. In the case of some problems or tracking interruptions, (for example due to sensor malfunction), The fusion module should make decision to correct the values of the generated data.

The second source of data for fusion can be context. For instance information about terrain configuration (elevation of stations and theirs surrounding areas) could be useful. For example in checking if use of interpolation method is useful in order to get value at point between two sensor's places (for mountainous area it is not desired). The second good example of context usefulness is detection of wrong values delivered by the sensors (for example temperatures below freezing point during heat waves in summer).

When fusion process is completed data delivered by fusion is stored in the cloud in distributed NoSQL Database (Hbase, RIAK etc) in form of tuple (place, time, type of measurement and value). Access to database and fusing module will be provided by external access service.

Figure 8 shows architecture schematic, proposed as the problem solution.

Implementation and testing of this system is still in progress. Results will be presented in future.

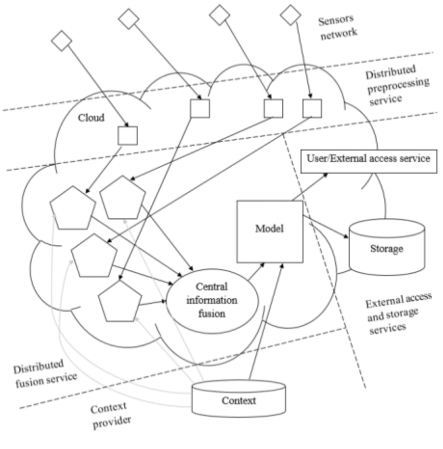

Sensors network
Distributed preprocessing service
Cloud
User/External access service
Model
Storage
Central information fusion
External access and storage services
Distributed fusion service
Context
Context provider

Fig. 8. Proposed architecture for meteorological problem. Direction of data and events flow is counterclockwise

CONCLUSIONS AND FUTURE WORK

BD analytic methods need to be designed with an appreciation of the significant trade-off that exists between Computational Time and Quality of Solutions. This trade-off between time and accuracy requires big data analytic architectures to support both Batch Data Analytic Processes (for latent but quality solutions) as well as Streaming Data Analytics (for near real time situation awareness). This exponential explosion has a tremendous effect to the information fusion community that constantly deals with Big Streaming Data.

The deployment and provision of Big data applications will greatly benefit from a cloud-like middleware infrastructure, which could be formulated to create a hybrid environment consisting of multiple private (municipalities, enterprises, research organisations) and public (Amazon, Microsoft) infrastructure providers to deliver on-demand access to such Big data applications.

ACKNOWLEDGMENTS

The inspiration for the research presented in this paper is the resutl of the STSM visit of Piotr Szuster in within STSM programme of the IC1406 COST Action Horizon 2020 project cHiPSet "High-Performance Modelling and Simulation for Big DataApplications", His work has been supervised by Prof. Molina Lopez and prof. Kołodziej.

REFERENCES

Besada J. A., García J., de Miguel G., Berlanga A., Molina J. M., Casar J.R., "Design of IMM Filters using Evolution Strategies for Radar Tracking Applications". IEEE Transactions on Aerospace and Electronic Systems. vol. 41, issue 3, pp 1109-1122, Julio 2005.

Molina J. M., García J., Jiménez F.J., Casar J.R., "Cooperative Management of a Net of Intelligent Surveillance Agent Sensors", International Journal of Intelligent Systems. vol 18, nº 3 pp 279 - 307, 2003.

Molina J. M., García J., Jiménez F.J., Casar J.R. "Fuzzy Reasoning in a MultiAgent System of Surveillance Sensors to Manage Cooperatively the Sensor-to-Task Assignment Problem". Applied Artificial Intelligence. vol 18, nº 8 pp 673-713, Septiembre 2004.

E. Waltz, J. Llinas. Multisensor Data Fusion, Artech House Inc., Norwood, MA, 1990.

J. Manyika, H. Durrant-Whyte, Data Fusion and Sensor Management a decentralized information-theoretic approach, Ellis Horwood, 1994.

Jitendra R. Raol, Data Fusion Mathematics: Theory and Practice CRC Press, pp 170-172,August 2015 - 600

Anthony K. Hyder, E. Shahbazian, E. Waltz, Springer Science & Business Media, pp 79-85, December 2012

Strauch, Christof, Ultra-Large Scale Sites, and Walter Kriha. "NoSQL databases." Lecture Notes, Stuttgart Media University (2011).

Piotr Szuster, Data Fixing Algorithm in Radiosonde Monitoring Process, Journal of Telecommunications and Information Technology, 2017, vol. 1, pp. 81-88

PIOTR SZUSTER graduated in Computer Science at Cracow University of Technology, Poland, in 2016. Currently, he is a research and teaching assistant at Cracow University of Technology and a Ph.D. student at AGH University of Science and Technology. The main topics of his research are Big Data, Data Fusion and Internet of Things. For more information, please visit: http://www.retsuz.pl. His e-mail address is: pszuster @pk.edu.pl.

JOSE MANUEL MOLINA LOPEZ received a degree in Telecommunication Engineering from the Universidad Politecnica de Madrid in 1993 and a Ph.D. degree from the same university in 1997. He joined the Universidad Carlos III de Madrid in 1993 where, actually, he is Full Professor at Computer Science Department. Currently he leads the Applied Artificial Intelligence Group (GIAA, http://www.giaa.inf.uc3m.es) involved in several research projects related with ambient intelligence, surveillance systems and context based computing. His current research focuses in the application of soft computing techniques (Multiagents Systems, Evolutionary Computation, Fuzzy Systems) to Data Fusion, Data Mining, Surveillance Systems (radar, Video, etc..), Ambient Intelligence and Air/Maritime Traffic Management. He is author up to 70 journal papers in JCR journals and 200 conference papers. He has 6000 cites in Google-Citation and h=27.

JESÚS GARCÍA-HERRERO
Jesus Garcia Herrero is Associate Professor at the Universidad Carlos III de Madrid, Computer Science Department. He received his M.Sc. and PhD in Telecommunications Engineering from Universidad Politecnica de Madrid. His main research interests are computational intelligence, sensor and information fusion, surveillance systems, machine vision, air traffic management and autonomous systems. Within these areas, including theoretical and applied aspects, he has co-authored more than 10 book chapters, 50 journal papers and 120 conference papers.

He has served on several advisory and programming committees in organizations such as IEEE (senior member), ISIF and NATO. He is the chair of the Spanish IEEEE Chapter on Aerospace and Electronic Systems since 2013 and appointed Spanish member of several NATO-STO Research Groups focused on information fusion

JOANNA KOŁODZIEJ has graduated in Mathematics from the Jagiellonian University in Cracow in 1992, where she also obtained the PhD in Computer Science in 2004. She is employed at Cracow University of Technology as a professor. She has served and is currently serving as PC Co-Chair, General Co-Chair and IPC member of several international conferences and workshops including PPSN 2010, ECMS 2011, CISIS 2011, 3PGCIC 2011, CISSE 2006, CEC 2008, IACS 2008-2009, ICAART 2009-2010. Prof. Kołodziej is a Managing Editor of IJSSC Journal and serves as a EB member and guest editor of several peer-reviewed international journals. For more information, please visit: www.joannakolodziej.org

PROFILING AND RATING PREDICTION FROM MULTI-CRITERIA CROWD-SOURCED HOTEL RATINGS

Fátima Leal
Universidade de Vigo
School of Telecommunications Engineering
INESC TEC, Porto
E-mail: fatimaleal2@gmail.com

Horacio González–Vélez
National College of Ireland
Cloud Competency Centre
E-mail: horacio@ncirl.ie

Benedita Malheiro
Instituto Politécnico do Porto
School of Engineering
INESC TEC, Porto
E-mail: mbm@isep.ipp.pt

Juan Carlos Burguillo
Universidade de Vigo
School of Telecommunications Engineering
E-mail: J.C.Burguillo@uvigo.es

KEYWORDS

Collaborative Filtering, Personalisation, Prediction Models, Multi-criteria Ratings, Tourism, Crowd-Sourcing, Recommender Systems, Data Analytics.

ABSTRACT

Based on historical user information, collaborative filters predict for a given user the classification of unknown items, typically using a single criterion. However, a crowd typically rates tourism resources using multi-criteria, *i.e.*, each user provides multiple ratings per item. In order to apply standard collaborative filtering, it is necessary to have a unique classification per user and item. This unique classification can be based on a single rating – single criterion (SC) profiling – or on the multiple ratings available – multi-criteria (MC) profiling. Exploring both SC and MC profiling, this work proposes: ($\imath$) the selection of the most representative crowd-sourced rating; and ($\imath\imath$) the combination of the different user ratings per item, using the average of the non-null ratings or the personalised weighted average based on the user rating profile. Having employed matrix factorisation to predict unknown ratings, we argue that the personalised combination of multi-criteria item ratings improves the tourist profile and, consequently, the quality of the collaborative predictions. Thus, this paper contributes to a novel approach for guest profiling based on multi-criteria hotel ratings and to the prediction of hotel guest ratings based on the Alternating Least Squares algorithm. Our experiments with crowd-sourced Expedia and TripAdvisor data show that the proposed method improves the accuracy of the hotel rating predictions.

INTRODUCTION

Information and Communication Technology has revolutionised the tourist behaviour. In particular, the mobile technology provides tourists with permanent access to endless web services which influence their decisions. Well-known tourism business-to-customer on-line platforms (*e.g.*, TripAdvisor, Expedia, airbnb, *etc.*) aim to support travellers by providing additional information regarding tourism resources. Furthermore, the on-line tourism-related services enable tourists to share (*e.g.*, photos or videos), comment (*e.g.*, reviews or posts) and rate (*e.g.*, ratings or likes) their travel experiences. Consequently, these tourism services become, while gatherers of voluntarily shared feedback information, crowdsourcing platforms (Egger et al. 2016). The value of the crowd-sourced tourism information is crucial for businesses and clients alike. In the case of this work, it enables the modelling of tourists and tourism resources using multi-criteria ratings to produce suitable recommendations.

Personalised recommendations are often based on the prediction of user classifications, whereby, accurate prediction is essential to generate useful recommendations. Typically, the crowd-sourced classification of hotels involves multi-criteria ratings, *e.g.*, hotels are classified in the Expedia platform in terms of *cleanliness*, *hotel condition*, *service & staff*, *room comfort* and *overall* opinion. We argue that the personalised combination of multi-criteria item ratings improves the tourist profile and, consequently, the accuracy of the collaborative predictions.

Collaborative filtering is a classification-based technique, *i.e.*, depends on the classification each user gave to the items he/she was exposed to (Breese et al. 1998). Typically, this classification corresponds to a unique rating. Whenever the crowd-sourced data holds multiple ratings per user and item, first, it is necessary to decide which user classification to use in order to apply collaborative filtering. This work explores both single criterion (SC) – chooses the most representative of the crowd-sourced user ratings (Leal et al. 2017) – and multi-criteria (MC) profiling approaches –combines the different crowd-sourced user ratings per item, using the Non-Null Rating Average (NNRA) or the Personalised Weighted Rating Average (PWRA), *i.e.*, based on the individual user rating profile.

This research contributes to guest and hotel profiling –

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

based on multi-criteria ratings – and to the prediction of hotel guest ratings – based on the the Alternating Least Squares with Weighted-λ-Regularisation (ALS-WR) matrix factorisation algorithm. Our experiments with crowd-sourced Expedia and TripAdvisor data proved that the proposed profiling method improves the ALS-WR prediction accuracy of unknown hotel ratings. In particular, when faced with null multi-criteria user ratings, the most accurate predictions were achieved with the Personalised Weighted Rating Average combination.

This paper is organised as follows. The related work section reviews personalisation via crowd-sourced ratings. The proposed method section describes the approach and algorithms used. The experiments and tests section reports the data set, tests performed and the results obtained. Finally, the conclusions section summarises and discusses the outcomes of this work.

RELATED WORK

Technology plays an important role in the hotel and tourism industry. Both tourists and businesses benefit from technology advances regarding communication, reservation and guest feedback services. Tourists use tourism Web services to organise trips, *i.e.*, to search, book and share their opinions in the form of ratings, textual reviews, photos, *etc.*, creating a digital footprint. This permanent interaction between tourists and tourism Web services and mobile applications generates large volumes of precious data.

The tourist profiles, which are based on the individual digital footprints, are used by recommendation systems to personalise recommendations. Thus, refined tourist profiles will increase the quality of the recommendations and, ultimately, the tourist experience.

Collaborative filtering is a popular recommendation technique in the tourism domain. It often relies on rating information voluntarily provided by tourists, *i.e.*, crowd-sourced ratings, to recommend unknown resources to other tourists. Well-known tourism crowdsourcing platforms, *e.g.*, TripAdvisor or Expedia, allow users to classify tourism resources using multi-criteria, *e.g.*, *overall*, *service*, *cleanliness*, *etc*. Thus, profiling and prediction using tourism crowd-sourced multi-criteria ratings is an important research topic for the hospitality industry.

Adomavicius and Kwon (2015), Bilge and Kaleli (2014), Lee and Teng (2007), Jhalani et al. (2016), Liu et al. (2011), Manouselis and Costopoulou (2007), and Shambour et al. (2016) have explored the integration of multi-criteria ratings in the user profile, mainly using multimedia data sets to validate their proposals. However, scant research considers crowd-sourced multi-criteria ratings for profiling and rating prediction applied to the tourism domain.

Jannach et al. (2012) apply the Adomavicius and Kwon (2007) methods to incorporate multi-criteria ratings in the tourist profile based on Support Vector Regression (SVR). It combines a user and item models, using a weighted approach, to provide better recommendations. The evaluation was performed with a data set provided by HRS.com.

Fuchs and Zanker (2012) perform multi-criteria rating analysis based on a TripAdvisor data set. First, they use Multiple Linear Regression (MLR) to identify correlations, patterns, and trends among the TripAdvisor data set parameters. Then, the authors apply the Penalty-Reward-Contrast analysis proposed by Randall Brandt (1988) to establish tourist satisfaction levels based on multi-criteria ratings. This work proposes a methodology for MC rating analysis.

Nilashi et al. (2015) propose a SC profiling approach together with a hybrid hotel recommendation model for multi-criteria recommendation. They employed: (*i*) Principal Component Analysis (PCA) for the selection of the most representative rating (dimensionality reduction); (*ii*) Expectation Maximisation (EM) and Adaptive Neuro-Fuzzy Inference System (ANFIS) as prediction techniques; and (*iii*) TripAdvisor data for evaluation.

Farokhi et al. (2016) explore SC profiling together with collaborative filtering. First, the authors selected the *overall* as the most representative rating after determining the correlation between the multiple ratings, then applied data clustering (Fuzzy c-means and k-means) to find the nearest neighbours and, finally, predicted the unknown hotel ratings using the Pearson Correlation coefficient. The evaluation was performed with TripAdvisor data.

Ebadi and Krzyzak (2016) developed an intelligent hybrid multi-criteria hotel recommender system. The system uses both textual reviews and ratings from TripAdvisor. Regarding the ratings, it adopts SC profiling to learn the guest preferences and Singular Value Decomposition (SVD) matrix factorisation to predict unknown ratings.

Contributions

This paper explores profiling and prediction using tourism crowd-sourced multi-criteria ratings. The main goal is to refine guest and hotel profiling by reusing the multiple hotel ratings each guest shares. According to Nilashi et al. (2015) and Adomavicius and Kwon (2015), collaborative filtering with multi-criteria item ratings has been unexplored when compared with its single criterion item rating counterpart. The current work proposes and compares different ways of utilising multi-criteria user ratings to improve the accuracy of predictions. Furthermore, when compared with other research found in the literature (Table 1), our work uses: (*i*) single and multiple rating profiling; (*ii*) employs ALS-WR as predictive technique; and (*iii*) Expedia (E) and TripAdvisor (TA) crowd-sourced data for evaluation.

TABLE 1: Comparison of Multi-Criteria Research Approaches

Approach	Evaluation	Profiling	Prediction
Jannach et al. (2012)	HRS	MC	SVR
Fuchs and Zanker (2012)	TA	MC	–
Nilashi et al. (2015)	TA	SC	ANFIS
Farokhi et al. (2016)	TA	SC	k-means
Ebadi and Krzyzak (2016)	TA	SC	SVD
Leal et al. (2017)	**TA & E**	**SC & MC**	**ALS-WR**

PROPOSED METHOD

Typically, a collaborative recommendation filter relies on an unique rating to produce recommendations, *i.e*, in standard rating-based recommendation systems, the user is modelled using a single rating. However, in tourism crowd-sourcing platforms, the tourist-related data encompasses multi-criteria ratings.

This paper addresses the problem of personalisation via crowd-sourced multi-criteria tourism ratings. Our proposed method has four modules: (*i*) Data Collection for gathering data from Expedia platform; (*ii*) Rating Analysis for exploring distinct profiling approaches based on multi-criteria ratings; (*iii*) Rating Prediction for predicting unknown ratings; and (*iv*) Evaluation Metrics for assessing the profiling and recommendation results.

Data Collection

Expedia (`http://www.expedia.com`) is a powerful platform which contains large volumes of crowd-sourced hotel opinions. Moreover, Expedia owns a host of on-line brands, including TripAdvisor, Hotels.com or trivago. According to Law and Chen (2000) (Law and Chen 2000), Expedia brands cover researching, booking, experiencing and sharing travels. The platform allows choosing flights or hotels, reading personal reviews of hotels, classifying hotels using textual reviews and ratings as well as planning new travels.

Taking into account these characteristics, we collected different crowd-sourced ratings via the Expedia API (`http://developer.expedia.com/directory`). In the Expedia platform tourists classify hotels using multi-criteria ratings: *overall*, *cleanliness*, *hotel condition*, *service & staff* and *room comfort*. Based on these multiple criteria classifications, we create, using different approaches, unique personalised ratings per tourist and hotel.

Rating Analysis

The rating analysis module explores different profiling approaches based on crowd-sourced multi-criteria ratings. First, we apply a Multiple Linear Regression (MLR) to identify the Most Representative Rating (MRR). Then, we combine the crowd-sourced multi-criteria user ratings into a single rating using NNRA and PWRA.

Multiple Linear Regression is typically applied to multivariate scenarios in order to predict one or more continuous variables based on other data set attributes, *i.e.*, by identifying existing dependencies among variables (Sykes 2000). First, we determine the correlation between the multi-criteria ratings to identify the dependent variable and, then, perform MLR to estimate the relation between the identified dependent variable and the remaining set of explanatory variables. Equation 1 displays the model of the MLR with k regression variables where ϵ_i is the disturbance, β_0 is the intercept and β_i ($i = 1$ to k) are the partial regression coefficients, representing the rate of change of Y as a function of the changes of $X = \{x_1, x_2, ..., x_k\}$ (Tranmer and Elliot 2008).

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_k x_{ki} + \epsilon_i \quad (1)$$

We use Ordinary Least Squares (OLS) to estimate the unknown parameters (β_{ki}) of this linear regression model. OLS minimises the distance between the observed responses and the responses predicted by the linear approximation (Stone and Brooks 1990). Equation 2 represents the OLS method where x_i and y_i are the observations and $\hat{x}$ and $\hat{y}$ the predictions.

$$\widehat{\beta} = \frac{\sum(x_i - \hat{x})(y_i - \hat{y})}{\sum(x_i - \hat{x})^2} \quad (2)$$

Rating Combination explores two multi-criteria item rating combination methods: (*i*) the Non-Null Rating Average (NNRA); and (*ii*) the Personalised Weighted Rating Average (PWRA). The non-null rating average $r_{u,i}$ is defined by Equation 3 where $r_{u,i,j}$ is the non-null rating of type j given by user u to item i and n is the number of non-null ratings given by user u to item i.

$$r_{u,i} = \frac{\sum_{j=1}^{n} r_{u,i,j}}{n} \quad (3)$$

Equation 4 displays the personalised weighted rating average $r_{u,i}$ where $r_{u,i,j}$ is the non-null rating of type j given by user u to item i, $n_{u,j}$ the number of non-null ratings of type j given by user u and n is the total number of non-null multi-criteria ratings given by user u.

$$r_{u,i} = \frac{\sum_{j=1}^{n} n_j r_{u,i,j}}{\sum_{j=1}^{n} n_{i,j}} \quad (4)$$

Rating Prediction

The rating prediction module aims to predict unknown hotel ratings, *i.e.*, hotels not yet classified by the tourists, by implementing a user-based collaborative recommendation filter. We use the Alternating-Least-Squares with Weighted-λ-Regularisation (ALS-WR) algorithm since, according to (Zhou et al. 2008, Hu et al. 2008), it provides better results than other matrix factorisation approaches despite its higher execution time. ALS-WR employs matrix factorisation to represent tourists and hotels as vectors of latent factors. The rating matrix (R_{u*i}) holds for all users and items the corresponding user u item i rating. For recommendations purposes, the algorithm factorises the matrix R_{u*i} into two latent matrices: (*i*) the user-factor matrix P; and (*ii*) the item-factor matrix Q. Equation 5 represents this factorisation where each row p_u of P or q_i of Q represents the relation between the corresponding latent factor and the user u or item i, respectively, and λ regularises the learned factors (Friedman et al. 2016).

$$\min_{P,Q} \sum_{r_{u,i} \in R} \left[(r_{u,i} - p_u q_i^\intercal)^2 + \lambda \left(||p_u||^2 + ||q_i||^2 \right) \right] \quad (5)$$

Finally, R is approximated by the product of P and Q, *i.e.*, each known rating $r_{u,i}$ is approximated by $\hat{r}_{u,i} = p_u \cdot q_i^\intercal$.

Algorithm 1 summarises the ALS-WR iterative implementation. In each iteration, P and Q are sequentially fixed to solve the optimisation problem. Once the latent vectors converge, the algorithm ends. We defined the regularisation weight λ, the dimensionality of latent feature space (k) and the number of iterations (n) based on the above mentioned research works. The final matrix holds all user item rating predictions used for recommendation.

Algorithm 1 ALS-WR

```
Inputs    User $u$, Item $i$ and $r_{u,i}$
Outputs   User $u$, Item $i$ and $\hat{r}_{u,i}$
Step 1    Matrix Factorisation with $\lambda = 0.1$, $k = 20$ and $n = 20$
Step 2    Create the $P$ and $Q$ latent matrices
Step 3    Fix $Q$ and estimate $P$
Step 4    Apply ALS-WR
Step 5    Fix $P$ and estimate $Q$
Step 6    Apply ALS-WR
Step 7    Calculate prediction matrix
```

Evaluation Metrics

The evaluation of recommendation systems involves predictive accuracy and classification metrics. On the one hand, the predictive accuracy metrics measure the error between the predicted rating and the real user rating. It is the case of the Mean Absolute Error (MAE), which measures the average absolute deviation among the predicted rating and the real rating, or the Root Mean Square Error (RMSE), which highlights the largest errors (Herlocker et al. 1999). Equation 6 and Equation 7 represent both error functions where $\hat{r}_{u,i}$ represents the rating predicted for user u and item i, $r_{u,i}$ the rating given by user u to item i, m the total number of users and n the total number of items.

$$MAE = \frac{1}{u} \times \sum_{u=1}^{m} \left(\frac{1}{n} \times \sum_{i=1}^{n} |\hat{r}_{u,i} - r_{u,i}| \right) \tag{6}$$

$$RMSE = \frac{1}{u} \times \sum_{u=1}^{m} \left(\sqrt{\frac{1}{n} \times \sum_{i=1}^{n} (\hat{r}_{u,i} - r_{u,i})^2} \right) \tag{7}$$

On the other hand, the classification accuracy metrics, which include Precision and Recall and range from 1 (best) to 0 (worst) (Basu et al. 1998). Recall determines the percentage of relevant items selected from the total number of relevant items available (Equation 9). Precision defines the percentage of relevant items selected from the total number of items (Equation 8). Equation 9 and Equation 8 detail both metrics where TP is the number of relevant items recommended by the system or true positives, FN is the number of relevant items not recommended by the system or false negatives and FP corresponds to the number of irrelevant items recommended by the system or false positives.

$$Precision = \frac{TP}{TP + FP} \quad (8) \quad Recall = \frac{TP}{TP + FN} \quad (9)$$

Finally, the quality of the top N recommendations can be determined using Recall@N metric. In particular, Nilashi et al. (2015) define Recall@N according to Equation 10, where TP is the number of true positive or relevant items and Top_N is the list of the top N recommended items.

$$Recall@N = \frac{TP \cap Top_N}{TP} \tag{10}$$

EXPERIMENTS AND RESULTS

We conducted several off-line experiments with the HotelExpedia data set (`http://ave.dee.isep.ipp.pt/~1080560/ExpediaDataSet.7z`) and the TripAdvisor data set (Wang et al. 2010) to evaluate the proposed method. The data processing was implemented in Python using the *scikit-learn* library (`http://scikit-learn.org`). Our system is running on a cloud OpenStack instance, holding 16 GB RAM, 8 CPU and 160 GB hard-disk. The experiments involved MRR, NNRA and PWRA profiling, rating prediction and rating prediction evaluation.

Data Sets

The experiments were performed with the HotelExpedia and TripAdvisor data sets. The data set was randomly partitioned into training (75 %) and test (25 %) in order to perform the off-line profiling and rating prediction.

Expedia Table 2 describes the contents of the data set. It is composed of 6276 hotels, 1090 identified users and 214 342 reviewers from 11 different locations. Each user classified at least 20 hotels and each hotel has a minimum of 10 ratings. Our experiments, which rely on the hotel, user and hotel user review data, use, specifically, the user nickname, the hotel identification and, as multi-criteria ratings, the *overall*, *cleanliness*, *service & staff*, *hotel condition* and *room comfort*. This data set does not contain null ratings, *i.e.*, all users rated the hotels according to the multiple criteria.

TABLE 2: HotelExpedia Data Set

Entities	Features
Hotels	`hotelId`, `description`, `latitude-longitude`, `starRating`, `guestReviewCount`, `price`, `amenity`, `overall`, `recommendedPercent`, `cleanliness`, `serviceAndstaff`, `roomComfort`, `hotelCondition`
Users & Reviews	`nickname`, `userLocation`, `hotelId`, `overall`, `cleanliness`, `hotelCondition`, `serviceAndStaff`, `roomComfort`, `reviewText` `timestamp`

TripAdvisor Table 3 describes the contents of the data set, which is composed of 9114 hotels, 7452 users and 235 793 hotel reviews. Our experiments reuse the user and hotel identification and, as multi-criteria ratings, the *overall*, *value*, *rooms*, *location*, *cleanliness*, *service*, and *sleep quality*. This data set contains 14 % of null ratings.

TABLE 3: TripAdvisor Data Set

Entities	Features
Hotels	`name`, `hotelURL`, `price`, `hotelID`, `imgURL`
Users & Reviews	`authorLocation`, `title`, `author`, `reviewID`, `reviewText`, `date`, `overall`, `value`, `rooms`, `location`, `cleanliness`, `service`, `sleepQuality`

Rating Analysis and Prediction

First, we analysed the available multi-criteria guest ratings per hotel and, then, applied Algorithm 1 to predict the unknown hotel ratings. The rating analysis comprised two different approaches: (*i*) the identification of the most representative hotel rating; and (*ii*) the combination of the multi-criteria guest ratings per hotel into a unique guest rating per hotel.

Most Representative Rating This rating analysis determines the correlation between the multiple hotel ratings to recognise the most correlated rating and, then, estimates and quantifies the relationship between this rating (dependent variable) and the remaining ratings (independent variables) using Multiple Linear Regression. The *overall* rating resulted as the most correlated rating (dependent variable) and, thus, can be estimated in terms of the remaining ratings (independent variables) for both HotelExpedia and TripAdvisor data sets. Table 4 displays the OLS MLR results where β_i are the regression coefficients and R^2 quantifies the response variable variation that is explained by the model.

TABLE 4: MLR Results for the Overall Rating

Data Set	Independent Features	β_i	R^2
Hotel Expedia	Service & Staff	0.32	0.80
	Hotel Condition	0.30	
	Room Comfort	0.29	
	Cleanliness	0.11	
Trip Advisor	Value	0.23	0.78
	Service	0.22	
	Rooms	0.18	
	Cleanliness	0.14	
	Location	0.12	
	Sleep Quality	0.10	

In the case of HotelExpedia, the results show that the independent variables (*cleanliness*, *hotel condition*, *room comfort* and *service & staff*) are capable of explaining approximately 80 % of the dependent variable. The regression was performed with 214 343 multi-criteria ratings. In the case of TripAdvisor, Leal et al. (2016) report that the independent variables (cleanliness, location, rooms, service, sleep quality and value) are capable of explaining approximately 78 % of the dependent variable (*overall*). Based on these results, we chose the *overall* rating as the Most Representative Rating (MRR) of both HotelExpedia and TripAdvisor and, then, performed the *overall* rating prediction. Figure 1 plots the Normalised RMSE (NRMSE) of the predictions for the training and test data partitions of both data sets. In both cases the NRMSE decreases monotonically and converges over time to approximately 0.138 (training) and 0.196 (test) using Expedia data and 0.05 (training) and 0.215 (test) using TripAdvisor data.

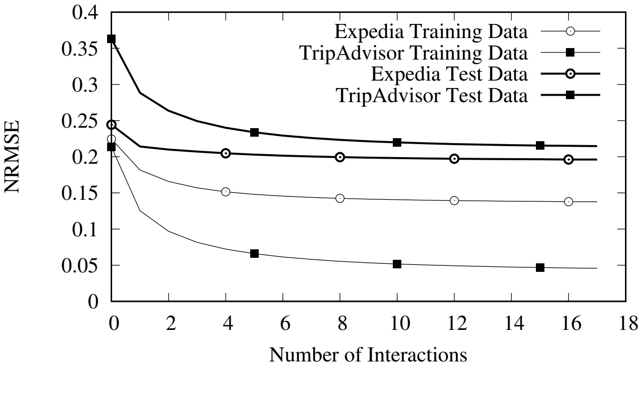


Fig. 1. NRMSE of the Predictions with MRR Profiling

Rating Combination The second rating analysis combines the multi-criteria guest ratings per hotel into a single guest rating per hotel. As a first approach, we calculated the Non-Null Rating Average (NNRA) with Equation 3 and performed the rating prediction using the NNRA rating. Figure 2 plots the NRMSE of the predictions for the training and test data partitions. In both cases the NRMSE decreases monotonically and converges over time to 0.123 (training) and 0.167 (test) using Expedia data and 0.045 (training) and 0.191 (test) using TripAdvisor data.

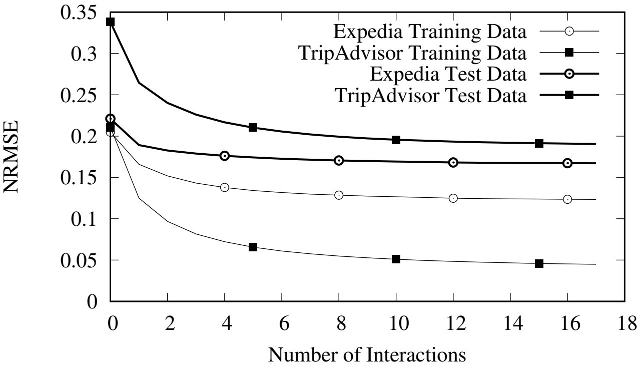


Fig. 2. NRMSE of the Predictions with NNRA Profiling

As an alternative combination approach, we applied the Personalised Weighted Rating Average (PWRA), according to Equation 4, and generated the predictions. Figure 3 plots the NRMSE of the training and test data rating predictions based on the PWRA rating. The NRMSE decreases monotonically and converges over time to 0.123 (training) and 0.167 (test) for Expedia and 0.045 (training) and 0.186 (test) for TripAdvisor.

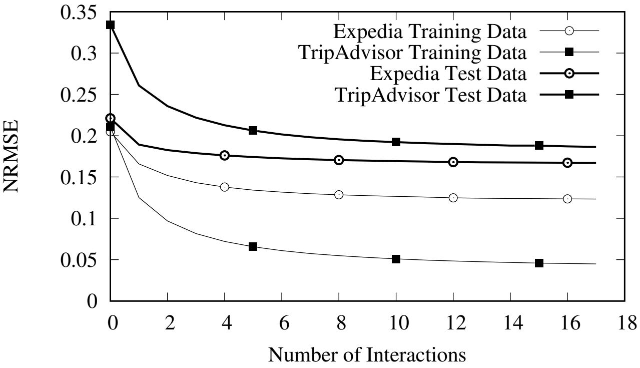


Fig. 3. NRMSE of the Predictions with PWRA Profiling

Recommendations

Finally, we recommend hotels to potential guests with the support of ALS-WR predictions and PWRA profiling. Our novel profiling approach reuses the complete collection of multi-criteria hotel classifications available. To obtain recommendations, the user introduces a desired location and the application provides the top 10 hotel predictions for the user and location. The effectiveness of this recommendation engine was measured using the Recall@10, *i.e.*, considering the top 10 hotel predictions per user.

Discussion

Table 5 compares the global predictive (NRMSE and NMAE), and classification (Recall and Recall@10) accuracy for the test data with the most representative rating (MRR), the Non-Null Rating Average (NNRA) and the Personalised Weighted Rating Average (PWRA) profiling approaches. The results correspond to the average of ten tests. Lower error values and higher classification values indicate higher prediction accuracy. Since the global Precision is one (1), we only present Recall-based classification results. The MMR profiling, which corresponds to the usage of the standard *overall* rating, is the base profiling approach.

The NNRA and PWRA results with the HotelExpedia data set, which has no null ratings, are naturally equal, whereas, with the TripAdvisor data set, which includes 14 % of null ratings, they are not only distinct, but favourable to PWRA. In particular, with HotelExpedia, the NNRA and PWRA profiling, when compared with the MMR approach, improve the NRMSE 14.8 %, the NMAE 26.6 %, the Recall 5.5 % and the Recall@10 17.9 %. In the TripAdvisor case, the results with the PWRA profiling, when compared with those of the MMR approach, improve the NRMSE 13.5 %, the NMAE 14.7 %, the Recall 6.6 % and the Recall@10 16.1 %.

TABLE 5: Prediction Metrics Results

	Profiling	NRMSE	NMAE	Recall	Recall@10
Hotel Expedia	MRR	0.196	0.173	0.254	0.801
	NNRA	0.167	0.127	0.268	0.944
	PWRA	0.167	0.127	0.268	0.944
Trip Advisor	MRR	0.215	0.143	0.351	0.753
	NNRA	0.191	0.125	0.363	0.822
	PWRA	0.186	0.122	0.374	0.874

In terms of the accuracy of the rating predictions, these results show that: (*i*) NNRA and PWRA are preferable to MRR profiling; and (*ii*) PWRA, when faced with null multi-criteria user ratings, outperforms both MMR and NNRA profiling.

CONCLUSIONS

Tourism crowdsourcing platforms, *e.g.*, Expedia and TripAdvisor, collect large volumes of feedback data regarding tourism resources, including multi-criteria ratings, textual reviews, photos, *etc*. The crowd-sourced tourist profile corresponds this individual digital footprint.

The present work explores crowd-sourced multi-criteria rating profiling together with collaborative filtering to provide hotel recommendations. In order to apply standard collaborative filtering, it is necessary to provide the filter with a single classification per user and item. To address this problem, *i.e.*, use multi-criteria ratings for profiling, we designed and experimented with two main approaches: (*i*) the identification of the most representative rating (MRR) with MLR; and (*ii*) the combination of the multi-criteria ratings into a single rating per user and item using NNRA and PWRA. The predictions were performed using the ALS-WR matrix factorisation technique.

The experiments, which were conducted with Expedia and TripAdvisor crowd-sourced multi-criteria hotel ratings, showed that the highest ALS-WR prediction accuracy occurs with the personalised weighted rating average profiling. Based on these results, we adopted the PWRA profiling for the prediction of hotel guest ratings.

In terms of contributions, this research work provides a novel profiling approach based on crowd-sourced multi-criteria ratings which improves the ALS-WR hotel rating prediction accuracy.

As future work, we intend to: (*i*) cluster hotels taking into account their crowd-sourced value for money; and (*ii*) explore multi-criteria recommendation using both textual reviews and multi-criteria ratings.

ACKNOWLEDGEMENTS

This work was partially financed by: (*i*) the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalisation - COMPETE Programme - within project ≪FCOMP-01-0202-FEDER-023151≫ and project ≪POCI-01-0145-FEDER-006961≫, and by National Funds through Fundação para a Ciência e a Tecnologia (FCT) - Portuguese Foundation for Science and Technology - as part of project UID/EEA/50014/2013; and (*ii*) ICT COST Action IC1406 High-Performance Modelling and Simulation for Big Data Applications (cHiPSet).

REFERENCES

Adomavicius, G. and Kwon, Y.: 2007, New recommendation techniques for multicriteria rating systems, *IEEE Intelligent Systems* **22**(3).

Adomavicius, G. and Kwon, Y.: 2015, Multi-criteria recommender systems, *Recommender systems handbook*, Springer, pp. 847–880.

Basu, C., Hirsh, H., Cohen, W. et al.: 1998, Recommendation as classification: Using social and content-based information in recommendation, *AAAI/IAAI*, pp. 714–720.

Bilge, A. and Kaleli, C.: 2014, A multi-criteria item-based collaborative filtering framework, *Computer Science and Software Engineering (JCSSE), 2014 11th International Joint Conference on*, IEEE, pp. 18–22.

Breese, J. S., Heckerman, D. and Kadie, C.: 1998, Empirical analysis of predictive algorithms for collaborative filtering, *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., pp. 43–52.

Ebadi, A. and Krzyzak, A.: 2016, A hybrid multi-criteria hotel recommender system using explicit and implicit feedbacks, *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering* **10**(8), 1377–1385.

Egger, R., Gula, I. and Walcher, D.: 2016, *Open Tourism: Open Innovation, Crowdsourcing and Co-Creation Challenging the Tourism Industry*, Springer.

Farokhi, N., Vahid, M., Nilashi, M. and Ibrahim, O.: 2016, A multi-criteria recommender system for tourism using fuzzy approach, *Journal of Soft Computing and Decision Support Systems* **3**(4), 19–29.

Friedman, A., Berkovsky, S. and Kaafar, M. A.: 2016, A differential privacy framework for matrix factorization recommender systems, *User Modeling and User-Adapted Interaction* **26**(5), 425–458.

Fuchs, M. and Zanker, M.: 2012, Multi-criteria ratings for recommender systems: an empirical analysis in the tourism domain, *International Conference on Electronic Commerce and Web Technologies*, Springer, pp. 100–111.

Herlocker, J. L., Konstan, J. A., Borchers, A. and Riedl, J.: 1999, An algorithmic framework for performing collaborative filtering, *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, ACM, pp. 230–237.

Hu, Y., Koren, Y. and Volinsky, C.: 2008, Collaborative filtering for implicit feedback datasets, *2008 Eighth IEEE International Conference on Data Mining*, Ieee, pp. 263–272.

Jannach, D., Gedikli, F., Karakaya, Z. and Juwig, O.: 2012, *Recommending Hotels based on Multi-Dimensional Customer Ratings*, Springer Vienna, Vienna, pp. 320–331.

Jhalani, T., Kant, V. and Dwivedi, P.: 2016, *A Linear Regression Approach to Multi-criteria Recommender System*, Springer International Publishing, Cham, pp. 235–243.

Law, R. and Chen, F.: 2000, Internet in travel and tourism-part ii: Expedia, *Journal of Travel & Tourism Marketing* **9**(4), 83–87.

Leal, F., Dias, J. M., Malheiro, B. and Burguillo, J. C.: 2016, Analysis and visualisation of crowd-sourced tourism data, *Proceedings of the Ninth International C* Conference on Computer Science & Software Engineering*, C3S2E '16, ACM, Porto, pp. 98–101.

Leal, F., Malheiro, B. and Burguillo, J. C.: 2017, Prediction and analysis of hotel ratings from crowd-sourced data, *in* A. Rocha, A. M. Correia, H. Adeli, S. Costanzo and L. P. Reis (eds), *New Advances in Information Systems and Technologies*, Springer International Publishing, Cham, pp. 1–10.

Lee, H.-H. and Teng, W.-G.: 2007, Incorporating multi-criteria ratings in recommendation systems, *Information Reuse and Integration, 2007. IRI 2007. IEEE International Conference on*, IEEE, pp. 273–278.

Liu, L., Mehandjiev, N. and Xu, D.-L.: 2011, Multi-criteria service recommendation based on user criteria preferences, *Proceedings of the fifth ACM conference on Recommender systems*, ACM, pp. 77–84.

Manouselis, N. and Costopoulou, C.: 2007, Analysis and classification of multi-criteria recommender systems, *World Wide Web* **10**(4), 415–441.

Nilashi, M., bin Ibrahim, O., Ithnin, N. and Sarmin, N. H.: 2015, A multi-criteria collaborative filtering recommender system for the tourism domain using expectation maximization (em) and pca–anfis, *Electronic Commerce Research and Applications* **14**(6), 542–562.

Randall Brandt, D.: 1988, How service marketers can identify value-enhancing service elements, *Journal of Services Marketing* **2**(3), 35–41.

Shambour, Q., Hourani, M. and Fraihat, S.: 2016, An item-based multi-criteria collaborative filtering algorithm for personalized recommender systems, *International Journal of Advanced Computer Science and Applications* **7**(8), 275–279.

Stone, M. and Brooks, R. J.: 1990, Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression, *Journal of the Royal Statistical Society* pp. 237–269.

Sykes, A. O.: 2000, An introduction to regression analysis, *in* E. A. Posner (ed.), *Chicago Lectures in Law and Economics*, Foundation Press, New York.

Tranmer, M. and Elliot, M.: 2008, Multiple linear regression, *Technical report*, The Cathie Marsh Centre for Census and Survey Research (CCSR).

Wang, H., Lu, Y. and Zhai, C.: 2010, Latent aspect rating analysis on review text data: a rating regression approach, *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 783–792.

Zhou, Y., Wilkinson, D., Schreiber, R. and Pan, R.: 2008, Large-scale parallel collaborative filtering for the netflix prize, *International Conference on Algorithmic Applications in Management*, Springer, pp. 337–348.

FÁTIMA LEAL holds a B.Sc. and an M.Sc. in Electrical and Computers Engineering (Major in Telecommunications) from the Instituto Superior de Engenharia do Porto, Portugal. She is currently enrolled in the "Information and Communication Technologies" Ph.D. programme at the University of Vigo, Spain as well as a researcher at INESC TEC, Porto. Her research, which is applied to crowd-sourced tourism data, is focussed on Trust and Reputation, Big Data and Context-aware Recommendation.

HORACIO GONZÁLEZ-VÉLEZ is an associate professor and head of the new Cloud Competency Centre at the National College of Ireland in Dublin. He spent over a decade working in engineering and product marketing for innovation-driven companies such as Silicon Graphics and Sun Microsystems. Award-winning lecturer and researcher, Horacio has also carried out applied research in parallel and distributed computing, funded by a number of public and industrial organisations including the European Commission, UK NESTA, NVidia, Dell, and Microsoft. He holds a Ph.D. in Informatics from the University of Edinburgh.

BENEDITA MALHEIRO holds a Ph.D. and an M.Sc. in electrical and computer engineering from the University of Porto (Faculty of Engineering). She is an adjunct professor at the Polytechnic Institute of Porto (School of Engineering) and a senior researcher at INESC TEC. Her research interests include multi-agent systems, conflict resolution, belief revision, personalisation, recommendation and context-aware systems.

JUAN CARLOS BURGUILLO holds an M.Sc. in telecommunications and a Ph.D. in telematics from the University of Vigo, Spain. He is currently an associate professor at the Department of Telematic Engineering at the same university. He has managed several R&D projects, and has published more than 100 papers in international journals and conferences. His research interests include multi-agent systems, evolutionary algorithms, game theory and telematic services.

Security Supportive Energy Aware Scheduling and Scaling for Cloud Environments

Agnieszka Jakóbik, Daniel Grzonka
Institute of Computer Science
Cracow University of Technology
Warszawska st 24, 31-155 Cracow, Poland

Joanna Kołodziej
Research and Academic Computer Network
(NASK)
Kolska st 12, 01-045 Warsaw, Poland

KEYWORDS

Computational clouds; Cloud computing; Tasks scheduling; Energy savings; Cloud services modelling; Cloud security; Genetic algorithm.

ABSTRACT

Energy consumption is one of the most important problems in the era of Computational Clouds (CC). CC infrastructures must be elastic and scalable for being accessible by huge population of users in different geographical locations. It means also that CC energy utilization systems must be modern and dynamic in order to reduce the cost of using the cloud services and resources.

In this paper, we develop the novel energy saving strategies for resource allocation and task scheduling in computational clouds. We present the new energy–aware scheduling policies and methods of scaling the virtual resources. The idea of the proposed models is based on Dynamic Voltage and Frequency Scaling (DVFS) techniques of modulation of the power of microprocessors. Additionally, the proposed model enables the monitoring of the energy consumption, which is necessary for providing the scheduling under the security criterion. The efficiency of the proposed models has been justified in the simple empirical analysis. The obtained results show the need to maintain a balance between energy consumption and task schedule execution.

I. INTRUDUCTION

The main idea of the efficient resource and service providing in computational clouds is based on the virtualization of the available resources. Typical cloud cluster is a multi-tenant environment, in which many virtual machines (VMs) may be implemented at the same physical computational server. The efficiency and performance of such VMs depend on the on virtualization policy and characteristics of the hardware, i.e. virtual disk configuration and allocation policy, the speed of the physical processor, etc. [18], [15].

Considering security and privacy, cloud system should deliver appropriate security operations for each uploaded task [21] or should deliver the tools for the administrators for building the security infrastructures [22].

In this paper, we define the novel energy optimization strategies for task scheduling in the cloud environment. The optimization of the consumption of the energy in computational clouds is achieved through specialized load balancing methods and scaling of the VMs. Additionally, model is used for monitoring the utilization of the energy consumed for processing the security operations. Based on the monitoring results, the users may take their decisions on security parameters (security levels) and configuration of VMs. For instance, the users may generate the large (long) or small (short) keys for cryptographic procedures. Such key scaling services are available in Amazon Cloud, RackSpace, OpenStack and Google Clouds ([1], [5], [4],[3]).

The paper is organized as follows. In section II, we present the methods for modelling and calculation energy consumption for the activated Virtual Machines. In Section III, we specify security policies and parameters related to the Independent Batch Scheduling in computational clouds. Section IV describes the methods of calculation total energy consumption, and scaling of VMs. In section V, the scheduling problem is defined as multiobjective optimization task with energy consumption and security as the major scheduling criteria. In Section (VI), the scenarios of the energy management in CC are presented. The proposed methodology is empirically evaluated through simple numerical tests in Section VII. The paper is summarized in Section VIII.

II. VM POWER CONSUMPTION

There are many methods of measurement and optimization of power and energy consumed by the VMs in clouds. Power consumption of a physical server in the cloud infrastructure can be easily measured by using the well known methods designed for microprocessors [27], [28], [20]. This Problem is different in the case of VMs [11], [30]. The power of the physical server which is necessary for the utilization of the VM allocated at this server (we will call it " VM power" in the rest of this paper) cannot be estimated by using just the hardware methods. Some specified technologies, such as Watts UP PRO Power or APIs enable the measurement of the power consumption of CPU, memory, IO, devices, disks and networks (see CloudWatch metrics service for Amazon Cloud [2]). The amount of energy consumed by the VMs in computational cloud depends on many factors. Most of them results from the virtualization itself, the implementation of VMs at

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

available servers and configuration of the physical infrastructure. Models of the energy utilization for VMs can be defined as modifications of the models for utilization of the physical resources, Such modifications are made by using the additional characteristics of the VMs architecture.

Let us denote by P_{Static} the power necessary for the preparation of the physical server for running the processes and implementation of the VMs (ready for work). Let $P_{Virtual}$ be the dynamic power consumed by VMs allocated at that server. The total power necessary for that physical server can be defined as follows [23]:

$$P_{Phys} = P_{Static} + \sum_{i=1,..,m} P(VM_i) = P_{Static} + P_{Virtual}, \quad (1)$$

where $P(VM_i)$ denotes the energy consumed by i-th instance of VM. Different methods are used to estimate this value. The non-observable variable $P(VM_i)$ is found based on observable P_{Phys}. Such methods are defined as relevant mathematical models with the most power-related resources as independent variables. The parameters of such models are estimated based on collected samples. The information may be collected by hyper-visor (black box method), or, in contrary, using white box method based on running proxy program on each VM [19].

In the model defined by Li, at al. in [31], the power consumption is calculated based on the utilization of CPU, operational memory and hard disk.

Bohra in [8] proposes a model that distinguish the baseline and active power consumption. In that paper, the independent variables are monitoring hardware performance of different Cloud system components and VMs energy consumption is estimated based on that measurements.

Krishnan in [30] used not only CPU utilization, but also memory consumption. Versick in [35] proposes a polynomial model. In this model network interface card (NIC) power consumption and hard disk power consumption is measured. The energy utilized by the VM depends on the above measurements.

Betran, at al. [7] used the linear model for the measurement of the VM power. The VM power consumption depends on 9 independent variables, such as activity of first level cache and number of accesses per cycle.

Most of the presented models are linear mathematical models with independent variables. Contrary, Gaussian Stochastic Mixture model is proposed in [12]. All of mentioned models are not good enough for the illustration of the realistic cloud virtual resource allocation and scheduling problems, [34].

There are some tools developed for the measurement of the virtual machine energy, such as FitGreen [13], Julemeter [26] or system proposed by Murwantara [32]. But those algorithms are not integrated with the cloud platforms. They need a special policy for the access to the cloud physical layer. Therefore, they may be implemented only from the cloud provider level or in the Infrastructure as a Service (IaaS) layer as a separate component.

Another approach is presented in [14], where the authors try to reduce energy consumption of applications running in cloud environments. The paper describes several deployment configurations based on queuing networks, and quantitative analysis for prediction of application performance and energy consumption.

Lot of attention was paid for the problem of the physical resources power consumption. D. Cerotti et al. [10] considered the issue of modelling power consumption in multicore CPUS with multithreading and frequency scaling. The authors present non-linear model for energy consumption that takes into account dynamic frequency scaling and Hyper-Threading, which have a significant impact on the model effectiveness.

This review shows that the Virtual Energy can be estimated. Additionally, the energy optimization methods presented so far are based on different methods comparing to those presented in this paper.

III. BATCH TASK SCHEDULING

In this paper, we considered the problem of Independent Batch Scheduling in computational clouds. We used Genetic Algorithm as the main mechanism for the cloud schedulers [29], [16], [17]. In this paper, we consider energy utilization as additional scheduling criterion. The main scheduling model is based on Expected Time to Compute (ETC) matrix, adopted to virtual machines (ETC_V). ETC_V matrix can be defined as follows:

$$ETC_V = [ETC_V[j][i]]_{j=1,\ldots,n;i=1,\ldots,m}, \quad (2)$$

where

$$ETC_V[j][i] = wl_j/cc_i, \quad (3)$$

in which cc_i is the computational capacity of i-th VM and wl_j is the workload of j-th task; n and m are respectively, number of tasks and number of VMs.

Based on the ETC_V matrix, we defined another $SBETC$ (Security Biased Expected Time to Compute) matrix which contains the additional security bias (SB) parameter b^i in order to reflect the security issues. Full description of this model can be found in [25].

The main objective of the scheduling is to find an optimal solution for specified criteria. The major objective for batch scheduling is the makespan that can be defined as follows:

$$C_{\max} = \min_{S \in Schedules} \left\{ \max_{j \in Tasks} C_j \right\}, \quad (4)$$

where C_j is the time when j-th task is finalized. $Tasks$ is the set of the tasks in the batch, and $Schedules$ is the set of all possible schedules, which can be generated for the tasks from that batch.

The batches of tasks are generated in non-deterministic time intervals. Therefore, each batch may have different number of tasks and the dimensions of the related SBETC matrices can be also various. The detailed description of that process is presented in [24] and [25].

The measure that may be used for energy optimization is energy efficiency. It is calculated for the particular schedule s in the following way:

$$E_{efficiency}(VM_i) = \frac{\sum_{j=1,...n,} wl_j \delta_{i,j}}{E(VM_i)} \quad (5)$$

where $\delta_{i,j}(s)$ is the binary factor. $\delta_{i,j}(s) = 0$ when the task number j is not scheduled for the machine i. $\delta_{i,j}(s) = 1$ when it is.

IV. Energy Calculation

A. Constant VM characteristics

Let E_{sec} denotes the energy necessary for running security operations. Then E_{total} is the total energy spend for particular batch of tasks. This energy is calculated for each schedule. In case of computational capacity of VM is constant and VMs are fully loaded during task running, only two energetic states are considered. They are: busy (100% computational power used for tasks calculations) and idle state. Let's assume that: t^i_{idle} - the time when i-th VM is idle; t^i_{busy} - the time when VM is calculating tasks; P^i_{idle} - the power necessary for VM to keep idle state; and P^i_{busy} - the power consumed by VM when is calculating tasks. The power necessary for security operations is assumed to be the same as in *busy* mode.

The above parameters are different for different schedules and can be defined as follows:

$$t^i_{busy} = max_{j \in Tasks\ scheduled\ \ for\ \ VM_i} C_j \quad (6)$$

$$t^i_{idle} = C_{max} - t^i_{busy} \quad (7)$$

$$t^i_{sec} = \sum_{j \in Tasks\ scheduled\ \ for\ \ VM_i} b^i_j \quad (8)$$

The overall energy may be expressed in the following way:

$$E_{total} = \sum_{i=1}^{m} \int_0^{C_{max}} Pow_{VM_i}(t)dt =$$
$$\sum_{i=1}^{m} (P^i_{idle} * t^i_{idle} + P^i_{busy} * (t^i_{busy} + t_{sec})) =$$
$$E_{task} + E_{sec} \quad (9)$$

B. VM scaling and re-provisioning

Very effective tool for saving energy in CC is VMs scaling. Scaling services allows to scale instances capacity up or down automatically or manually according to the user's needs. Cloud providers offers the following two main scaling methods:

- with re-provisioning: when scaling is done with change of allocation of virtual CPUs (vCPUs), memory, storage, or network resources. The scaled VM has to be shutdown (that lasts t_{close} sec.), new VM have to be configured and created (that lasts t_{open} sec.). It consumes relevant portion of system power.
- without re-provisioning, when VM is not reallocated, but only the computational power of it is changed. It is done by the change of capacity and redeploying VM from the template.

After each modification of the computational capacity parameter, the model (9) have to be updated. The computational capacity transition may be realized with or without re-provisioning of the VMs. In the second case, the time for the deactivation of a given VM and implementation of the new one and the related energy for that operations have to be considered. Also the dimensions of ETC_V and $SBETC$ matrices have to be changed.

There are two scenarios for VM scaling:

- scenario α: before calculating the schedule for the new batch, when all old tasks were executed and the system is idle, waiting for the next batch,. It may be done with re-provisioning ($\alpha 1$) or without ($\alpha 2$);
- scenario β: after calculating the schedule for the new batch, when workload is known to adapt to it. It may be executed with re-provisioning ($\beta 1$) or without ($\beta 2$).

Decision of the Cloud provider about choosing the number of VMs and their computational capacities for the next stage of system functioning may be based on different criteria and objectives. This problem is beyond the scope of this paper. The example of such decision process was presented in [36]. It was based on Stalkerberg games strategies. In this paper we assume that such decision was made earlier and the number of VMs and their computational capacities is set. Therefore the scheduling is based on:

- $\alpha 1$: when old VM has to be closed to open the new one

$$SBETC[j][i] =$$
$$w_j/cc_i + b(sd_j, w_j, tl_i, cc_i) + t^i_{close} + t^i_{open}, \quad (10)$$

$$E^{\alpha 1}_{total} = E_{total} + \sum_{i=1}^{m} (P^i_{close} t^i_{close} + P^i_{open} t^i_{open}); \quad (11)$$

is the energy consumed;

- $\alpha 2$: when old VM do not have to be to be closed but it needs rescaling only:

$$SBETC[j][i] = w_j/cc_i + b^i_j + t^i_{scale} \quad (12)$$

$$E^{\alpha 2}_{total} = E_{total} + \sum_{i=1}^{m} P^i_{scale} t^i_{scale}; \quad (13)$$

- $\beta 1$: the scheduling is made for

$$SBETC[j][i] = w_j/cc_i + b^i_j + t^i_{close} + t^i_{open}, \quad (14)$$

but the tasks are done according to the new re-provisioned computational capacity $\overline{cc_i}$, therefore $E^{\beta 1}_{total}$ consists elements related to the new version of VM.

- $\beta 2$: the schedule is calculated for

$$SBETC[j][i] = w_j/cc_i + b^i_j + t^i_{scale}, \quad (15)$$

but the tasks are run according to the re-scaled computational capacity: $\overline{cc_i}$, therefore $E^{\beta 1}_{total}$ also includes values for re-scaled version of VM.

In all above cases, the number of tasks in the batch may be of the range $i = 1, 2, \ldots, \overline{m}$ and the number of VMs may be of the range $j = 1, 2, \ldots, \overline{n}$.

V. Energy Aware Scheduling Objectives

The problem of finding the schedule that minimizes the makespan with constant computational capacities may be written in the form:

$$argmin_{s \in Schedules} \sum_{i=1,..,m \quad j=1,...,n} (\frac{wl_j}{cc_i} + b^i)\delta_{i,j}(s) \tag{16}$$

The problem of finding the schedule that minimizes the total energy in that case may be formulated as:

$$argmin_{s \in Schedules} \sum_{i=1,...,m} (\sum_{j=1,\delta_{i,j}(s)=1}^{n} P^i_{busy}(\frac{wl_j}{cc_i} + b^i) + \sum_{j=1,\delta_{i,j}(s)=0}^{n} P^i_{idle} t^i_{idle}) \tag{17}$$

For cases α and β the proper equations are constructed in the similar way, taking into account relevant energy levels given in section IV.

The energy consumption may also be considered as a complementary scheduling criterion together with the makespan as the main objective.

We may also be interested in finding the rate of energy spend on the security operation to energy spend on bare task calculation: E_{task}/E_{sec}.

The usage of the SBECT matrix enables to test different energy savings strategies by lowering or rising the trust level of VM. Furthermore, the complex simulation may be performed before real cloud environment modifications.

VI. Energy Saving Scenarios

A. Strategies for scheduler

We proposed four concurrent models for monitoring energy and makespan during the scheduling process. They reflect the importance of short time of tasks calculation and energy savings.

1. Makespan based scheduling and monitoring of the energy. In any case when two schedules have the same (or close) makespan, the scheduler chooses that one with smaller energy level. This case is suitable for the situation when the makespan is the priority. We want to save the energy not compromising the makespan.

2. Energy based scheduling and monitoring of the makespan. When the two schedules have the same (or close) energy level, the scheduler chooses that with smaller makespan. This case will be executed when we would like to calculate energy efficient schedules and we may afford to wait for our tasks longer.

3. Makespan based scheduling until the desired level is reached, then tasks are scheduled according to the energy objective. The search is performed only among the schedules that has the desired or smaller makespan.

4. Energy based scheduling until the desired level is reached, then tasks are scheduled according to the makespan objective. We are looking among the schedules that has the desired or smaller energy level.

The reference model is the makespan based scheduling only.

B. Strategies for VM scaling

Each VM may rated according to the energy efficiency, see eq. 5. This monitoring supports the decision making process about cancelling particular VM and replacing it with more effective one.

VM scaling may be done according to the $\alpha 1$, $\alpha 2$, $\beta 1$ or $\beta 1$ models.

For example, the new computational capacities after the schedule was calculated but before the tasks are executed, ($\beta 1$), may be found by solving the following problem:

$$C_{max} - t^i_{scale} - \overline{t}^i_{scale} = \frac{1}{\overline{cc}_i}(\sum_{j=1,\delta_{ij}(s)=1}^{n} wl_j + \overline{b}^i_j) + \overline{t}^i_{idle} \tag{18}$$

where the 'bar' values are computed for the new configuration of the environment. The new computational capacity for the machine is accepted if is it profitable according to the energy expenditure. The scheduling problem may be NP-complete due to many tasks [29]. The problem of finding the particular VM is not so time consuming. Cloud providers offers limited versions of instances.

VII. Evaluation of Developed Models

A. Tests of energy aware scheduling vs makespan scheduling

The aim of the test was to examine the makespan and energy consumption for the schedules that was calculated using different criteria.

The tests were evaluated using platform for simulation the Cloud environments called SimGrid [33].

Tested strategies are described in sec. VI-A Five types of VMs were assumed (see tab.I). The SURF component was used to simulate the execution of activities on resources [9]. For simulating the VM starting, running, scaling and closing the *pstates* were used.They allow to declare power states when the VM is switched off, the idle state power consumption and energy necessary for fully loaded VM. In case of Frequency scaling of the physical CPU of the VM the linear model in between fully load and idle state is assumed. The $watt^{state}_{per}$ function was used to measure the power consumption of the VM in each state. The simulator enables also to get current speed (in FLOPS) for each energetic state, total energy consumed so far (see tab. I). Additionally, the number of tasks to distribute, the computation size of each task, the size of the files associated to each task and a list of VMs that will accept those tasks may be specified. The scheduling algorithm was implemented using C++. Optimization module for solving problems 16 and 17 was implemented in MATLAB programming environment.

TABLE I: Characteristics of VMs declared in SimGrid used for simulation

VM	Speed	Energetic profile
number	GFLOPS	min:max in Watts
1	0.02	90:105
2	0.05	93:110
3	0.1	100:120
4	0.2	150:170
5	0.3	200:230

TABLE II: Measured energy consumption for benchmark task where $wl = 10$ GFLOPS and total simulation time is equal 510 seconds.

VM	After sleep for 10 sec.	After exec. 10 Gflops	Exec. time	Total energy
nr	Joules	Joules	Seconds	Joules
1	900	53400	500	53400
2	930	22930	200	50830
3	1000	13000	100	53000
4	1500	10000	50	77500
5	2000	9666.66	33.34	103000

The computational capacities of VMs were measured by the benchmark test (see tab. II). They were computed using execution time of benchmark task having workload of 10 GFLOPS and was stated to be equal to the declared VMs speed (see tab. III). The P^i_{idle} and P^i_{busy} per second were calculated using energy consumed in benchmark execution time. Idle machine was simulated using sleep mode.

TABLE III: Calculated VMs characteristics

cc_i	P^i_{idle}	P^i_{busy}	E_{eff}	P^i_{open}	P^i_{close}	P^i_{scale}
0.02	90	106.8	0.09	63	27	54
0.05	93	114.6	0.08	65	28	56
0.1	100	130	0.07	70	30	60
0.2	150	200	0.05	105	45	90
0.3	200	290	0.03	140	60	120

The time of VM opening was assumed to be the same for all VMs: $t^{scale} = t_{close} = t_{open} = 1sec$. Additionally, the environment was modelled so that $P^i_{open} = 70\% P^i_{idle}$, $P^i_{close} = 30\% P^i_{idle}$, $P^i_{scale} = 60\% P^i_{idle}$.

Fist test considered very simple VM loading of 5 tasks scheduled for 5 machines. Testing workloads were: $wl_1 = 1, wl_2 = wl_3 = wl_4 = 2, wl_5 = 8GFLOPS$. Moreover, tasks having the same workload are indistinguishable and may be computed interchangeably. For such a case direct search of best schedules was implemented. Scheduling according to the makespan objective (see model 1) resulted in the makespan=50 sec. For that schedule, energy consumed by VMs was 258004 Watts. The resulted mapping of tasks for consecutive VMs (1 to 5, according to the table I) was: 1, 2, 2, 2, 8. This mapping was denoted as schedule 1. Scheduling according to the energy objective (see model 2) resulted in the schedule with energy consumption equal 257596 Watts. This schedule makespan was 160 sec. The resulted mapping of tasks for consecutive VMS was: 1, 8, 2, 2, 2. They formulated schedule 2.

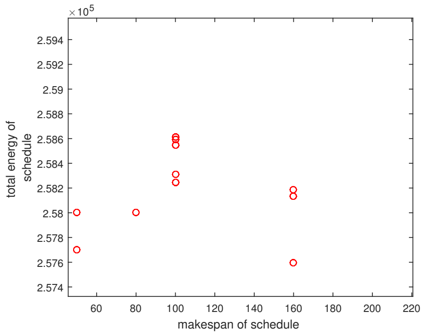


Fig. 1: Energy consumed for different schedules

Fig. 1 shows that for the same makespan there is possible to find schedules consuming less energy. Conversely, for the same energy there are several schedules having more or less beneficial makespan. Therefore finding suboptimal (or optimal) solutions according to the models 3 and 4 is possible and may be profitable.

B. Tests of strategies for VM scaling

The aim of these tests was to examine the influence of proper re-scaling of VMs on the energy consumption. Tested strategies are described in sec. VI-B. For these chosen schedules presented in the previous subsection, the re-scaling strategy $\beta 2$ was incorporated. For schedule 1. only the first VM was busy for the whole makespan time, see tab. IV. Therefore MVs 2, 3 and 4 was examined for the possibility off re-scaling, see tab. V. The most beneficial re-scaling was: VM 3 scaled VM 2, VM 4 scaled VM 2 and VM 5 scaled VM 4. It enabled to save energy, keeping makespan of batch at the same level. For schedule 2. second machine was not considered for re-scaling because is not idle. First VM machine was also excluded, because there is no VM with smaller energy consumption that VM number 1. The most beneficial re-scaling was: VM 3 scaled VM 2, VM 4 scaled VM 2 and VM 5 scaled VM 1.The example result of scaling in tern of idle time of the machine and best schedule 2. is the following:

- before re-scaling the VM 5 was busy for 6.6 sec consuming 290 Watt/sec. and was idle for 200 sec. consuming 200 Watts/sec.
- after re-scaling the VM 5 into VM 4, it was busy for 10 sec consuming 200 Watt/sec. and was idle for 149 sec. consuming 150 Watts/sec., additionally 1 sec. and 120 Watt was consumed into re-scaling operation

Incorporating scaling scenarios when schedule is ready is very beneficial. The scaled VMs will do the assigned work in the same makespan, but the energy consumption is much lower.

TABLE IV: VMs loading for the chosen optimal schedules

Tasks	Idle time for consecutive VMs	Working time for consecutive VMs
Schedule 1. 1, 2, 2, 2, 8	0,10,30,40,23.4	50,40,20,10,26.6
Schedule 2. 1, 8, 2, 2, 2	110,0,140,150,153.4	50,160,20,10,6.6

TABLE V: VMs power consumption and time of tasks running time, before and after scaling for generated best suboptimal schedules: best schedule 1: 1,2,2,2,8 and best schedule 2: 1,8,2,2,2. Only the beneficial scaling possibilities are presented.

Schedule 1 old VM → new VM	Energy before scaling	Energy after scaling
3 → 2	5600	5481
4 → 3	8000	5590
4 → 2	8000	5511
5 → 4	12394	9470
Schedule 2		
3 → 2	16600	15711
3 → 1	16600	16050
4 → 3	24500	16590
4 → 2	24500	15741
4 → 1	24500	16040
5 → 4	32594	24470
5 → 3	32594	23570
5 → 2	32594	15747
5 → 1	32594	12510

C. Tests of energy aware scheduler with advanced workload

Scheduling many tasks require using advanced method for searching the optimal schedule. We implemented Genetic Algorithm (GA) for finding the schedules given by strategies presented in sec. VI. Four strategies for scheduler are described in sec VI-A.

The models were implemented in C++. Due to the fact that makespan and energy values are *double* type, the 3% difference between values was chosen as the increment distinguishing one type of *double* real value from the second one.

The model was tested for set of 20 VMs, see tab. VI. These characteristics were obtained by test on real infrastructure (see [6]). Each batch consists 200 tasks. Every run of the GA was tested for 20 set of parameters configurations. The single gene represents the schedule. In our approach, the final result is the whole population that represents the suboptimal schedule (see: 2. Crossover operation means swapping tasks among virtual resources from respectively, set with best and worse fitted individuals. Mutation operation was not used. During each epoch the population is changed, but the size of population remains the same.

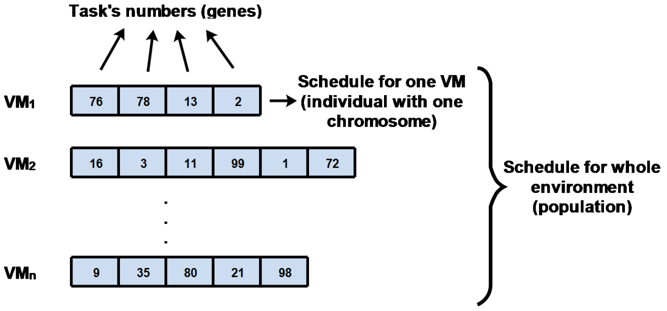

Task's numbers (genes)
VM1
76
78
13
2
Schedule for one VM (individual with one chromosome)
VM2
16
3
11
99
1
72
VMn
9
35
80
21
98
Schedule for whole environment (population)

Fig. 2: Mapping tasks into genes

TABLE VI: Tested VMs

i	cc_i	P^i_{idle}	P^i_{busy}	i	cc_i	P^i_{idle}	P^i_{busy}
1	0.2	71.9	57.52	2	0.7	75.4	60.32
3	1.0	74.5	59.6	4	1.9	82.0	65.6
5	1.7	78.6	62.88	6	2.4	71.0	56.8
7	10.6	73.0	58.4	8	62.9	128.7	102.96
9	71.8	124.7	99.76	10	50.4	122.5	98.0
11	57.8	123.6	98.88	12	0.4	55.9	44.72
13	2.2	60.1	48.08	14	2.7	60.4	48.32
15	4.2	62.7	50.16	16	4.3	62.5	50.0
17	9.8	60.6	48.48	18	47.7	64.9	51.9
19	1.71	17.1	13.68	20	1.73	17.4	13.92

After some iterations, average individuals are frozen and tasks are assigned to the recourses permanently (see: fig. 3). Therefore the number of swapped tasks is decreasing. The best and worst individuals are chosen to be crossovered. During each epoch the new population is created and evaluated. The fitness function for the GA was assumed to be equal the makespan or total energy consumed per schedule.

Exchanging two tasks according the makespan influences the energy consumption. This is due to the fact that makespan depends only on the tasks scheduled on the machine that needs the longest tasks to complete the work. Energy consumption differs for the schedules having the same makespan, but different task distribu-

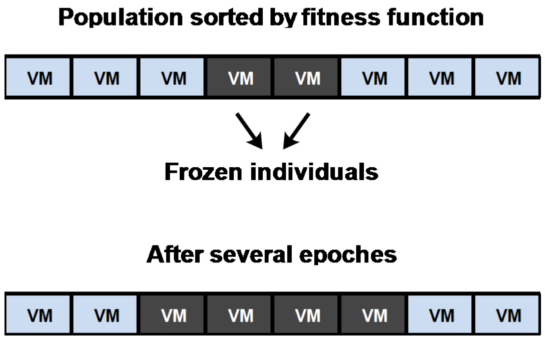


Fig. 3: Excluding average individuals from exchanging tasks (freezing)

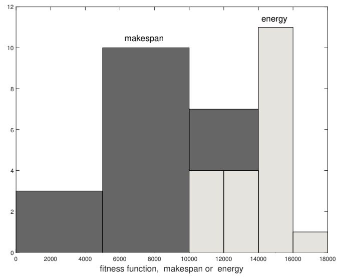


Fig. 4: Histogram of generated fitness function for scheduling according to the makespan and total energy

tion. It depends on the working (busy) and idle time of all the recourses in the system. Therefore, larger number of crossovers modifies the value of total energy fitness GA function. GA was also more sensitive for scheduling according to the energy objective. Fig. 4 presents the value of fitness function for 20 GA runs after 2000 epochs. Incorporating energy objective after initial stage of makespan scheduling (model 3) results in both smaller mean makespan, and lowering energy consumption of the system, comparing to the scheduling based on energy criterion only (model 2). The results of testing energy saving scenarios 1 - 4 are presented in the tab. VII.

The best result for the proposed GA was obtained where scheduling was done according to the energy criterion first, and makespan criterion as second (model 4).

TABLE VII: Mean values of makespan and total energy for 20 initial populations of GA running, tested 4 scenarios (see sec. VI-A)

Model	Mean makespan	Mean total energy
1	8732.2 sec.	86149.3 W
2	13844.3 sec.	136546 W
3	13337.4 sec.	118166 W
4	7345.5 sec.	76902.7 W

The experimental results show, that the crossover for scheduling according to the energy fitness function is less effective. In this scenario we are exchanging tasks between worst and best individuals. There is a need for formulation another criterion for finding better populations offspring.

VIII. Summary

In this paper we developed and implemented a new model of energy and security aware Independent Batch Scheduler. We defined four scenarios for monitoring of the energy utilization and makespan during scheduling process. We also developed four models for scaling VMs in order reduce the energy consumption. Additionally, we presented the short overview of methods for estimation energy consumption in virtualized environments, and existing energy scaling methods for VMs.

The experimental results presented in the paper demonstrate and confirm the effectiveness of proposed models. The best result for scheduling of advanced workload was obtained for the last model, where the major scheduling objective was energy and the second makespan.

In the future, we would like to avoid the work on the cloud simulators and implement our models on the realistic CC platforms working with the OpenStack standards.

Acknowledgement

This article is based upon work from COST Action IC1406 "High-Performance Modelling and Simulation for Big Data Applications" (cHiPSet), supported by COST (European Cooperation in Science and Technology)

References

[1] Amazon Cloud Scaling Service.
[2] Amazon CloudWatch.
[3] Google Cloud Scaling Service.
[4] OpenStack Cloud Scaling Service.
[5] Rackspace Cloud Scaling Service.
[6] P. Benner, P. Ezzatti, E. Quintana-Ortí, and A. Remón. *On the Impact of Optimization on the Time-Power-Energy Balance of Dense Linear Algebra Factorizations*, pages 3–10. Springer International Publishing, 2013.
[7] R. Bertran, Y. Becerra, D. Carrera, V. Beltran, M. Gonzalez, X. Martorell, J. Torres, and E. Ayguade. Accurate energy accounting for shared virtualized environments using pmc-based power modeling techniques. In *2010 11th IEEE/ACM International Conference on Grid Computing*, pages 1–8, Oct 2010.
[8] A. E. H. Bohra and V. Chaudhary. Vmeter: Power modelling for virtualized clouds. In *2010 IEEE International Symposium on Parallel Distributed Processing, Workshops and Phd Forum (IPDPSW)*, pages 1–8, April 2010.
[9] H. Casanova, A. Giersch, A. Legrand, M. Quinson, and F. Suter. Versatile, scalable, and accurate simulation of distributed applications and platforms. *Journal of Parallel and Distributed Computing*, 74(10):2899 – 2917, 2014.
[10] D. Cerotti, M. Gribaudo, P. Piazzolla, R. Pinciroli, and G. Serazzi. *Modeling Power Consumption in Multicore CPUs with Multithreading andFrequency Scaling*, pages 81–90. Springer International Publishing, Cham, 2016.
[11] M. Colmant, M. Kurpicz, P. Felber, L. Huertas, R. Rouvoy, and A. Sobe. Process-level power estimation in vm-based systems. In *Proceedings of the Tenth European Conference on Computer Systems*, EuroSys '15, pages 14:1–14:14, New York, NY, USA, 2015. ACM.
[12] G. Dhiman, K. Mihic, and T. Rosing. A system for online power prediction in virtualized environments using gaussian mixture models. In *Design Automation Conference*, pages 807–812, June 2010.
[13] C. Dupont, T. Schulze, G. Giuliani, A. Somov, and F. Hermenier. An energy aware framework for virtual machine placement in cloud federated data centres. In *2012 Third International Conference on Future Systems: Where Energy, Computing and Communication Meet (e-Energy)*, pages 1–10, May 2012.
[14] M. Gribaudo, T. T. N. Ho, B. Pernici, and G. Serazzi. *Analysis of the Influence of Application Deployment on Energy Consumption*, pages 87–101. Springer International Publishing, 2015.
[15] D. Grzonka. The Analysis of OpenStack Cloud Computing Platform: Features and Performance. *Journal of Telecommunications and Information Technology*, 3:52–57, 2015.
[16] D. Grzonka, J. Kołodziej, and J. Tao. Using artificial neural network for monitoring and supporting the grid scheduler performance. In *28th European Conference on Modelling and Simulation, ECMS 2014, Brescia, Italy, May 27-30, 2014*, pages 515–522, 2014.
[17] D. Grzonka, J. Kołodziej, J. Tao, and S. U. Khan. Artificial neural network support to monitoring of the evolution-

ary driven security aware scheduling in computational distributed environments. *Future Generation Computer Systems*, 51:72–86, 2015.

[18] D. Grzonka, M. Szczygiel, A. Bernasiewicz, A. Wilczynski, and M. Liszka. Short analysis of implementation and resource utilization for the openstack cloud computing platform. In *29th European Conference on Modelling and Simulation, ECMS 2015, Albena (Varna), Bulgaria, May 26-29, 2015. Proceedings.*, pages 608–614, 2015.

[19] C. Gu, H. Huang, and X. Jia. Power metering for virtual machine in cloud computing-challenges and opportunities. *IEEE Access*, 2:1106–1116, 2014.

[20] T. W. Harton, C. Walker, and M. O'Sullivan. Towards power consumption modeling for servers at scale. In *2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC)*, pages 315–321, Dec 2015.

[21] A. Jakobik. *Big Data Security*, pages 241–261. Springer International Publishing, Cham, 2016.

[22] A. Jakobik. A cloud-aided group rsa scheme in java 8 environment and openstack software. *Journal of Telecommunications and Information Technology : JTIT*, (2):53–59, 2016.

[23] A. Jakóbik, D. Grzonka, J. Kołodziej, A. E. Chis, and H. Gonzalez-Velez. Energy Efficient Scheduling Methods for Computational Grids and Clouds. *Journal of Telecommunications and Information Technology*, 1, 2017.

[24] A. Jakobik, D. Grzonka, J. Kolodziej, and H. González-Vélez. Towards secure non-deterministic meta-scheduling for clouds. In *30th European Conference on Modelling and Simulation, ECMS 2016, Regensburg, Germany, May 31 - June 3, 2016, Proceedings.*, pages 596–602, 2016.

[25] A. Jakobik, D. Grzonka, and F. Palmieri. Non-deterministic security driven meta scheduler for distributed cloud organizations. *Simulation Modelling Practice and Theory*, in press. (available online 4 November 2016).

[26] A. Kansal, F. Zhao, J. Liu, N. Kothari, and A. A. Bhattacharya. Virtual machine power metering and provisioning. In *Proceedings of the 1st ACM Symposium on Cloud Computing*, SoCC '10, pages 39–50, New York, NY, USA, 2010. ACM.

[27] H. Kataoka, D. Duolikun, T. Enokido, and M. Takizawa. Power consumption and computation models of a server with a multi-core cpu and experiments. In *2015 IEEE 29th International Conference on Advanced Information Networking and Applications Workshops*, pages 217–222, March 2015.

[28] H. Kataoka, A. Sawada, D. Duolikun, T. Enokido, and M. Takizawa. Energy-aware server selection algorithms in a scalable cluster. In *2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, pages 565–572, March 2016.

[29] J. Kołodziej. *Evolutionary Hierarchical Multi-Criteria Metaheuristics for Scheduling in Large-Scale Grid Systems.* Springer Publishing Company, Incorporated, 2012.

[30] B. Krishnan, H. Amur, A. Gavrilovska, and K. Schwan. Vm power metering: Feasibility and challenges. *SIGMETRICS Perform. Eval. Rev.*, 38(3):56–60, Jan. 2010.

[31] Y. Li, Y. Wang, B. Yin, and L. Guan. An online power metering model for cloud environment. In *2012 IEEE 11th International Symposium on Network Computing and Applications*, pages 175–180, 2012.

[32] I. M. Murwantara and B. Bordbar. A simplified method of measurement of energy consumption in cloud and virtualized environment. In *Proceedings of the 2014 IEEE Fourth International Conference on Big Data and Cloud Computing*, BDCLOUD '14, pages 654–661, Washington, DC, USA, 2014. IEEE Computer Society.

[33] M. Quinson. Simgrid: A generic framework for large-scale distributed experiments. In *2009 IEEE Ninth International Conference on Peer-to-Peer Computing*, pages 95–96, Sept 2009.

[34] J. Read. What is an ECU? CPU Benchmarking in the Cloud.

[35] I. Wassmann, D. Versick, and D. Tavangarian. Energy consumption estimation of virtual machines. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, SAC '13, pages 1151–1156, New York, NY, USA, 2013. ACM.

[36] A. Wilczyński and A. Jakóbik. Using Polymatrix Extensive Stackelberg Games in Security–Aware Resource Allocation and Task Scheduling in Computational Clouds. *Journal of Telecommunications and Information Technology*, 1, 2017.

AUTHOR BIOGRAPHIES

AGNIESZKA JAKÓBIK (KROK) received her M.Sc. in the field of stochastic processes at the Jagiellonian University, Cracow, Poland and Ph.D. degree in the field of neural networks at Tadeusz Kosciuszko Cracow University of Technology, Poland, in 2003 and 2007, respectively. From 2009 she is an Assistant Professor. Her e-mail address is: agneskrok@gmail.com

DANIEL GRZONKA received his B.Sc. and M.Sc. degrees with distinctions in Computer Science at Cracow University of Technology, Poland, in 2012 and 2013, respectively. Currently, he is Research and Teaching Assistant at Cracow University of Technology and Ph.D. student at Jagiellonian University in cooperation with Polish Academy of Sciences. He is also a member of Polish Information Processing Society and IPC member of several international conferences. His e-mail address is: grzonka.daniel@gmail.com. For more information please visit: www.grzonka.eu

JOANNA KOŁODZIEJ is an associate professor in Research and Academic Computer Network (NASK) Institute and Department of Computer Science of Cracow University of Technology. She is a vice Head of the Department for Sciences and Development. She serves also as the President of the Polish Chapter of IEEE Computational Intelligence Society. She is also a Honorary Chair of the HiPMoS track of ECMS. Her e-mail address is: joanna.kolodziej68@gmail.com. The detailed information is available at www.joannakolodziej.org

A Low-cost Distributed IoT-based Augmented Reality Interactive Simulator for Team Training

Pietro Piazzolla[1], Marco Gribaudo[1], Simone Colombo[2], Davide Manca[2], and Mauro Iacono[3]

[1]Dipartimento di Elettronica, Informazione e Bioingegneria
[2]CMIC Chemical Engineering Department
Politecnico di Milano, P.zza Leonardo da Vinci 32, 20133 Milano, Italy
{pietro.piazzolla, marco.gribaudo, simone.colombo, davide.manca}@polimi.it
[3]Dipartimento di Matematica e Fisica, Università degli Studi della Campania "L. Vanvitelli"
Viale Lincoln 5, 81100 Caserta, Italy
mauro.iacono@unina2.it

KEYWORDS

Distributed simulation, Real time, Virtual reality, Education, Immersive training, Firefighting

ABSTRACT

The performance over cost ratio of last generation off the shelf devices enables the design of heterogeneous distributed computing systems capable of supporting the implementation of an immersive Virtual Reality, Internet of Things based training support architecture.

In this paper we present our work in progress on a low cost distributed immersive simulation system for the training of teams by means of Virtual Reality and off the shelf mobile and prototyping devices. In this case, performance prediction is crucial, because the generation of the scenario have to be performed in real time and synchronization problems may disrupt the result.

The approach is demonstrated by a prototypical case study, that consists in a distributed simulator for the interactive training of groups of people that have to coordinate to face a fire emergency, and features advanced immersivity thanks to CGI-enabled stereoscopic 360 degree 3D Virtual Reality and ad-hoc devised interaction interfaces. In particular, we focus on the subsystem that is related to a single trainee, providing a reference implementation and a performance evaluation oriented model to support the design of the complete system.

INTRODUCTION

The performance-over-cost ratio of last generation off-the-shelf devices enables the design and implementation of cost effective complex interactive cyberphysical systems for non-critical applications. The consumer market offers: powerful smartphones with many cores processors, gigabytes of RAM and native power management and communication features; low cost augmented reality or virtual reality apparels; customizable prototyping oriented computer systems, such as Raspberry Pi and Arduino; Internet of Things (IoT) enabled smart location and motion aware sensors. Conversely, the most of these devices should not be considered reliable, due to the fact that they are meant to be sold on the consumer or makers market.

This equipment is anyway a resource to build heterogeneous distributed computing systems capable of supporting the implementation of non critical applications oriented to interactive simulation in a real environment, integrating a sensor part, by means of the IoT enabled components, an Virtual Reality (VR) part, supported by smartphones or prototyping platforms, and a coordination part, that is implemented by means of a peer to peer or server based distributed layer. A possible application is the interactive training of groups of people that have to coordinate to face an emergency in a given fictional location or on the field. In such a scenario the coordination part also includes a non trivial distributed workload that should generate in real time the VR elements: this implies a local generation of the point of view of a single user and a global management of the virtual objects and what is connected to the positions and the actions of all other users. Although such workloads are not a problem for modern high performance computer systems, the nature of the target reference architecture, the available computing power, the low reliability of common consumer devices, the interconnection problems and the real time requirements make the design of the system non trivial.

We propose an immersive VR IoT based training support architecture, based on off the shelf components, for low cost real time implementations. In this paper we focus on the subsystem that is related to a single trainee, providing a reference implementation and a performance evaluation oriented model to support the design of the system, with reference to a fire extinguisher use training case study.

This paper is organized as follows: in the next Section we present a quick survey of related literature, highlighting the main features that characterize our single-trainee subsystem compared to other similar solutions, proposed in the literature or currently on market. In the subsequent Section the technical details of our proposed simulator are presented, focusing on both the hardware both the software solutions adopted.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Then, this single-trainee subsystem is placed inside the wider perspective of an IoT collaborative system, which is described into details. Finally, a performance evaluation model of our system for the support of the design process is discussed.

RELATED WORKS

The use of virtual reality for training and learning has been largely explored by many papers (See e.g.:[1]). In this section we highlight the main features that characterize our simulator compared to other similar solutions, proposed in the literature or currently on market. While all solutions strove to provide the best real time 3D immersive computer graphics available at the time of their development, not all of them concerned about maximizing visual quality while containing the final application cost, like our proposed simulator. Other main differences between these products and our solution: the type of sensors used to detect user gestures and inputs (from ad-hoc devised controllers to on-market solutions like Microsoft Kinect II or Nintendo Wii controllers), and the hardware adopted to visually display the 3D environment (from Cave systems to Head Mounted Displays like Oculus Rift or Google Daydream). Recently in [2], the authors propose firefighting scenarios based training system on virtual reality platform. Their simulator software engine runs on a dedicated machine instead of, as in our solution, on smartphones. Our solution tend to be more cost effective with a lower physical space requirement to deploy the simulator. In [3] the authors aim for a solution that use software and hardware already on market to implement their simulator. The final product leverages on a cave system for providing immersivity, a solution which may not be widely adopted due to space constraints. One of the most comprehensive works on virtual reality for fire extinguisher use training is [4]. The author implement a simulator relying on two infrared cameras for tracking user gestures. Empirical evidence taught us that camera tracking alone (i.e.: not coupled with other sensors) may tend to be of limited effectiveness in some circumstances when the process of marker detection introduces delays.

The several on-market solutions can be sorted in two big categories: those using physical world video footage to present the virtual environment for training, like e.g.: [5], [6], [7], [8] and those leveraging on 3D computer graphics to generate it, like e.g.: [9], [10]. The latter solutions are intended for firefighters training, not just for the training of fire extinguishers.

About Internet-Of-Things main aspects and challenges: [11] offers a comprehensive review of the functional complexity of connecting real-world objects to the Internet with tiny sensors, while in [12] a Cloud centric vision for worldwide implementation of Internet of Things is presented and the key enabling technologies are discussed.

ARCHITECTURE OF THE SIMULATION SYSTEM

The system is composed of personal nodes, that collaborate on a peer to peer basis to enact the overall real time simulation. The hardware architecture of each node includes a smartphone, a device coordinator, a smart sensing subsystem and a AR/VR subsystem. The device coordinator, that may be implemented as an Arduino system as in our case study or by another analogous technology, is in charge of managing the smart sensing subsystem, that is composed of one or more smart sensors; the smartphone is in charge of running the distributed simulation and of coordinating with the other nodes, eventually managing dependability issues of the interconnection, by using the services that are provided by the device coordinator; according to the complexity of the AR/VR workload, either the smartphone or the device coordinator is in charge of piloting the AR/VR device. In Figure 1 the the software architecture of the system is presented. It is structured into 4 layers. The bottom layer is the data acquisition and management layer, that is distributed among the smart sensors and the device manager: smart sensors provide a first level of data processing, while the device manager provides a post processing that is based on a synthesis of data from smart sensors. The second layer is node management and coordination: it is executed on the smartphone and it is responsible of running a node and providing the integration into the distributed system. The third level is the simulation services layer: it runs the distributed simulation and generates and maintains the internal abstract description of the environment and of the participants and all needed synchronization. The fourth layer is the AR/VR layer, and is responsible of generating what needed for visualization (eventually supporting, besides the devices of the AR/VR subsystem, external monitors).

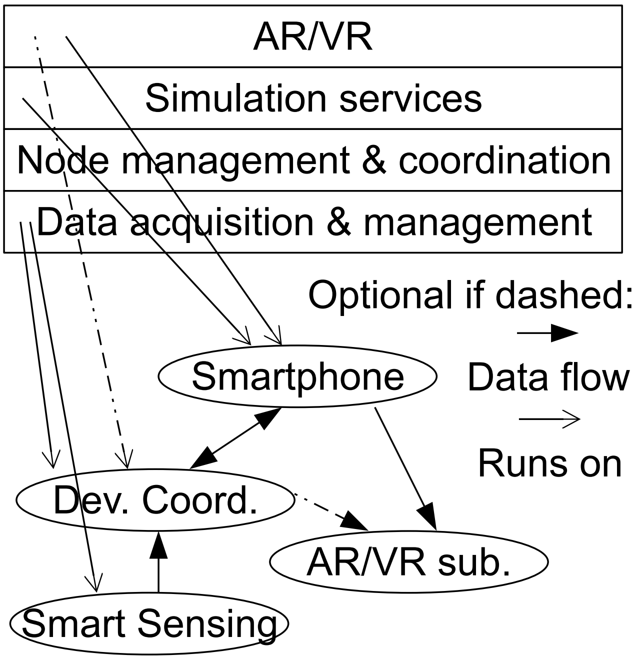


Fig. 1. The software architecture of the simulation system

IMPLEMENTATION

In this Section we present the technical details of a node of the proposed simulation system. Its main goal is to allow trainees to learn the correct use of a fire extinguisher in a 3D

Virtual Reality (VR) environment. Trainees benefit of the use of a virtual environment in avoiding them all the risks and costs associated with physical world training.

A VR head-mounted display (VR HMD) enable an immersive, stereoscopic, 360 degree view of the fire emergency scenario, while a specifically developed interface system that can be mounted on a physical fire extinguisher manage the interaction between the user-trainee and the virtual simulation.

As seen, from a technical perspective, each node of the simulation system is composed by a 3D *software application* containing all the simulator's logic and its 3d graphic engine, and a *hardware system* that is in charge of handling a trainee's input gestures and providing them to the application. In Figure 2 we present the overall structure of the real implementation of a node.

Node hardware

The core element of the hardware system is an *Arduino Leonardo*[13], a microcontroller board that appears to a connected computer as a mouse and keyboard. It can be easily programmed using a dedicated programming language, in order to adapt its behavior to users needs. Connected to the Arduino Leonardo input pins there are two devices: an *accelerometer* with 9 degrees of freedom and an *analog control stick*. In particular, we use the first one to determine the direction of the fire extinguisher nozzle and the latter for handling the movements of the trainee in the simulated environment. For the purposes of our simulator prototype development, for the accelerometer component, we use a SparkFun 9DOF Sensor Stick [14], a very small sensor board with 9 degrees of freedom that includes an accelerometer and a magnetometer. In order to solve the challenge of real-time 3D orientation tracking of the nozzle, we resort to the implementation of an AHRS (Attitude and Heading Reference System): a 3-axis sensor system that provides attitude position (e.g.: pitch, roll, and heading) by fusing accelerometer data and magnetometer data.

The analog control stick used is a Nintendo Nunchuk [15] compatible controller. Its relative small size allow to easily position it over a physical fire extinguisher without incurring in handling problems. The trainee is able to move inside the virtual environment intuitively using the thumb of same hand that carries the extinguisher. The controller allow forward/backward movement along the user's viewing direction, while the left/right axis allow strafe accordingly.

Since our training goal can greatly benefit from immersivity, we resort to an head-mounted display for the trainee to visualize the virtual environment. We decided to avoid vendors lock-ins, so we opt for a generic HMD mount able to house an any-brand smartphone.

We exploit the smartphone operating system and resources to host and run the application implementing the simulator logic. Thanks to their displays and graphics adapters an high visual quality can be obtained, able to satisfy immersivity requisites about visual realism. The vast majority of newest generation smarthphone are equipped with an internal accelerometer that is used to track trainee's head movements to determine his or her looking direction in the virtual environment. The Arduino Leonardo sends data from the input devices it handles to the smartphone by means of a standard micro USB cable, but a wireless solution is currently under study. Since the Arduino board is actually powered through the USB by the phone, renouncing to the cabled solution will introduce a new challenge. Figure 3 shows how the hardware components of the proposed node fit on the trainee. The main advantage of this setup is in its independence from external sensors (e.g: Microsoft Kinect II) or computational resources.

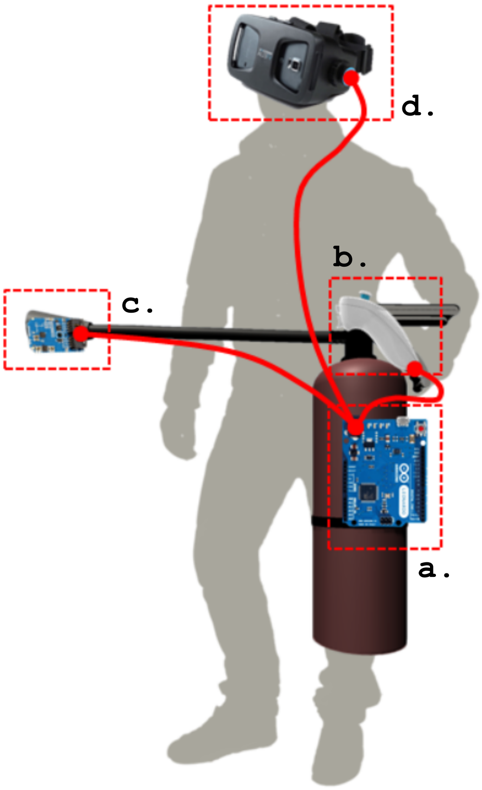


Fig. 3. Simulator' input and output devices wear by the user. The connected devices are: a. The Arduino Leonardo; b. Analogic Control Stick; c. Accellerometer for nozzle rotation tracking; d. Smartphone housed in the HMD mount.

Node software

The proposed simulator logic is implemented as an application developed using the Unity cross-platform game engine[16] and has been tested both on Windows and Android operating systems. It features different scenarios of possible fire related disaster that challenge the trainee acting to avoid fire escalation, by using in the proper way the fire extinguisher.

According to our in-the-field experience, a fire extinguishing system simulator should cover and include a number of features. Based on experimental training, the trainees should get trained to wear some protective devices and should experience the side effects when they forget to. The burning substance could be different and consequently also the typology of extinguisher both in terms of compounds (e.g., foam, powder,

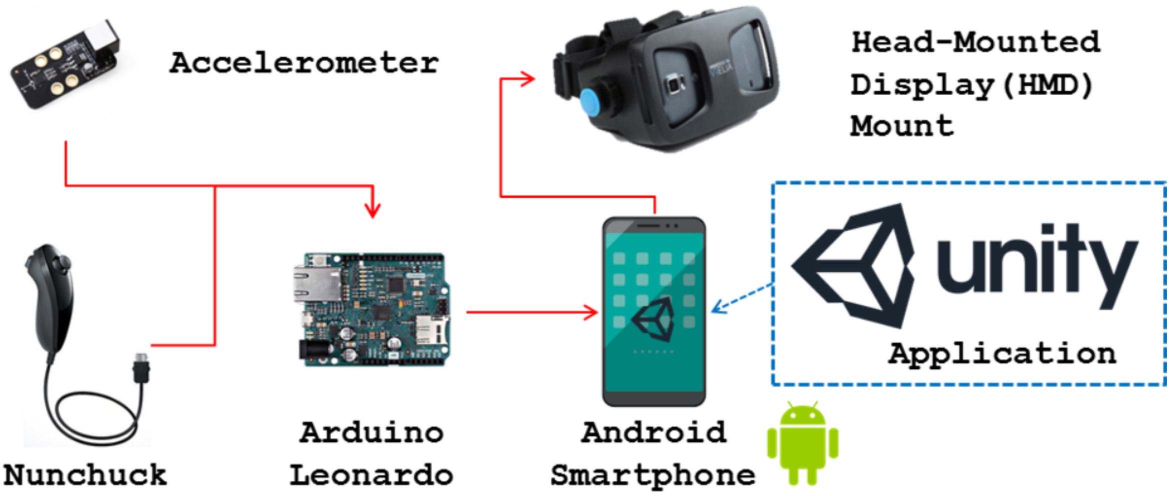


Fig. 2. Simulator structure, highlighting hardware interaction.

gas, liquid) and dimensions. The fire produced by the burning substance would consequently be different and its behavior subject to the surrounding environment. In case of outdoor fires, the flame should bend to winds. For the sake of training purposes, the trainee should take advantage from a virtual augment reality (AVR) feature. AVR would provide details about the heat radiation and the thermal load, which is the amount of heat received after exposure, i.e. the absorbed dose. A real-time graph would also provide details about the physical thresholds for first, second, and third degree burns to inform the trainee about the exposure risk to flames as a function of the donned self-protective devices and the distance from the fire.

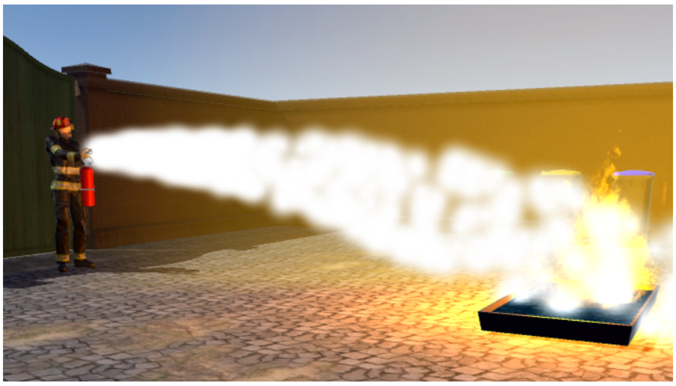


Fig. 4. Particle system used for fire extinguisher spray.

Each user is represented by an avatar handling the extinguisher, as can be seen from Figure 4. We invested a great deal of effort making the particle effects reproducing fires and sprays in a realistic way.

Note that during the training simulation, the trainee will see the environment through the avatar's eyes to ensure the immersivity.

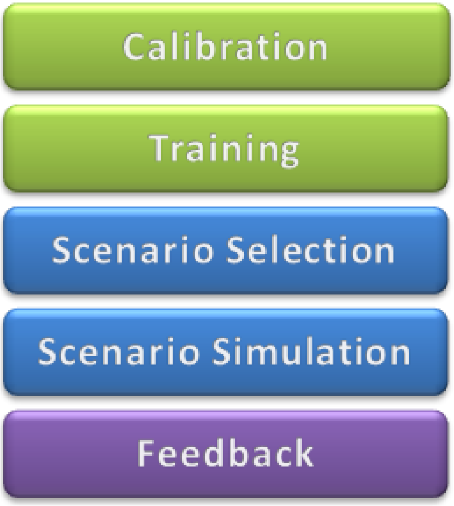


Fig. 5. Simulator stages.

In Figure 5 the stages that a user encounters during his interaction with our simulator are presented. When the application is launched the first time by a new learner it is necessary to go through a quick *calibration* process. It is necessary to correctly align input sensors to the virtual character and extinguisher for direction and pointing at purposes.

When the technical equipment is set up, the learner can access a first *training scenario* (See Figure 6) to familiarize with the basics of moving in the virtual environment and the nozzle pointing, as well as with the first very basic notion of extinguisher handling. These information are displayed to the trainee as a sequence of tutorial-like steps to be followed. After the training scenario, the learner can *select* one of the several fire disaster scenario, different both in terms of parameters of fire diffusion and fire intensity, both in terms of extinguisher typology and in terms of the environment virtually reproduced. The selected scenario is *simulated* and the training session can begin. The third layer introduced in the previous Section is in

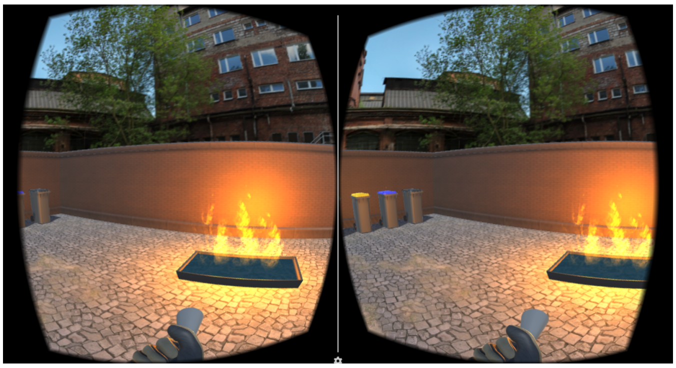


Fig. 6. Stereoscopic 360 view of the training yard level of the simulation.

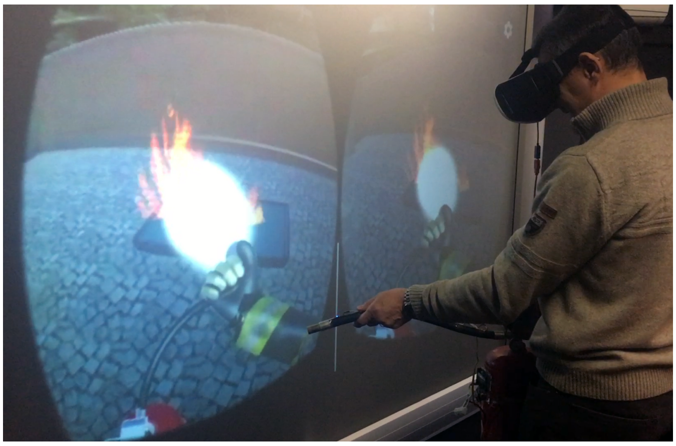


Fig. 7. Testing our simulator, with a screencast of the image as seen by the user-trainee.

its development stage and actually not yet used to interconnect the nodes. It is instead implemented on a stand-alone machine that, by means of a web browser, connects to an App (Screen Mirror[17], in particular) on the same smartphone on which the simulator is running. In this way, during training session it is possible to follow trainees' progresses by web-casting the visual result of the simulator on a screen (See Figure 7). After the session it is also possible to receive a *feedback* from the simulator in the form of collected statistics about trainees' performances.

The performance assessment of the trainees is accomplished by tracking and processing their decisions and actions. These include their distance from the fire, the impact angle of the gas/liquid/foam/powder jet on the flame, the emission time, and the total time spent to extinguish the fire. These bits of information, once automatically processed at the end on the experiment, provide some valuable details for either a self or a trainer assessment. The AVR feature can be activated or deactivated by the user or the trainer depending respectively on a training or assessment session.

MODEL AND EVALUATION

In order to support the design process of a simulation setup, it is important to predict the performances of each single node, in order to be able to verify that real time requirements are achievable, before experimenting on a prototype (in the system development phase) and before setting up a simulation (in the application of the system to a certain simulation scenario). A convenient approach is the use of a queuing network based model.

The queuing network needed for the task is composed by three modules per each node. For each node, two modules describe the device coordinator and the smartphone, respectively. These two modules are structured after the main tasks that they have to accomplish in the system, and form a pipeline. The third module, that is parametric in the number of IoT sensors, models the workload that the device coordinator has to manage from the IoT sensors. Figure 8 describes a possible implementation of the queuing network for a node of the simulation system.

With reference to Figure 8, from left to right and from top to down, the first depicted module is the IoT module. This module is composed of one queue per IoT sensor that is connected to the node ($IoTsens_i$, with i that assumes values between 1 and the number of connected sensors N). All the outputs are directed to the device coordinator module.

This module is composed of a queue that accounts for the management of the communications with the IoT sensors, a queue that accounts for the workload generated by passive sensors (that need a local processing of data, differently from IoT sensors, and that may need to be explicitly activated and controlled by the device coordinator, e.g. for polling) and a queue that accounts for the workload due to the software tasks connected to local data processing needs. The output of last queue is directed to the smartphone module.

The smartphone module is composed of a queue that models the local sensors (i.e. the ones integrated in the smartphone, such as accelerometers), a queue that models the data management operations that allow the interactions with the device coordinator, a queue that accounts for the tasks that allow the integration in the distributed simulation system by processing the events from the other nodes, and the related networking workloads, a queue that executes all the local tasks that allow to perform the simulation, including the generation of the additional VR/AR workloads to be executed by the dedicated hardware and the management of feedbacks, and a queue that models the behavior of the dedicated hardware.

The overall queuing network model is then obtained by a composition of one such model per node, and the interactions between the nodes are modeled by means of the dangling arcs showed in the figure, that are to be interpreted as one per kind towards and from any node to all others. This approach seamlessly allows to model a simulation system composed of heterogeneous nodes, that differ in the configuration and in the hardware parameters; moreover, the performances of a node may be assessed in isolation, by substituting the queuing networks representing the other nodes with a single fictional queue, that can be easily reconfigured to study system scaling problems. This approach has been designed as a support for overall simulation system assessment, node design parameter

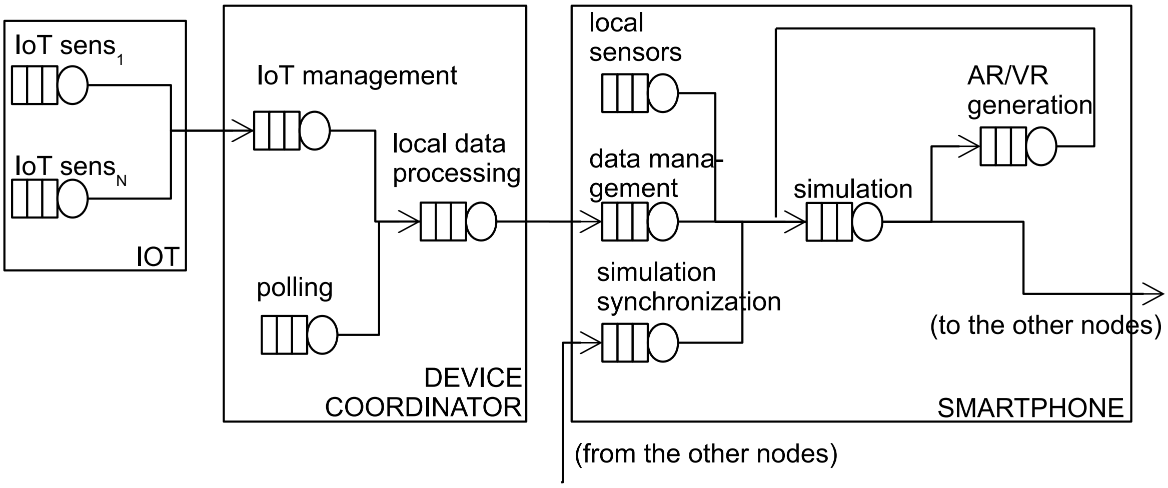


Fig. 8. The queuing network parametric model of a node

exploration and scaling problems.

To show the effectiveness of the approach, we applied the description in Figure 8 to the described case study, and simulating the presence of the other nodes by means of a single queue.

CONCLUSIONS AND FUTURE WORKS

In this paper we presented an immersive VR IoT based training support simulator architecture, which, by exploiting off-shelf components, can be implemented at low costs. Even if the presented results are very interesting, our work on the simulator is still in progress. In this paper we focused on the subsystem that is related to a single trainee, providing a performance evaluation oriented model to support the design of the system. In the future we plan to extend our study by adding the interaction of several single trainee subsystem, as well as performing a test campaign using fire department personal to asses the validity of the proposed architecture.

REFERENCES

[1] A. Craig, W. R. Sherman, and J. D. Will, *Developing Virtual Reality Applications: Foundations of Effective Design*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2009.

[2] P. Vichitvejpaisal, N. Yamee, and P. Marsertsri, "Firefighting simulation on virtual reality platform," in *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, July 2016, pp. 1–5.

[3] F. Poschner, "Fire fighting and related simulations in a cave using off-the-shelf hardware and software," in *Proceedings of SIGRAD 2014, Visual Computing, June 12-13, 2014, Göteborg, Sweden*, no. 106. Linköping University Electronic Press, Linköpings universitet, 2014, pp. 33–40.

[4] M. Maschek, "Real time simulation of fire extinguishing scenarios," Master's thesis, Institute of Computer Graphics and Algorithms, Vienna University of Technology, Favoritenstrasse 9-11/186, A-1040 Vienna, Austria, 2010. [Online]. Available: https://www.cg.tuwien.ac.at/research/publications/2010/maschek-2010-rts/

[5] iBG fire simulator. [Online]. Available: http://www.feuersimulator.com/index_en.html

[6] ViFeLoe. [Online]. Available: https://www.vrvis.at/research/projects/vifeloe/

[7] Pyrosoft - fire safety training. [Online]. Available: http://www.pyrosoft.ca/

[8] Bullex laser-driven fire extinguisher training. [Online]. Available: http://bullex.com/product/bullseye-publiceducation/

[9] ETC Training Systems - ADMS-Fire. [Online]. Available: http://www.trainingfordisastermanagement.com/products/adms-fire/

[10] Ludus firefighter services. [Online]. Available: http://www.ludus-vr.com/en/portfolio/firefighter-services/

[11] M. Swan, "Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0," *Journal of Sensor and Actuator Networks*, vol. 1, no. 3, p. 217–253, Nov 2012. [Online]. Available: http://dx.doi.org/10.3390/jsan1030217

[12] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645 – 1660, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167739X13000241

[13] Arduino Leonardo webpage. [Online]. Available: https://www.arduino.cc/en/Main/ArduinoBoardLeonardo

[14] Spark Fun 9DOF Sensor Stick. [Online]. Available: https://www.sparkfun.com/products/retired/10724

[15] Nunchuk Wii Remote controller. [Online]. Available: https://en.wikipedia.org/wiki/Wii_Remote#Nunchuk

[16] Unity cross-platform game engine. [Online]. Available: https://unity3d.com/

[17] Screen mirror androids app. [Online]. Available: https://play.google.com/store/apps/details?id=com.ajungg.screenmirror

PERFORMANCE EVALUATION OF MASSIVELY DISTRIBUTED MICROSERVICES BASED APPLICATIONS

Marco Gribaudo
DEIB
Politecnico di Milano
via Ponzio 51
20133, Milano, Italy
marco.gribaudo@polimi.it

Mauro Iacono
DMF
Università degli Studi della
Campania ”Luigi Vanvitelli”
viale Lincoln 5
81100 Caserta, Italy
mauro.iacono@unicampania.it

Daniele Manini
DI
Università degli Studi di Torino
corso Svizzera 185
10129, Torino, Italy
manini@di.unito.it

KEYWORDS

cloud infrastructures; data center performances; performance modeling; containers; microservices

ABSTRACT

Microservice-based software architectures are a recent trend, stemming from solutions that have been designed and experimented in big software companies, that aims to support devops and agile development strategies. The main point is that software architectures, similarly to what happens in SOA, are decomposed into very elementary tasks, that can be developed, maintained and deployed in isolation by small independent teams, and that compose an application by means of simple interactions. The resulting architecture is advocated to be more maintainable, less prone to failures, more agile, but obviously impacts on performances. In this paper we provide a simulation based approach to explore the impact of microservice-based software architectures in terms of performances and dependability, given a desired configuration. Our approach aims at giving a first approximation estimation of the behavior of different classes of microservice-based applications over a given system configuration, to characterize the infrastructure from the point of view of the service provider under a randomly generated realistic overall workload: to the best of our knowledge, there is not any other analogous decision support tool available in literature.

I. INTRODUCTION

Microservice-based software architectures are a technical solution that emerged from the industry sector to face the challenges that the market of cloud applications has created. Competition requires a continuous update and upgrade of applications that become larger and larger and serve simultaneously a very big number of users and requests, while the design, maintenance and deployment of larger and larger code bases results in more and more complex development and management cycles, exposing applications to a higher risk of fault propagation or more extended consequences of erroneous behaviors, due to coding errors. Microservice-based architectures support the decomposition of complex monolithic applications into a (high) number of very simple services, each responsible of elementary actions within the logic of the execution flow of an application, and interacting by means of simple (generally HTTP based or socket based) communication protocols.

A first advantage of this choice is a decoupling of microservices, each of which may be developed, maintained and administered by a different team, and each of which may be managed in a continuous integration mode, typical of agile development paradigms. Moreover, the small size of a microservice allow a small code base (based on an independent stack, seamlessly with respect to the stack used by the other microservices), a smaller team, and an easy integration of new team members with a shorter training.

A second advantage is the potential improvement of application resilience and scalability, as, being each service run independently, faults will not result in a crash of the whole application, and a different number of replicas of each microservice may be executed when and if needed, depending on the workload. Together with the use of containers, the increase in workload may be mitigated: anyway, in general, lower performances are reasonably expected with respect to monolithic applications.

While there is a simplification of the development cycle of microservices, the general structure of the application becomes more complex, because of the fact that planning and general design must suffer a lack of control, and, potentially, of optimization, on what is delegated to each microservice development team. The overall management of the execution environment and of the infrastructure, that is the most of what is left to the global level, heavily affects the general performances of the application, but with a high level of decentralization of responsibilities and control leverages. In this paper we investigate, with a stochastic simulation based

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

modeling approach, the general behavior of microservice-based applications with respect to performances, dependability and scalability, on a given infrastructure. The goal is to obtain a conceptual tool to provide methodological guidelines for the management of the infrastructure.

This paper is organized as follows: next Section provides related works, while Section III gives an overall introduction to microservice-based software architectures and applications; Section IV describes the simulation approach and the modeling framework; Section V offers the results of the test experiments that have been performed; finally, conclusions close the paper.

II. RELATED WORKS

At the state, there is not a wide academic literature about microservices, while there is plenty of good technical references related to implementation, cases and practical issues. Microservices architectures, as they are intended in the currently agreed definition, have been introduced in [1]. For a fast and readable introduction to the main themes about microservices in the cloud we suggest the reader to refer to [2], while for a systematic mapping of existing literature about the microservice architecture we suggest [3], to which we also redirect the readers for a more extensive reference list. In [4] the authors discuss, with a quite complete and solid analysis of all the aspects related to the executing architecture, the workload characterization of microservice architectures, with an experimental approach that benchmarks a monolithic application versus two different microservice versions, one based on a monothread support and one based on a multithread support, with very interesting results. In [5] and [6] the analysis focus instead on the costs and the benefits of the deployment of monolithic versus microservices and of monolithic versus AWS Lambda versus microservices architectures, considering both cloud customer or cloud provider operated systems. In [7] the scalability of the Docker container is evaluated, in different conditions, considering it as a new type of system workload. A similar study has been developed in [8], that identifies the challenges for a full development of containers based systems. For what concerns the operating condition, [9] describes a proposal for resilience testing, while [10] formulates a proposal for a decentralized autonomic behaviour for microservice architectures. Finally, for what concerns applications, the web offers a lot of proposals and descriptions: we rather prefer here to refer to a couple of peer reviewed papers, [11] and [12], as a starting point for readers, for their clear and systematic presentation.

III. MICROSERVICE ARCHITECTURES

A microservices based software architecture is an application composed of a number of software services that may be independently deployed and that directly interact with one another with lightweight mechanisms [4]. Each microservice is executed in a separate process. In general, a microservice is executed as a native process of the host, by means of a container, that is an abstraction layer capable of virtualizing resources with a low, but non negligible, impact on performances, and providing isolation. The technology stack may or may not allow multithreading, introducing a further element of complexity for the evaluation of performances: a multithreading solution may be more efficient by exploiting pooling, but suffers e.g. from blocking caused by I/O request waits; a single thread solution does not, but processes need more resources than threads. The use of containers allows a microservices based application to be seamlessly migrated as a whole, as it is commonly used for an agile deployment from the development to the production environment, that is generally cloud based. However, in the practice an application spans over a large number of different containers, that may theoretically get up to the number of microservices, to allow the enactment of agile development processes. This has an impact on performances as well, because the interactions between the microservices need a virtualized network between containers. Finally, the number of active containers is also a leverage to scale performances up and down when needed to fit the dynamics of the workload. As described in [2], Fig. 1 shows a graphical representation of the relations between microservices (μs), virtual machines (VM) and nodes (HM - Hardware Machines). As usual in cloud computing, each node can host several VMs and partition resources among them. Moreover, each VM can host several microservices that are used to implement the application.

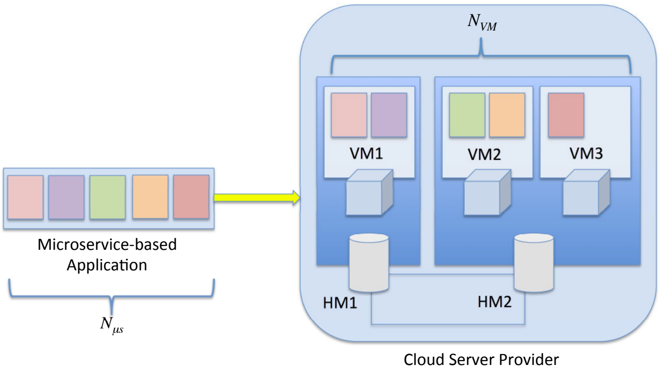


Fig. 1. General schema of a microservices based architecture

Each container is executed within the operating system (OS) of which it virtualizes the resources, as a process, and provides its services by means of a client-server logic. When run in cloud environments, containers are executed within VMs. Consequently, containers are executed on the OS provided by a VM that in turn is executed on the computing nodes by means of the host OS or a hypervisor, eventually together with other VMs. The computing resources of the node, that is generally a multiprocessor and/or multicore architecture, are so managed in order to map each thread of a microservice to a core and each process to a processor.

Within the cloud infrastructure, VMs are managed according to the internal policies that provide elasticity and power management features. VMs may be migrated, shelved or launched when needed, similarly to what happens to threads within containers and to containers within a VM. A correct estimation of the best policies for the provider, consequently, requires models that may allow to understand the overall effects of the interactions within and between the different levels, and evaluate the role of the various available parameters. We already dealt with performance modeling of cloud architectures [13] [14] and multithreaded applications [15]: in the following we will focus on the microservice architecture, including all the architectural details of the whole cloud stack that are relevant for performance evaluation.

IV. SIMULATION APPROACH AND MODEL

As, to the best of our knowledge, there is not at the moment a general simulation approach for microservice architectures, nor there are extensive characterizations of their parameters available, the simulation approach adopted in this paper is designed to produce a first approximation glance on the general behavior of these systems without a single reference scenario: the goal is to provide an estimation of performance and dependability for different configurations. Consequently, the approach is based on a parametric generation of a large number of different possible microservices applications, defined as oriented graphs, with a parametric random resource usage and fault probability per microservice, that are mapped onto a parametric architecture, in which the number of servers (in a cloud), the number of VMs per server, the number of containers per VM can be varied: moreover, a fault probability is assigned to every component of the architecture. The workflow of the approach is depicted in Fig. 2. A set of applications is generated, according to chosen parameters, by a *scenario generator*, that instantiates a simulation per case. Simulations are run by an event based *simulator* that has been specificly designed for this paper. The simulator produces performance and dependability metrics, and the results of simulations are then processed by a *statistics* processor that produces an overall performance profile of the given system configuration(s), described as a function of the different parameters and the scenarios.

More in details, in this work we used Montecarlo simulation to generate several random application topologies and study their performance and availability, and proper maximum entropy probability distributions, to avoid biasing due to the lack of assessed models, or the Zipf distribution in analogy to web traffic characterizations.

In each simulated scenario (see Fig. 1) an application is split in $N_{\mu s}$ microservices, and it can be executed by users at rate of λ_u requests per second. Microservices are allocated within proper containers in VMs executed on the top of host machines provided by a cloud infrastructure, the total number of available VMs, according to the contract, is denoted with N_{VM}. We defined most parameters with stochastic numbers generated with probability distributions. In this case study, we assume that the number of microservices $N_{\mu s}$ has a Poisson distribution with parameter $\lambda_{\mu s}$. Next, we consider that the number of VMs on which microservice containers are deployed N_{VM} is a fraction of β of $N_{\mu s}$, to represent that subsets of containers can be allocated on the same VM. In particular, we define:

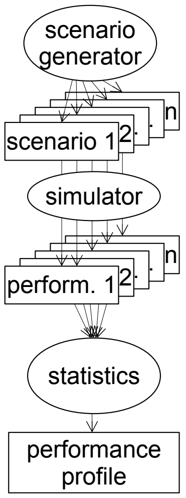


Fig. 2. The workflow of the simulation approach

$$N_{VM} = \lceil \beta \cdot N_{\mu s} \rceil \tag{1}$$

Each microservice can be invoked a random number of times during the execution of the application. We assume this random number to be geometrically distributed, with an average v_i for each microservice i. We assume that not all the microservices are equally used in serving a request. Some can be essential, and will consequently be called several times during the execution (e.g. verification of user identity), while some other ones might be required only in special circumstances. We thus assign to each microservice i, $1 \leq i \leq N_{\mu s}$, a random average number of executions v_i, according to a popularity level. In particular, we assume that microservices with a lower index i are more popular (i.e. have a larger average number of executions) than services with a high index. Popularity follows a modified Zipf distribution, characterized by 4 parameters: c (the scale parameter), s (the shape parameter), q (the shift parameter) and α (the randomness parameters). Let $\mathbf{u}_i$ be a random number, uniformly distributed in the range $[0,1]$. We compute, for each randomly generated scenario, the average number of calls to a microservice i as:

$$v_i = \frac{c}{(i + q + \alpha \cdot \mathbf{u})^s} \tag{2}$$

We assume that the execution time for each microservice i is exponentially distributed with average S_i. In each scenario for any micorservice i the value of S_i is randomly defined according to an Erlang distribution with k_S stages, and average λ_S. In order to consider VM faults, Mean Time To Failure (MTTF) and Mean Time To Repair (MTTR) parameters are sampled for both the VMs (infrastructure faults) and the microservices (software faults). In particular, such parameters are $MTTF_{VM}$, $MTTR_{VM}$, $MTTF_{\mu s}$ and $MTTR_{\mu s}$. For each random scenario generated, we sample the previous parameters from four different Erlang distributions, each one characterized by its number of stages ($k_{MTTF_{VM}}, \ldots, k_{MTTR_{\mu s}}$), and average time ($\lambda_{MTTF_{VM}}, \ldots, \lambda_{MTTR_{\mu s}}$).

A. VMs allocation

Many allocation policies can be described in our model. In particular, for each VM j we denote with $\mathbf{M}_j \subseteq \{1, \ldots, N_{\mu s}\}$ the set of microservices that are executed over it. The set $\mathbf{M}_j$ forms a partition: $\bigcup_{1 \leq N_{\mu s}} A_j = \{1, \ldots, N_{\mu s}\}$, and $A_j \cap A_k = \emptyset$, $\forall 1 \leq j, k \leq N_{VM}$. In this work we have considered two policies, addressed in the following as *case I* and *case II*. The *case I* policy defines an elementary strategy according to which containers are cyclically assigned to VMs taken from a list obtained as a random permutation of available VM number identifiers. When the application execution starts, a container is assigned to the first VM in list, then the next is assigned to the following VM in list, and so on; if the VM list is over, containers are assigned starting again from the top of the list. In this way each VM has at least one container, and services are assigned to computing resources in a random way.
The *case II* strategy instead tries to compact the containers according to their demand. The two containers with the smallest requirements are merged together on the same VM, creating a new single ”equivalent” container whose demand is the sum of the ones of the services that are combined. The equivalent container replaces the two merged services, and the process is repeated until there is just one equivalent container per VM. In this way, the average utilization of the VMs is maximized, creating a more balanced system.
Other strategies that can be easily included in the model can for example be based on the actual load, overall utilization, or containers can be assigned to VMs taking into account their task and requirements.

B. Performance indexes

Given these scenario and parameters we are able to evaluate system performance by computing the indexes we report in the following. The load introduced in the cloud infrastructure by any microservice i is easily derived as:

$$D_i = v_i \cdot S_i \tag{3}$$

By counting the number of containers assigned to each VM we can compute VMs load and utilization. Hence, for any VM j we have the load defined as

$$D_j = \sum_{i \in \mathbf{M}_j} D_i \tag{4}$$

Remembering that λ_u is the rate at which users requests for the application arrives to the system, the utilization U_j for any VM j, when the system is stable, can be computed as:

$$U_j = \lambda_u \cdot D_j \tag{5}$$

In the same way, one particular configuration is stable only if Equation 5 is strictly less then one for all VMs, or equivalently if:

$$\lambda_u < \frac{1}{\max\limits_{1 \leq j \leq N_{VM}} D_j} \tag{6}$$

Throughput X_i of a container i is computed as $X_i = \lambda_u \cdot v_i$; the average response time of a VM j and the average system response time R as:

$$R_j = \frac{D_j}{1 - U_j} \qquad R = \sum_{k=1}^{N_{VM}} R_j \tag{7}$$

The availability of the application $A = A_{VM} \cdot A_{\mu s}$ is computed as the product of the availabilities of the VMs (A_{VM}) and of the microservices ($A_{\mu s}$) used during one application execution. Since each VM j and each microservice i is characterized by its own mean time to failure and mean time to repair, we can define their corresponding availability as:

$$A_{VM_j} = \frac{MTTF_{VM_j}}{MTTF_{VM_j} + MTTR_{VM_j}} \tag{8}$$

$$A_{\mu s_i} = \frac{MTTF_{\mu s_i}}{MTTF_{\mu s_i} + MTTR_{\mu s_i}} \tag{9}$$

However, since not all microservices are required during each application execution, the fault of a machine will not always cause a failure. Let us call p_{VM_j} and $p_{\mu s_i}$ respectively the probabilities that VM j or microservice i are used during a call to the application. Then effective availability $\hat{A}$ can be computed as:

$$\begin{aligned} \hat{A}_{VM_j} &= 1 \cdot (1 - p_{VM_j}) + A_{VM_j} \cdot p_{VM_j} \\ &= 1 - (1 - A_{VM_j}) \cdot p_{VM_j} \qquad (10) \\ \hat{A}_{\mu s_i} &= 1 - (1 - A_{\mu s_i}) \cdot p_{\mu s_i} \qquad (11) \end{aligned}$$

Due to the geometric assumption, and since each microservice i is called an average of v_i times, we have:

$$p_{\mu s_i} = Pr\left\{ Geom\left(\frac{1}{v_i + 1}\right) > 0 \right\} = \frac{v_i}{v_i + 1} \tag{12}$$

The probability that VM j is used during a call to the application must instead consider the fact that at least one microservices i allocated over it is required, which could be expressed as:

$$p_{VM_j} = 1 - \prod_{i \in \mathbf{M}_j} \left(1 - p_{\mu s u_i}\right) \tag{13}$$

The final values of A_{VM} and $A_{\mu s}$ can then be computed as:

$$A_{VM} = \prod_{j=1}^{N_{VM}} \hat{A}_{VM_j}, \qquad A_{\mu s} = \prod_{i=1}^{N_{\mu s}} \hat{A}_{\mu s_i} \tag{14}$$

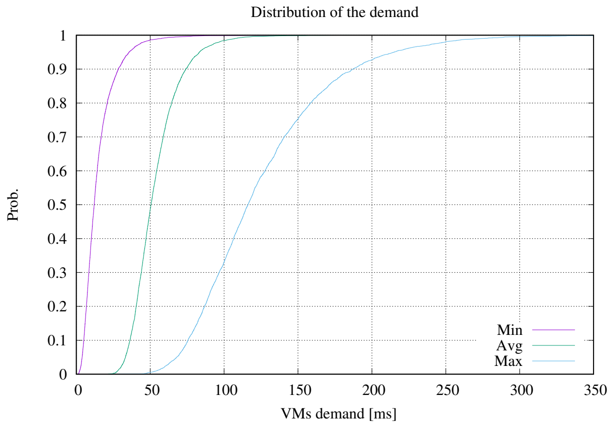


Fig. 3. Probability distribution of VMs demand

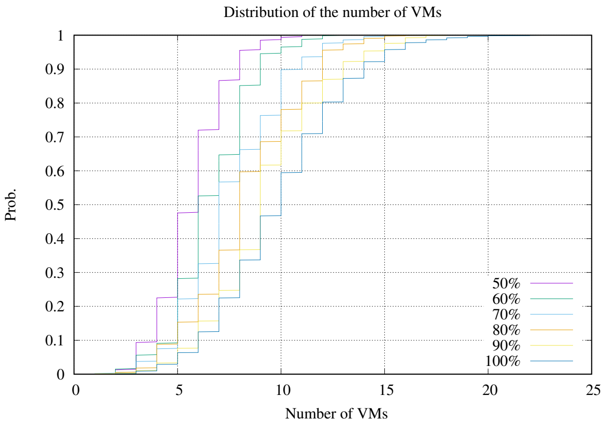


Fig. 4. Probability distribution of the number of VMs

V. EXPERIMENTS

In the following we report the results obtained from a set of experiments. For each configuration, we have generated between 5000 to 100000 random applications (depending on the fraction of stable cases), and we have collected both average values and distributions. We show the outcome in three different classes: infrastructure, performance, and availability. In all the following case studies we have set the parameters corresponding to the popularity and the average service time as reported in Table I.

TABLE I
COMMON PARAMETERS FOR ALL THE EXPERIMENTS

c	6	q	2
α	1	s	1.5
λ_S	100 ms.	k_s	4
$\lambda_{MTTF_{VM}}$	1000 h.	$\lambda_{MTTR_{VM}}$	2 h.
$\lambda_{MTTF_{\mu s}}$	500 h.	$\lambda_{MTTR_{\mu s}}$	6 min.

A. Infrastructure

The figures of this section plot indexes related to the system structure: their purpose is to show the main features that the applications generated with the considered set of parameter distributions have. Fig. 3 shows minimum, average, and maximum of the demand load on each VMs. The parameters setting is $\lambda_u = 1$ call / sec., $\lambda_{\mu s} = 10$ ms., $\beta = 66\%$, the VM selection policy is Case I. As we can see, the popularity mechanism creates a few microservices that are heavily loaded: for this reason the average demand is more skewed towards the minimum.

The probability distribution of the number of VMs for different values of β is reported in Fig. 4: the savings that could be made in number of VMs with a lower value of β become important only when the application is composed by a large number of microservices.

The probability distribution of the number of microservices is showed in Fig. 5 as function of the $\lambda_{\mu s}$: the Poisson distribution provides a good way to generate meaningful topologies when only a rough idea on the average number of microservices is available.

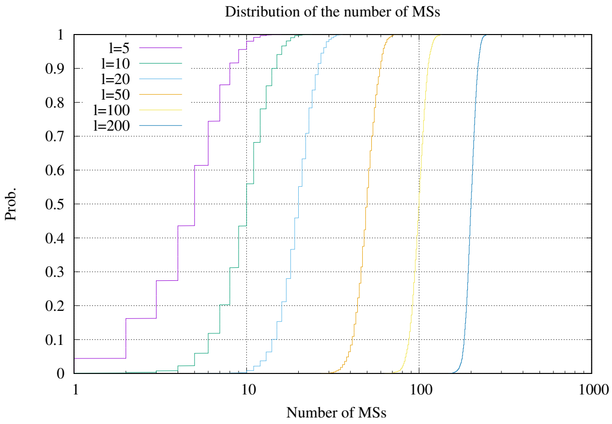


Fig. 5. Probability distribution of the number of microservices

B. Performance

In this section performance indexes are reported. We first analyze the system response time versus the user demand. Fig. 6 shows the probability distribution of the response time with increasing values of λ_u, $\lambda_{\mu s} = 10$, and $\beta = 66\%$, the VM selection policy is Case I. As expected the system reacts more slowly when microservices execution demand is higher. Moreover, the probability distributions are defective, since as the load increases, there is a higher chance of obtaining unstable topologies that are excluded from the output.

In Fig. 7 mean response time and average percentage of stable runs are compared in case I and II versus the user load λ_u, with $\lambda_{\mu s} = 10$ and $\beta = 66\%$. Since mean values are considered, confidence intervals are reported. Case II outperforms the response time of Case I, with the exception of the case with the highest value of λ_u. The reason lies in the fact that with this value the model has a relevant number of unstable runs.

After the evaluation versus the user load, we studied the indexes as a function of the number of VMs. Fig. 8 reports

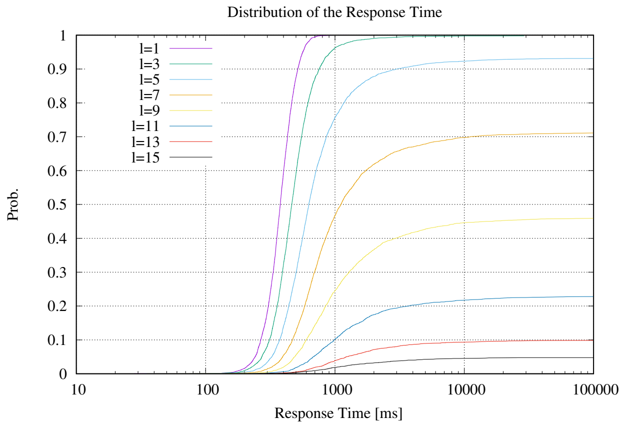


Fig. 6. Probability distribution of the response time

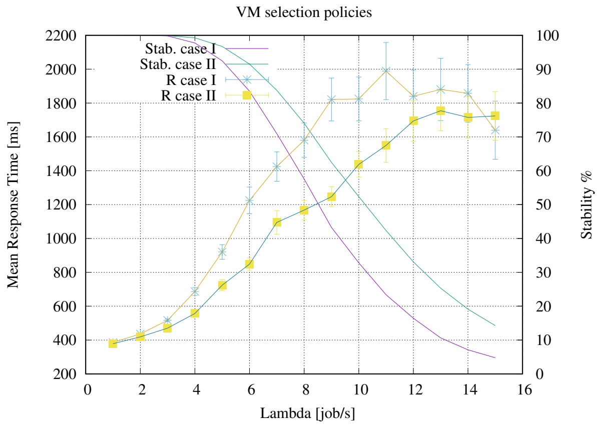


Fig. 7. Comparison of VMs policies

the probability distribution of VMs utilization with different values of β.Parameters are $\lambda_u = 12$, $\lambda_{\mu s} = 10$, the VM selection policy is Case I. When the percentage of VMs is lower, each VM get more microservices to be executed and hence the overall VMs utilization is higher.

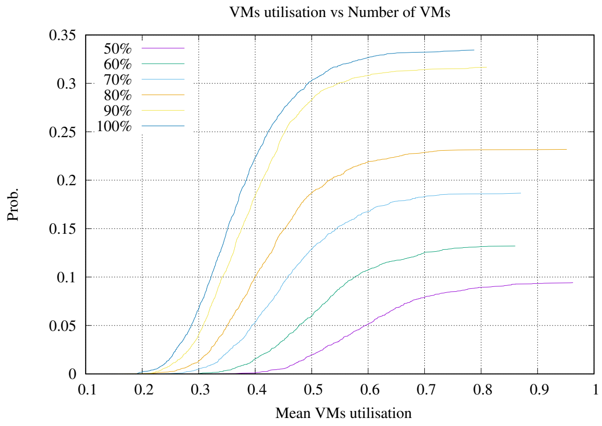


Fig. 8. Probability distribution of VMs utilization

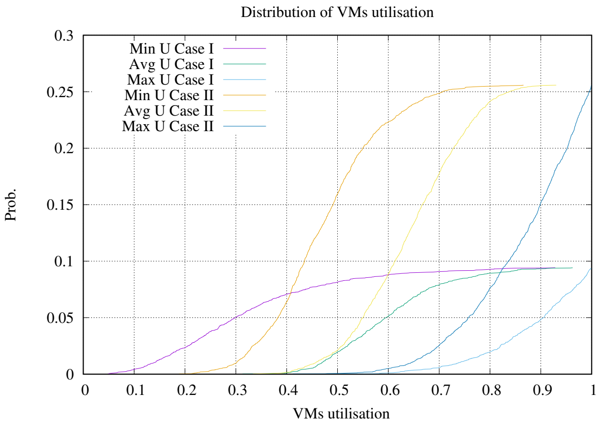


Fig. 9. VMs utilization with different policies

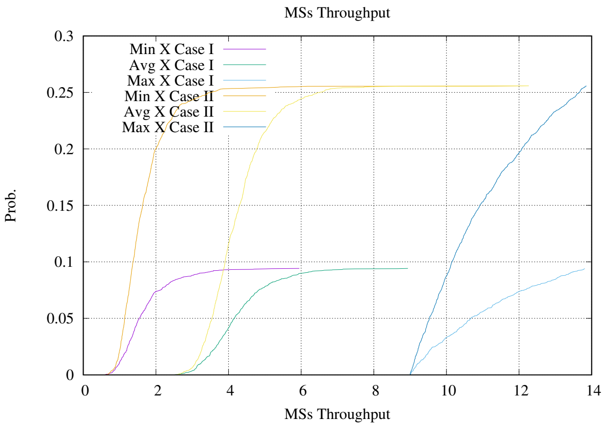


Fig. 10. MSs throughput

Fig. 9 shows the distribution of the minimum, average, and maximum VMs utilization in each topology, with different policies. Parameters are $\lambda_u = 12$, $\lambda_{\mu s} = 10$, and $\beta = 50\%$. The utilization in Case I is lower, but as it is showed in Fig. 7 this policy is less unstable, indeed the VM selection is random and it can happen that some VM get higher load then the others.

The distribution of the minimum, average and maximum throughput of the microservices in each simulation run is reported in Fig. 10. As it can be seen, although both cases has the same throughput (since this parameter is determined by the application and not by its deployment on the infrastructure) Case II policy performs better than the Case I, having a larger number of configuration in which the system is stable.

C. Availability

Finally, we evaluated the system from the availability point of view. Fig. 11 shows the overall, microservices, and VMs unavailability, with $\lambda_u = 1$, $\lambda_{\mu s} = 10$, and $\beta = 66\%$. The Erlang distribution used to generate the MTTF and MTTR are all characterized by 10 stages. The other parameters are reported in Table I. As it can be seen, in this scenario most of the faults are caused by software error in microservices rather than problems with the VMs.

The probability distribution of the availability versus VMs and microservices MTTFs are reported in Fig. 12 and Fig.

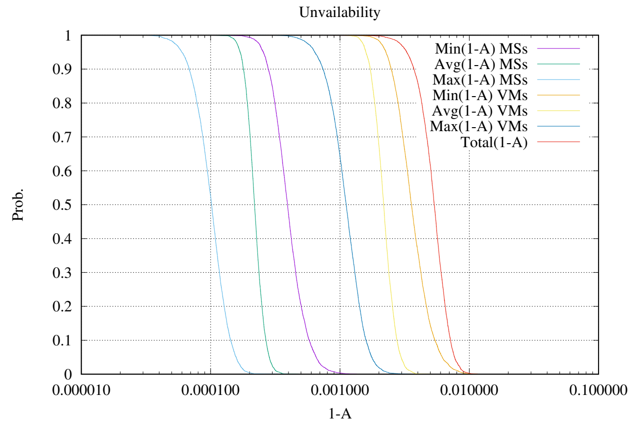


Fig. 11. Total, microservices, and VMs unavailability

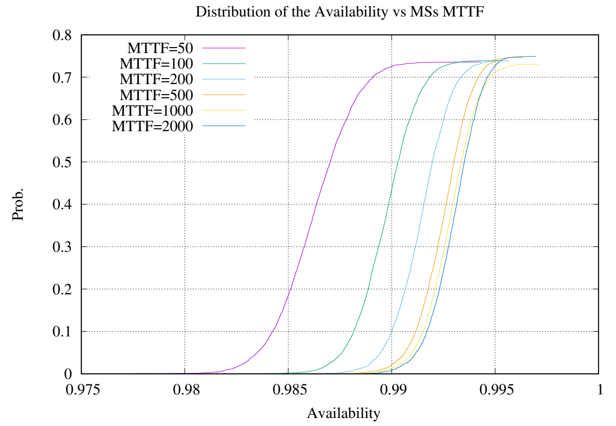


Fig. 12. Probability distribution of the availability versus microservice MTTF

13 respectively, with $\lambda_u = 8$, $\lambda_{\mu s} = 20$, and $\beta = 66\%$, Case II policies is used. As expected, the higher the MTTF, the higher the probability to have better availability.

VI. Conclusions

In this paper we proposed an approach for performance evaluation of infrastructures that support the execution of a mix of microservice-based software applications. Our approach, to the best of our knowledge, is the first parametric simulation approach that allows providers to model such architectures in a general case of an aggregated heterogeneous tunable workload mix, to support decisions in the design, maintenance and management of microservice-based infrastructures. Future works include further parameterization of the simulation approach, the exploration of massive real workloads for a better modeling approach, the extension of the simulation support for more infrastructural configurations, and a more accurate validation campaign when sufficient data will be available about traces from real, production infrastructures. Finally, the simulator will be extended in order to include energy issues.

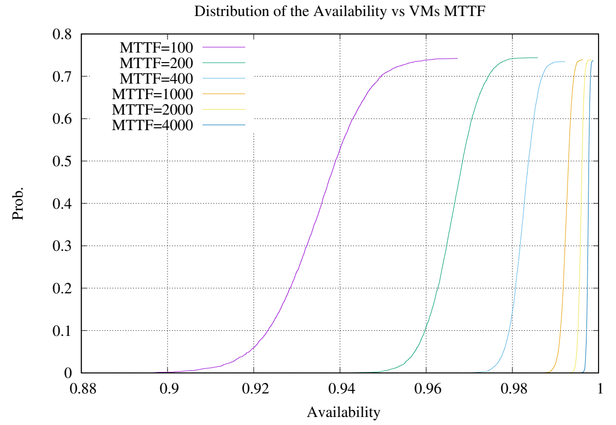


Fig. 13. Probability distribution of the availability versus VM MTTF

References

[1] "Microservices (a definition of this new architectural term)," https://martinfowler.com/articles/microservices.html, accessed: 2017-01-25.

[2] C. Esposito, A. Castiglione, and K. K. R. Choo, "Challenges in delivering software in the cloud as microservices," *IEEE Cloud Computing*, vol. 3, no. 5, pp. 10–14, Sept 2016.

[3] N. Alshuqayran, N. Ali, and R. Evans, "A systematic mapping study in microservice architecture," in *2016 IEEE 9th International Conference on Service-Oriented Computing and Applications (SOCA)*, Nov 2016, pp. 44–51.

[4] T. Ueda, T. Nakaike, and M. Ohara, "Workload characterization for microservices," in *2016 IEEE International Symposium on Workload Characterization (IISWC)*, Sept 2016, pp. 1–10.

[5] M. Villamizar, O. Garcs, H. Castro, M. Verano, L. Salamanca, R. Casallas, and S. Gil, "Evaluating the monolithic and the microservice architecture pattern to deploy web applications in the cloud," in *2015 10th Computing Colombian Conference (10CCC)*, Sept 2015, pp. 583–590.

[6] M. Villamizar, O. Garcs, L. Ochoa, H. Castro, L. Salamanca, M. Verano, R. Casallas, S. Gil, C. Valencia, A. Zambrano, and M. Lang, "Infrastructure cost comparison of running web applications in the cloud using aws lambda and monolithic and microservice architectures," in *2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, May 2016, pp. 179–182.

[7] T. Inagaki, Y. Ueda, and M. Ohara, "Container management as emerging workload for operating systems," in *2016 IEEE International Symposium on Workload Characterization (IISWC)*, Sept 2016, pp. 1–10.

[8] H. Kang, M. Le, and S. Tao, "Container and microservice driven design for cloud infrastructure DevOps," in *2016 IEEE International Conference on Cloud Engineering (IC2E)*, April 2016, pp. 202–211.

[9] V. Heorhiadi, S. Rajagopalan, H. Jamjoom, M. K. Reiter, and V. Sekar, "Gremlin: Systematic resilience testing of microservices," in *2016 IEEE 36th International Conference on Distributed Computing Systems (ICDCS)*, June 2016, pp. 57–66.

[10] L. Florio and E. D. Nitto, "Gru: An approach to introduce decentralized autonomic behavior in microservices architectures," in *2016 IEEE International Conference on Autonomic Computing (ICAC)*, July 2016, pp. 357–362.

[11] A. Melis, S. Mirri, C. Prandi, M. Prandini, P. Salomoni, and F. Callegati, "Crowdsensing for smart mobility through a service-oriented architecture," in *2016 IEEE International Smart Cities Conference (ISC2)*, Sept 2016, pp. 1–2.

[12] B. Butzin, F. Golatowski, and D. Timmermann, "Microservices approach for the internet of things," in *2016 IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sept 2016, pp. 1–6.

[13] A. Castiglione, M. Gribaudo, M. Iacono, and F. Palmieri, "Modeling performances of concurrent big data applications," *Software: Practice and Experience*, vol. 45, no. 8, pp. 1127–1144, 2015.

[14] M. Gribaudo, M. Iacono, and D. Manini, "Three layers network influence on cloud data center performances," 2016, pp. 621–627.

[15] D. Cerotti, M. Gribaudo, M. Iacono, and P. Piazzolla, "Modeling and analysis of performances for concurrent multithread applications on multicore and graphics processing unit systems," *Concurrency and Computation: Practice and Experience*, vol. 28, no. 2, pp. 438–452, 2016, cpe.3504.

MODELING A SESSION-BASED BOTS' ARRIVAL PROCESS AT A WEB SERVER

Grażyna Suchacka
Institute of Mathematics and Informatics
Opole University
ul. Oleska 48
45-052 Opole, Poland
E-mail: gsuchacka@uni.opole.pl

Daria Wotzka
Faculty of Electrical Engineering, Automatic Control and Informatics
Opole University of Technology
ul. Prószkowska 76
45-758 Opole, Poland
E-mail: d.wotzka@po.opole.pl

KEYWORDS
Web traffic, Web workload, Web server, log file, user session, Web bot, Internet robot, analysis and modeling, regression analysis

ABSTRACT

The paper deals with the problem of modeling key features of the Web traffic generated by Internet bots, observed at the input of a Web server. Based on real log data of an online store, a set of bot sessions was prepared and analyzed. Three session features connected with bots' arrival process at the server were analyzed: session interarrival time, request interarrival time, and the number of requests in session. Distributional models for these bot session features were developed using regression analysis and validated through graphical comparisons of histograms for the empirical data and simulated values. As a result, interarrival times of bot sessions and interarrival times of requests in bot sessions were modeled by a Weibull and a Pareto distribution, respectively, and the numbers of requests in session were modeled by a function being a combination of a sigmoid and exponential distributions. The aim of our analysis was to develop a model of a session-based bot arrival process on a Web server which may be then implemented in a bot traffic generator integrated with a Web server simulator.

INTRODUCTION

The continuous growth of the Internet community and development of Web-based technologies have been accompanied by a persistent growth and proliferation of Internet robots. According to (Geroimenko 2004), an *Internet robot*, also called *Web bot*, *Web agent*, *software agent*, or *intelligent agent*, is "a software tool that carries out a task on behalf of a user or computer, typically relatively autonomously". Search engine crawlers, shopbots, link archivers, and other autonomous software agents are constantly crawling the Web, making it possible to provide Web users with a variety of fast, up-to-date, and reliable information. There are also harmful bots, however, which hack computer systems or user accounts, steal Web content, carry out DoS (Denial of Service) attacks, or perform other malicious actions.

The increasing activity of bots has motivated research into bot traffic characterization and analysis. In particular, many studies have focused on the detection of various types of malware and hacking bots activities, both in computer networks (Kołaczek and Juszczyszyn 2012; Skrzewski 2014) and on end hosts (Liu et al. 2008; Stevanovic et al. 2011; Skrzewski 2016).

The analysis of bot traffic has been typically performed using data recorded in Web server access logs (Doran and Gokhale 2010; Doran et al. 2013; Suchacka 2014). As regards the characterization of various kinds of bots, much attention has been paid to Web crawlers (Calzarossa and Massari 2013; Dikaiakos et al. 2005). Some studies have tried to classify different types of bots, including text crawlers, link checkers and icon crawlers (Lee et al. 2009) based on the finding that various types of bots exhibit different traffic characteristics.

A variety of statistical and machine learning techniques have been applied to detect bot traffic from Web server log data (Saputra et al. 2013; Stassopoulou and Dikaiakos 2009; Stevanovic et al. 2011; Suchacka and Sobków 2015).

Although many studies have addressed the issue of modeling Web traffic features for the overall traffic incoming to the server, relatively few such studies have been dedicated to bot traffic. Therefore, the arrival process of Web clients on a Web server has been well characterized and modeled but this is not the case for bots' arrivals. What is more, the character and properties of bot traffic is subject to more rapid changes than the traffic resulting from human users' visits on the Web, mainly due to the dynamic development of Web technologies relying on bots' indexing and monitoring activities – Web analytics, Internet marketing, price and product comparison Web services, etc.

As regards distributional models describing bots' arrival process on a Web server, some traffic features have been analyzed and modeled for known bots (Doran and Gokhale 2010) and for specific bots, including crawlers and shopbots (Almeida et al. 2001; Calzarossa and Massari 2013). In (Doran and Gokhale 2010) interarrival times of known bots' sessions were shown to be heavy-tailed and follow a hybrid lognormal-Pareto distribution. On the other hand, request interarrival times in known bots' sessions were not heavy-tailed and

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

such distributions as lognormal, Weibull or Pareto did not fit the analyzed data set. Request interarrival times were modeled by a lognormal distribution for crawler sessions (Almeida et al. 2001; Calzarossa and Massari 2013) and with an exponential distribution for shopbot sessions (Almeida et al. 2001).

The literature review has shown that the main goal of previous analyses of a bot arrival process on a Web server has been to characterize and detect Web bots to cope with the consequences of their visits. In contrast, the motivation for our study was the need for differentiation between bots and human users in synthetic Web traffic being generated during simulation experiments testing the performance of a Web server system under various overload conditions.

Based on real log data of a Web store site we analyze and model key features of Web bot traffic in order to create a simulation model of bot arrivals on a Web server. The model is intended to be implemented in a bot traffic generator and used in simulation experiments to generate a stream of HTTP requests emulating the real bot traffic. An assumption underlying the planned simulation model is the ability to emulate many independent clients (including bots) interacting with the server. Thus, the implementation of such a model will make it possible to monitor and treat the emulated sessions in the server system independently from one another – which is not possible when using aggregated Web traffic models or reproducing real Web traces.

RESEARCH METHODOLOGY

Data Collection and Preprocessing

Source data of the Web traffic are Web server access log files. Basic log data describing each HTTP request coming to the server includes:

- information on the Web client sending the request (an IP address, a user agent string, a version of HTTP protocol, a referrer),
- information on the requested server resource (an URI of the resource, an HTTP method);
- information on request's processing at the server (a timestamp indicating request's arrival time, an HTTP status code, and a size of the file transferred to the client).

Processing and analysis of the requests' data using a computer program allows one to describe Web clients' sessions in an observation window. A *session* may be defined as a sequence of HTTP requests sent by a Web client during a single visit to the website (in some cases a session may contain only one request). A client is identified based on two data combined: the IP address and the user agent string. To identify multiple visits of the same client a minimum value of a time gap between any two consecutive requests of the client is defined: in the literature this value was established as 30 minutes (Catledge and Pitkow 1995; Bomhardt et al. 2005; Stevanovic et al. 2011).

Preparation of a Dataset of Web Bot Sessions

In the set of all sessions only the ones accomplished by Web bots were used for the analysis. Identifying bot sessions is not a trivial task. Only some bots may be recognized by requesting the file "robots.txt" in their sessions or by a bot name declared in the user agent string.

Furthermore, we assume that a Web client is a bot if the mean time per page in session is less than 0.5 second or by identifying some atypical behavioral patterns (Suchacka 2014): the empty referrer of the first request in session, all page requests or even all HTTP requests in session with empty referrers, combined with such observations as all requests containing the HTTP method *HEAD* (instead of the most popular method *GET*) or all requests with the erroneous HTTP status code and the image-to-page ratio in session equal to zero.

Moreover, each session containing only one request may be attributed to a Web bot as well. Even if a human user visits only one page of the online store website, their client – Web browser – generates and sends to the server multiple HTTP requests (a hit for a page description file and the following hits for embedded objects, which in the case of an e-commerce site are typically image files).

To avoid truncation of reconstructed bot sessions on the edges of the observation window, the target *bot sessions' assembling phase* was extended by two additional phases: an *initial phase* and a *final phase* (Fig. 1). Bot sessions which started in the *initial* phase are not used for the analysis. A *final phase* was introduced to allow time for completion of sessions initiated in the previous phase (sessions initiated in the *final phase* are not taken into consideration).

Developing Distributional Models for Bot Sessions' Features Using Regression Analysis

Three features of bot sessions, connected with the character of bots' arrival process on the server, were modeled in our study:

- session interarrival time,
- request interarrival time (in session),
- the number of requests in session.

These features are analyzed while maintaining the integrity of sessions. The way of measuring session interarrival times and request interarrival times is illustrated in Fig. 2.

Before the analysis a dataset of all bot sessions is divided into *N* subsets corresponding to consecutive days (because in the case of huge datasets there might be a problem with data processing). Each subset contains samples from sessions which were initiated on the corresponding day (although duration of some sessions might extend to the following day). Thus, for each session feature there are *N* subsets of data samples and each subset is analyzed separately.

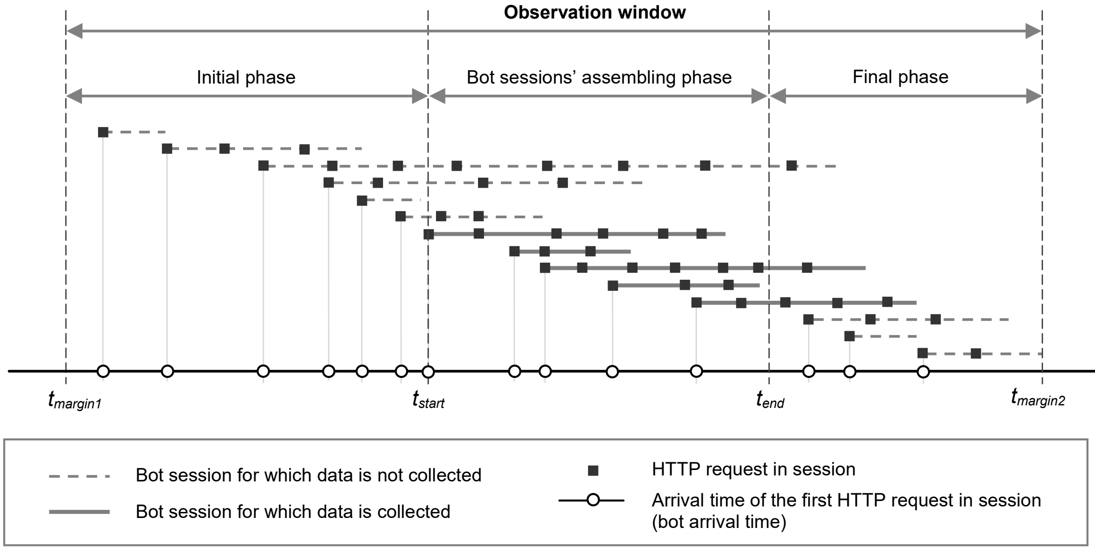

Observation window
Initial phase
Bot sessions' assembling phase
Final phase
$t_{margin1}$
t_{start}
t_{end}
$t_{margin2}$
time
Bot session for which data is not collected
Bot session for which data is collected
HTTP request in session
Arrival time of the first HTTP request in session (bot arrival time)

Figure 1: The Way of Collecting Bot Sessions' Data to Be Analyzed

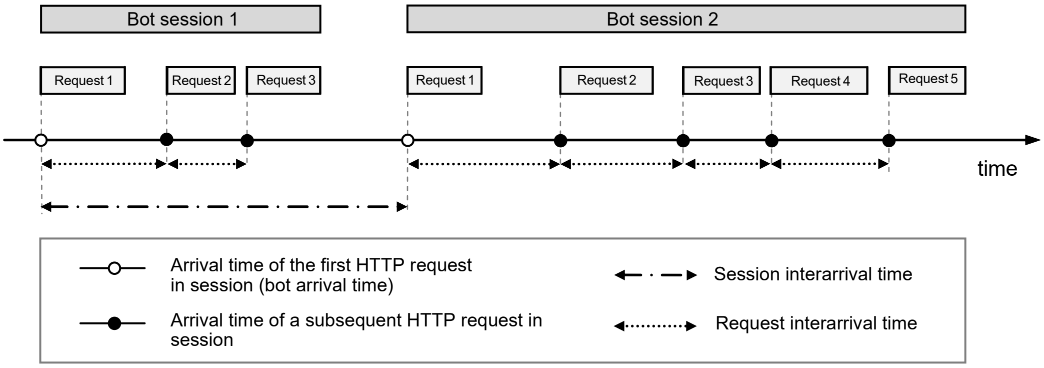

Bot session 1
Bot session 2
Request 1
Request 2
Request 3
Request 1
Request 2
Request 3
Request 4
Request 5
time
Arrival time of the first HTTP request in session (bot arrival time)
Arrival time of a subsequent HTTP request in session
Session interarrival time
Request interarrival time

Figure 2: The Way of Measuring Interarrival Times

A distributional model for each session feature for a given subset is developed using statistical analysis software in the following way. First, for empirical data contained in the subset a histogram is built and based on its shape some candidate regression functions are chosen. The suitability of the functions is then analyzed, resulting in the selection of the function $f(x)$ characterized by the highest correlation of the model with the empirical data.

Values of parameters of a regression function are estimated by applying the least squares method and optimized by applying the Nelder-Mead simplex method. The optimization criterion is minimization of the residual norm, given by the formula:

$$\delta = \|\hat{\mathbf{y}} - \mathbf{y}\| = \sqrt{\sum_{i=1}^{n}(\hat{y}_i - y_i)}, \qquad (1)$$

where:

n – the number of samples in a given subset,

y_i – the ith sample, $i = 1, 2, \ldots, n$,

$\hat{y}_i$ – the ith estimate of the regression function.

The suitability of a model to the empirical data is measured with the correlation coefficient, R^2, given by the formula:

$$R^2 = \left(\frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}} \right)^2, \qquad (2)$$

where:

$\bar{y}$ – the average over y_i,

$\bar{\hat{y}}$ – the average over $\hat{y}_i$.

For each of the three session features values of R^2 are calculated for all N subsets of samples (N days) and then the mean value of the coefficient for a potential model is determined. The optimal model for each feature is selected based on the highest mean value of R^2.

RESULTS OF MODELING A BOTS' ARRIVAL PROCESS AT A WEB SERVER

Data used for the analysis was log data of a Web server hosting a middle-sized online bookstore. The store software was osCommerce, a PHP-based electronic commerce platform.

A timespan of the used log data corresponded to the observation window ranging from 31st March 2014, 7:00 pm to 1st May 2014, 5:00 am with a 5-hour *initial phase* and a 5-hour *final phase*. Thus, the *bot sessions' assembling phase* covered the period from midnight preceding 1st April 2014 to midnight on 30th April 2014, i.e., 30 days ($N = 30$).

As a result of log data processing and analysis, a set of bot sessions was prepared. Basic information about the traffic on the server for the analyzed log data are given in Tab. 1. The entire set contained 42,782 Web bot sessions, in which 1,148,863 HTTP requests were sent to the server in total. Basic statistics for the three bot sessions' features are presented in Tab. 2.

Most active bots in terms of the number of sent requests were search engine crawlers (Googlebot, Yahoo! Slurp, MJ12bot, Bingbot, Nekstbot), SEO spybots (AhrefsBot and BLEXBot), online advertising bot Google AdsBot, e-commerce bots (ShopWiki, WillyBot, DotBot), and the Facebook bot called FacebookExternalHit.

Table 1: Basic Statistics for the Server Log Data (Timespan: 1st – 30th April 2014)

	All sessions	Bot sessions
Number of sessions	51,121	42,782 (83.7%)
Number of requests	3,200,228	1,148,863 (35.9%)
MB transferred	25,123.4	10,782.5 (42.9%)

Table 2: Statistics for Bot Sessions' Features

Statistics	**Session interarrival time [s]**	**Request interarrival time [s]**	**Number of requests in session**
Minimum	0	0	1
Maximum	6,178	1,799	47,624
Mean	61.0	28.2	26.9
Median	36	0	2
Std. dev.	83.3	142.0	242.4

As a result of the regression analysis of bot session features, three distributional models were best fitted to the corresponding empirical data: a Weibull model for the session interarrival time, a Pareto model for the request interarrival time and a model based on sigmoid and exponential type functions for the number of requests in session.

Session Interarrival Time

As regards the session interarrival time (*SIT*), the mean, equal to 61 seconds, is much higher than the median, equal to 36 seconds – in fact, almost 66% of samples are below the mean. It suggests that a distribution of the session interarrival time is right-skewed. This result confirms findings reported in previous studies, e.g. in (Doran and Gokhale 2010).

The distribution of the bot session interarrival time is modeled by a Weibull function given by the formula:

$$f_{SIT}(x) = A\frac{k}{\lambda}\left(\frac{x+\mu}{\lambda}\right)^{k-1} e^{-((x+\mu)/\lambda)^{k}}, \qquad (3)$$

where:

A – the amplitude parameter,
λ – the scale parameter, $\lambda > 0$,
k – the shape parameter, $k > 0$,
μ – the location parameter.

The independent variable x is the value of the time interval in seconds, $x \in \mathbb{N}$.

The shape of the regression curve differs slightly between the 30 subsets (measurement days) so the parameters of (3) are different for each day. Fig. 3 presents a histogram of session interarrival times along with the probability density function (PDF) of the estimated distribution for one sample day (27th April). To improve the clarity of the graph only 400 first bars of the histogram are shown in the figure in a semi-log scale.

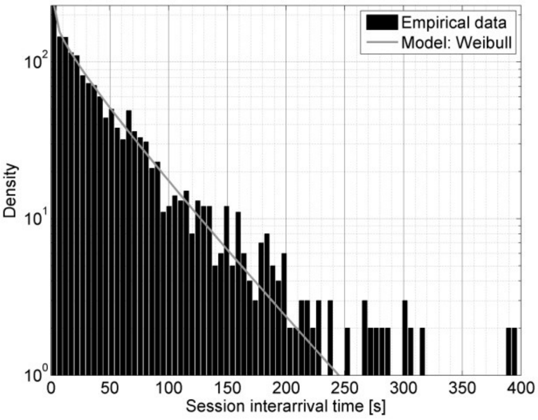


Figure 3: Histogram of the Empirical Session Interarrival Times and PDF of the Estimated Distribution (a Semi-Log Plot)

The correlation coefficient, R^2, varies from 0.77 to 0.99 and its average over all 30 datasets is 0.87. This shows a good fit of the distribution model to the empirical data.

Request Interarrival Time

Request interarrival times (*RIT*) are much more differentiated than session interarrival times, with the mean of 28.2 seconds and the standard deviation of 142 seconds. It is worth noting that the median is zero – in reality, above 73% of samples are equal to zero. This is a consequence of the fact that inter-request times in session are usually very short, even of the order of milliseconds (especially between hits for embedded objects) and on the other hand, request timestamps are registered in Web server log with an accuracy of one second.

Fig. 4 presents a histogram of request interarrival times and PDF of the estimated distribution for one day (27th

April). To capture data in a heavy-tail of the distribution, y-axis is plotted on a logarithmic scale.

Due to the very large number of zero intervals the empirical distribution is extremely heavy-tailed, similarly to distributions fitted in previous studies, e.g. in (Calzarossa and Massari 2013). We model this distribution by a Pareto function given by the formula:

$$f_{RIT}(x) = A\frac{k(x-\mu)^k}{(x-\mu)^{k+1}}, \quad (4)$$

where:

- A – the amplitude parameter,
- k – the shape parameter, $k > 0$,
- μ – the location parameter.

The independent variable x is the value of the time interval in seconds, $x \in \mathbb{N}$.

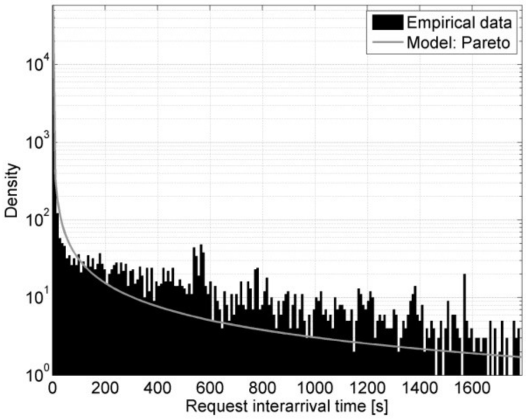


Figure 4: Histogram of the Empirical Request Interarrival Times and PDF of the Estimated Distribution (a Semi-Log Plot)

The correlation coefficient varies from 0.96 to 0.99 and its average over all 30 datasets is 0.99. So the correlation is even higher than that for the session interarrival time model.

Number of Requests in Session

Statistics for the numbers of HTTP requests in sessions (*NR*) generated by Web bots, included in Tab. 1, show that the minimum is one and the maximum is 47,624. In fact, as much as 41% of bot sessions contained only one request. The mean is 26.9 requests per session but the median is only two requests.

Fig. 5 shows a histogram of the numbers of requests in session (150 first bars only) and the corresponding PDF graph of the estimated distribution for one day (27th April). *Y*-axis is plotted on a logarithmic scale. One can notice in Fig. 5 that although the histogram is evidently right-skewed, there are some peaks, especially a very significant peak with a maximum value at about 120 requests per session (such a high peak is present in all 30 analyzed subsets).

We propose a model of the number of requests in bot session based on a regression function which is a combination of one sigmoid and two exponential type functions and is expressed by the formula:

$$f_{NR}(x) = e^{-\left(\frac{x-\mu_1}{c}\right)^2}\frac{A}{1+ke^{-b(x-\mu_2)}} + Be^{-\left(\frac{x-\mu_3}{d}\right)^2}, \quad (5)$$

where:

- k – the scale parameter of the sigmoid function, $k > 0$,
- b – the shape parameter of the sigmoid function,
- c, d – scale parameters of the first and the second exponential functions, respectively,
- A – the amplitude parameter of the sigmoid function,
- B – the amplitude parameter of the second exponential function,
- μ_1, μ_2, μ_3 – location parameters of the first exponential, sigmoid, and the second exponential functions, respectively.

The independent variable x is the number of requests in session, $x \in \mathbb{N}_+$.

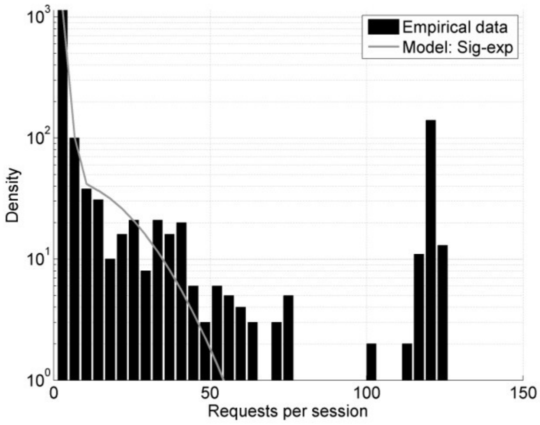


Figure 5: Histogram of the Empirical Numbers of HTTP Requests in Session and PDF of the Estimated Distribution for a Sample Day (a Semi-Log Plot)

This model is very well fitted to the empirical data as evidenced by the high correlation: R^2 ranges from 0.88 to 0.99 with the average equal to 0.97.

Summary of the Model

Since for each traffic feature each of the 30 datasets (measurement days) was analyzed and modeled separately, shapes of the regression curves differ slightly between the datasets. Thus, the parameters of the model for each feature are different for each day. Comparison of mean values of the correlation coefficient over all three features determined for each day showed that the highest overall correlation was achieved for the 27th April dataset (mean R^2 = 0.97). Therefore, parameters of the distribution functions in our model are given just for that day. The proposed model, including three features of Web bot arrival process, is summarized in Table 3.

The suitability of fit of the three distributional models to the empirical data is illustrated in Fig. 6. The best average fit was achieved for the request interarrival time model (R^2 equal to 0.99). The lowest average value of R^2, equal to 0.87, was achieved for the session interarrival time model.

Table 3: Model of a Web Bots' Arrival Process

Traffic feature	Distribution	Parameters
Session interarrival time [s]	Weibull (3)	A = 3,176.4 λ = 43.43 k = 0.86 μ = 0
Request interarrival time [s]	Pareto (4)	A = 1228.6 k = 2.50 μ = 1.43
Number of requests in session	Sig-exp (5)	k = -0.0003 b = 64801.55 c = 4.32 d = 27 A = 4988.86 B = 46.1 μ_1 = -2.45 μ_2 = 2.9 μ_3 = 1.48

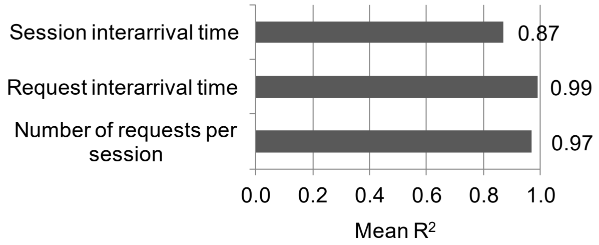


Figure 6: Suitability of Fit of the Model to the Empirical Data

VALIDATION OF THE MODEL

To check if modeled data will have a similar character to the real data, the proposed model was implemented in a C++ computer program. Numbers generated according to the model functions were recorded and then their distributions were compared with those of the corresponding empirical data (recorded in subsets for the 27th April). Results are shown in Fig. 7-9.

A visual inspection of the real data and the simulated results allows one to observe a very high degree of similarity between empirical and model data (note that the data is presented in a semi-log scale). One can notice some discrepancies, however. The Pareto model of the request interarrival time (Fig. 8), for which the highest mean R^2 was achieved, seems to produce too many extremely low values in the range of about 30 to 100 seconds and too few high values, exceeding 150 seconds. A big advantage of this model, however, is its ability to capture values in the distribution tail.

In the case of two other kinds of bot data, inter-session interarrival times generated according to the Weibull model (Fig. 7) and numbers of requests per session generated according to the sigmoid-exponential model (Fig. 9), one can notice that bodies of the distributions are very well fitted but the distribution tails are not well reflected.

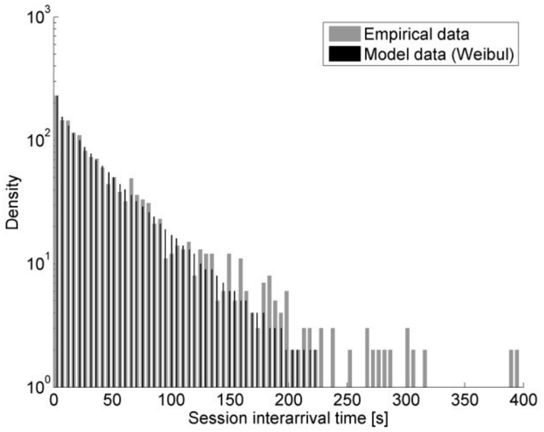


Figure 7: Histograms of the Session Interarrival Times for the Weibull Model Data and the Real Data (a Semi-Log Plot)

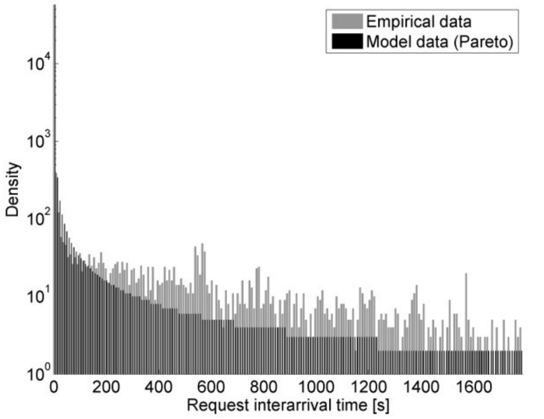


Figure 8: Histograms of the Request Interarrival Times for the Pareto Model Data and the Real Data (a Semi-Log Plot)

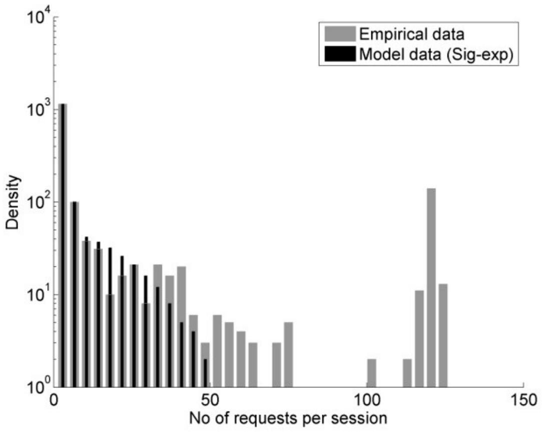


Figure 9: Histograms of the Numbers of Requests in Session for the Sigmoid-Exponential Model Data and the Real Data (a Semi-Log Plot)

A possible reason for these discordances may be the fact that no outliers have been eliminated before the analyses and the presence of the outliers might have negatively affected the resulting distributions of the modeled data. Besides, in the case of the numbers of requests per session the accurate modeling of a distribution peak with a maximum value at about 120 requests in session would probably require a combination of the bigger number of functions.

CONCLUSIONS AND FUTURE WORK

A result of the study presented in the paper is a mathematical model of a session-based bots' arrival process at an e-commerce Web server. The advantage of the model is that it is completely session-based and was developed based on real log data obtained from an online retailer. The model may be implemented in a discrete-event simulator, where the Web traffic is generated by many independent traffic sources (including Web bots). Thus, a feedback-based interaction of the server with the clients may be simulated.

A comparison of histograms for the real data and simulated results shows that the estimates generated according to the proposed model are characterized by the similar distributions as real data. However, further research is required to improve the quality of the model, to increase the value of the correlation coefficient and to capture data in heavy tails of the distributions.

Although this is a preliminary study, its results proved the efficiency of the proposed approach and provided some conclusions for our future work. The first issue to improve is connected with the coarse time resolution of log files, which causes most inter-request times in session to be equal to zero. Thus, it seems reasonable to eliminate zero intervals from the set of the analyzed request interarrival times by converting them to values of the order of milliseconds. Similarly, due to the high number of bot sessions containing only one request, it would be worth separating one-request sessions from the longer ones and to analyze the longer sessions separately. Another step which could improve the suitability of fit of the three distributional models to the empirical data might be the elimination of outliers.

The developed models describing a bots' arrival process on a Web server may be used in a traffic generator at the input of a Web server simulator. Such a simulation tool will make it possible to test Web service degradation under various load levels. We leave these issue to our future work.

ACKNOWLEDGEMENT

Grażyna Suchacka is a member of the consortium ICT COST Action IC1406 "High-Performance Modeling and Simulation for Big Data Applications" (cHiPSet).

REFERENCES

Almeida, V.; D. Menascé; R. Riedi; F. Peligrinelli; R. Fonseca; and W. Meira Jr. 2001. "Analyzing Robot Behavior in E-Business Sites". In *Proceedings of ACM SIGMETRICS* (Cambridge, Massachusetts, USA, Jun.16-20). ACM, New York, NY, USA, 338-339.

Bomhardt C.; W. Gaul; and L. Schmidt-Thieme. 2005. "Web Robot Detection - Preprocessing Web Log Files for Robot Detection". In *New Developments in Classification and Data Analysis. Studies in Classification, Data Analysis, and Knowledge Organization*, H.-H. Bock et al. (Eds.). Springer, Berlin-Heidelberg, 113-124.

Calzarossa, M.C. and L. Massari. 2013. "Temporal Analysis of Crawling Activities of Commercial Web Robots." *Computer and Information Sciences* III, 429-436.

Catledge, L.D. and J.E. Pitkow. 1995. "Characterizing Browsing Strategies in the World-Wide Web." *Computer Networks and ISDN Systems* 27, No.6, 1065-1073.

Dikaiakos, M.D.; A. Stassopoulou; and L. Papageorgiou. 2005. "An Investigation of Web Crawler Behavior: Characterization and Metrics." *Computer Communications* 28, No.8, 880-897.

Doran, D. and S.S. Gokhale. 2010. "Searching for Heavy Tails in Web Robot Traffic". In *Proceedings of the 7th International Conference on the Quantitative Evaluation of Systems* QEST'10 (Williamsburg, Virginia, USA, Sep.15-18). IEEE, Piscataway, N.J., 282-291.

Doran, D.; K. Morillo; and S.S. Gokhale. 2013. "A Comparison of Web Robot and Human Requests". In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* ASONAM'13 (Niagara, ON, Canada, Aug.25-29). IEEE, Piscataway, N.J., 1374-1380.

Geroimenko, V. 2004. *Dictionary of XML Technologies and the Semantic Web*. Springer-Verlag, London, UK.

Kołaczek, G. and K. Juszczyszyn. 2012. "Traffic Pattern Analysis for Distributed Anomaly Detection". In *Proceedings of the 9th International Conference on Parallel Processing and Applied Mathematics* PPAM'12 (Toruń, Poland, Sep.11-14, 2011), *Lecture Notes in Computer Science* 7204, R. Wyrzykowski; J. Dongarra; K. Karczewski and J. Waśniewski (Eds.). Springer, Berlin-Heidelberg, 648-657.

Lee, J.; S. Cha; D. Lee and H. Lee. 2009. "Classification of Web Robots: An Empirical Study Based on Over One Billion Requests." *Computers & Security* 28, No.8, 795-802.

Liu, L.; S. Chen; G. Yan; and Z. Zhang. 2008. "BotTracer: Execution-Based Bot-Like Malware Detection". In *Proceedings of the 11th International Conference on Information Security* (Taipei, Taiwan, Sep.15-18, 2008), *Lecture Notes in Computer Science* 5222. Springer, Berlin-Heidelberg, 97-113.

Saputra, C.H.; E. Adi; and S. Revina. 2013. "Comparison of Classification Algorithms to Tell Bots and Humans Apart." *Journal of Next Generation Information Technology* 4, No.7, 23-32.

Skrzewski, M. 2014. "System Network Activity Monitoring for Malware Threats Detection". In *Proceedings of the International Conference Computer Networks* (Lwówek Śląski, Poland, Jun.23-27), *Communications in Computer and Information Science* 431, A. Kwiecień; P. Gaj and P. Stera (Eds.). Springer, Berlin-Heidelberg, 138-146.

Skrzewski M. 2016. “About the Efficiency of Malware Monitoring via Server-Side Honeypots”. In *Proceedings of the International Conference Computer Networks* CN’16 (Lwówek Śląski, Poland, Jun.14-17), *Communications in Computer and Information Science* 608, P. Gaj; A. Kwiecień and P. Stera (Eds.). Springer, Cham, 132-140.

Stassopoulou, A. and M.D. Dikaiakos. 2009. “Web Robot Detection: A Probabilistic Reasoning Approach.” *Computer Networks* 53, No.3, 265-278.

Stevanovic, D.; N. Vlajic; and A. An. 2011. “Unsupervised Clustering of Web Sessions to Detect Malicious and Non-Malicious Website Users.” *Procedia Computer Science* 5, 123-131.

Suchacka, G. 2014. “Analysis of Aggregated Bot and Human Traffic on E-Commerce Site.” In *Proceedings of Federated Conference on Computer Science and Information Systems* FedCSIS'14 (Warsaw, Poland, Sep.7-10), Annals of Computer Science and Information Systems (ACSIS), Vol. 2. IEEE, Piscataway, N.J., 1123-1130.

Suchacka, G. and M. Sobków. 2015. “Detection of Internet Robots using a Bayesian Approach”. In *Proceedings of the 2nd IEEE International Conference on Cybernetics* CYBCONF'15 (Gdynia, Poland, Jun.24-26). IEEE, Piscataway, N.J., 365-370.

AUTHOR BIOGRAPHIES

GRAŻYNA SUCHACKA received the M.Sc. degrees in Computer Science and in Management, as well as the Ph.D. degree in Computer Science from Wroclaw University of Technology, Poland. Now she is an assistant professor in the Institute of Mathematics and Informatics at Opole University, Poland. Her research interests include analysis and modeling of Web traffic, Web mining and Quality of Web Service with special regard to electronic commerce. Her e-mail address is: gsuchacka@uni.opole.pl.

DARIA WOTZKA received the M.Sc. degree in Computer Science from the Technische Universität Berlin, Germany and the Ph.D. degree in Electrical Engineering from the Opole University of Technology, Poland. She is a lecturer and research fellow at the Opole University of Technology, Poland. Her research interests include digital data processing, modeling and simulation of multiphysical phenomena occurring in electric power devices. Her e-mail address is: d.wotzka@po.opole.pl.

Probability and Statistical Methods for Modelling and Simulation of High Performance Information Systems

-

Special Session

Modelling of the underwater targets tracking with the aid of pseudomeasurements Kalman filter

Alexander B. Miller
IITP RAS

Boris M. Miller
IITP RAS

KEYWORDS

TMA, passive tracking, bearing-only observation, fusion of passive and active measurements, pseudomeasurements Kalman filter

ABSTRACT

Target motion analysis of the underwater target tracking by the UUV (Unmanned underwater vehicle) usually based on the bearing-only observations including azimuth and elevation angles. However, low angular resolution of hydro acoustic sonars is not enough for the good quality of tracking. Moreover, angular observations lead to nonlinear filtering such as Extended Kalman Filtering (EKF) which usually produces estimations with unknown bias and quadratic errors. As it was mentioned long ago in a case of bearing-only observations target unobservability may take place, therefore, some special observer's motion become necessary. Other filters like the particle or unscented ones need the additional computer resources and also may produce the tracking loss. At the same time the pseudomeasurements Kalman filtering (PKF) method which transforms the estimation problem to the linear one and gives the current coordinates estimation with almost same accuracy could be modified to evaluate the moving target coordinates and velocities without bias. Since PKF gives unbiased estimate for the motion and the quadratic error it provides the good means for integration of various measurements methods such as passive (bearing-only) and active (range) metering. Using this filtering approach the good quality of target motion analysis (TMA) for randomly moving target may be achieved.

INTRODUCTION

Typical observations of underwater targets nowadays are based on so-called hydroacoustic imaging which simultaneously provides various modes of the observations such as passive ones (bearing-only) and/or active (range metering) [Sullivan (2015)]. In first works related to this class of observations it was mentioned that in the tracking of moving targets on the basis of bearing-only observations the unobservability may occur [Nardone and Aidala (1981)]. In the so-called target motion analysis (TMA) it is usually assumed that the motion of target is almost deterministic (just the initial position and velocity are unknown). This leads to the estimation procedure like the least square method on the basis of relatively long period of observations [Nardone et al. (1984)]. Most of results in this area demonstrate the difficulty of high quality TMA without observer's movement [Helbling (1988)] in [Chan (1988)]. Moreover, as was shown in [Jauffret et al. (2008)], [Pignol et al. (2010)] it is typical situation without observer's maneuvering. So the usage of so-called optimal moving observer performs to minimize the tracking error and even to solve some additional tasks such as minimizing the proximity between pursuer and evader [Rubinovich (2001)], [Andreev and Rubinovich (2016)]. There are plenty of filtering approaches on the basis of bearing-only observations, such as the particle filter [Fei et al. (2008)], cubature Kalman filter [Leong et al. (2013)], [Xin-Chun and Cheng-Jun (2013)], unscented filters [Barisic et al. (2012)], some versions of interpolation filtering [(Gupta et al., 2015)] and many others [Xin et al. (2004)]. The principal specific feature of all these nonlinear filters is the unknown bias and impossibility of correct estimation of the quadratic error which is one of the principal advantages of classical Kalman filtering. Some approaches which permit in principle to avoid the presence of bias lead to the serious increasing of necessary computational means since they lead to the usage of multiple filtering and dynamic equations like in the case of unscented Kalman filtering [Wan and van der Merwe (2000)]. Meanwhile, the comparison of various versions of nonlinear Kalman filters shows that in case of bearing-only observations most of existing filters give almost the same level of the estimation accuracy [Lin et al. (2002)]. At the same time [Lin et al. (2002)] the pseudomeasurements filter is easier for implementation and usually more stable than others. However, this filter, as it was mentioned long ago, also produces a bias due to the nonlinear dependence of the noise variance from unobservable parameters [Aidala and Nardone (1982)]. Recently we developed the version of pseudomeasurements Kalman filter (PKF) based on the idea of the best linear estimation such as used in the conditionally optimal filtering of V. S. Pugachev [Pugachev and I. N (1987)], [Miller and Pankov (2007)]. This filter does not have the bias and gives the unbiased estimation of current square error which is extremely important in the data fusion of optical measurements and inertial navigation system of the UAV. This filter serves as a basis for the UAV control in GPS denied environment [Amelin and Miller (2014)], [Miller and Miller (2014a)], [Miller (2015)] where it permits to develop the navigation system based on the observation of so-called singular points on the earth surface [Konovalenko et al. (2015)] and comparison of their position with template map uploaded to the UAV before flight. This method may be extended to the usage of 3-dimensional template map which increases the accuracy of the altitude evaluation [Karpenko et al. (2015a)], [Karpenko et al. (2015b)]. This PKF is used for control of the UAV in GPS denied environment and for landing with the aid of terrain bearing tools like

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

optic and radio-locators [Miller and Miller (2015)]. In the case of underwater navigation 3D measurement are one of the most important tools since modern hydro-locators (sonars) produce 3D image of the sea floor which may be used for own position and navigation estimation [Zhang et al. (2014)].

The aim of this work is to show that PKF may be used for underwater targets tracking with relatively high accuracy and without special maneuvering which permits to avoid the unobservability. The phenomenon of unobsrvability was first mentioned in [Nardone and Aidala (1981)] for the case of rectilinear target motion. In most successive works authors are using the special maneuvering of the observer to reduce the estimation errors [Rubinovich (2001)], [Andreev and Rubinovich (2016)]. Indeed, the unobservability is inherent to the case of almost deterministic target motion, where just measuring of the azimuth angle does not permit to distinguish the real and one of possible target's motions. However, in the case of the random perturbations in the target motion such situation could occur with null probability only. We observed that PKF provides good quality of the target tracking on the basis of bearing-only observations without special maneuvering [Miller and Miller (2014b)], here we apply this method to the tracking of the underwater targets. Of course, PKF demonstrates the randomness in the estimation error, however, one can observe good correspondence of the real tracking error and its mean quadratic error. The PKF can be easily extended to the case of additional range metering, which gives, of course, much better estimation's quality, though may be unacceptable in some special areas of applications. Meanwhile, the addition of range metering in a reasonable way can increase the estimation precision and prevent the tracking loss, inherent to the bearing-only tracking. Therefore, in the real case of possible obscurity observation constraints it may be used together with bearing-only observations in the optimal manner, providing the balance between obscurity and the precision of the TMA.

MODELS FOR PSEUDOMEASUREMENTS KALMAN FILTER (PKF)

Model of the observer motion

We assume the pursuer motion described by three coordinates $X(t_k)$, $Y(t_k)$, $Z(t_k)$, velocities $V_x(t_k)$, $V_y(t_k)$, $V_z(t_k)$. At times $t_k = k\Delta t, \quad k = 1, 2, ...$ pursuer state vector

$$\mathbf{X}(t_k) = (X(t_k), Y(t_k), Z(t_k), V_x(t_k), V_y(t_k), V_z(t_k))^T$$

satisfies the following dynamic equation:

$$\mathbf{X}(t_{k+1}) = F\mathbf{X}(t_k) + B\mathbf{A}(t_k), \qquad (1)$$

where

$$\mathbf{A}(t_k) = (A_x(t_k), A_y(t_k), A_z(t_k))^T$$

is the vector of the pursuer accelerations, and the F, B matrix have the following form:

$$F = \begin{pmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} \frac{\Delta t^2}{2} & 0 & 0 \\ 0 & \frac{\Delta t^2}{2} & 0 \\ 0 & 0 & \frac{\Delta t^2}{2} \\ \Delta t & 0 & 0 \\ 0 & \Delta t & 0 \\ 0 & 0 & \Delta t \end{pmatrix}.$$

Model of the motion of maneuvering target

The stochastic target's motion is given by three coordinates $x(t_k)$, $y(t_k)$, $z(t_k)$, velocities $v_x(t_k)$, $v_y(t_k)$, $v_z(t_k)$ and accelerations $a_x(t_k)$, $a_y(t_k)$, $a_z(t_k)$, which are forming the 9-dimensional vector of the target motion components (TMC). At times $t_k = k\Delta t, \quad k = 1, 2, ...$ the TMC vector satisfies the following equation:

$$\mathbf{x}(t_{k+1}) = Q\mathbf{x}(t_k) + \mathbf{W}(t_k), \qquad (2)$$

where $\mathbf{W}(t_k)$ is a vector of motion current perturbations

$$\mathbf{W}(t_k) = (0, 0, 0, 0, 0, 0, \sigma_{\mathbf{x}} W_x(t_k), \sigma_{\mathbf{y}} W_y(t_k), \sigma_{\mathbf{z}} W_z(t_k))^T$$

and matrix Q:

$$Q = \begin{pmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha_x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha_y & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha_z \end{pmatrix}.$$

We assume also that the target accelerations are random and satisfy the following equations:

$$\begin{pmatrix} A_x(t_{k+1}) \\ A_y(t_{k+1}) \\ A_z(t_{k+1}) \end{pmatrix} = \begin{pmatrix} \alpha_x A_x(t_k) \\ \alpha_y A_y(t_k) \\ \alpha_z A_z(t_k) \end{pmatrix} \Delta t + \begin{pmatrix} \sigma_{\mathbf{x}} W_x(t_k) \\ \sigma_{\mathbf{y}} W_y(t_k) \\ \sigma_{\mathbf{z}} W_z(t_k) \end{pmatrix}. \qquad (3)$$

The coefficient $|\alpha| < 1$, therefore the target random accelerations constitute the stationary processes [Pugachev and I. N (1987)].

Model of bearing+range measurements

Here we give the general measurements model including asimuth, elevation and possible range metering in the universal form. At time t_k pursuer direction finder (DF) produces a set of three target position measurements (see Fig. 1). It measures azimuth angle $\phi(t_k)$:

$$I(t_k)\frac{y(t_k) - Y(t_k)}{x(t_k) - X(t_k)} = I(t_k)(\tan\phi(t_k) + \varepsilon_k^{\phi}), \qquad (4)$$

elevation angle $\lambda(t_k)$:

$$I(t_k)\frac{z(t_k)-Z(t_k)}{y(t_k)-Y(t_k)}\sin\phi(t_k) = I(t_k)(\tan\lambda(t_k)+\varepsilon_k^{\lambda}) \tag{5}$$

and range $l(t_k)$:

$$I(t_k)\frac{z(t_k)-Z(t_k)}{\sin\lambda(t_k)} = I(t_k)(l(t_k)+\varepsilon_k^{l}), \tag{6}$$

where one can assume that $\varepsilon_k^{\phi} \sim \mathcal{WN}(0,\sigma_\phi^2)$, $\varepsilon_k^{\lambda} \sim \mathcal{WN}(0,\sigma_\lambda^2)$, $\varepsilon_k^{l} \sim \mathcal{WN}(0,\sigma_l^2)$ are uncorrelated random variables with zero means and variances $\sigma_\phi^2, \sigma_\lambda^2, \sigma_l^2$, defined as errors in measurements of DF and forming the white noise ($\mathcal{WN}$) sequences. $I(t_k)$ is an indicator function which is equal to 1 if at time t_k target is in the coverage area of the pursuer DF and zero otherwise.

At given time t_k either angular or angular+range measurements may be used together. However, due to the obscurity constraints the range measurements may be used in case of necessity to get more precise measurements when the estimated quadratic error is less than given threshold or with larger period than angle measurement.

DEVELOPMENT OF UNBIASED PKF

Using the pseudomeasurements method [Aidala and Nardone (1982)], [Lin et al. (2002)] we separate in (4),(5,)(6) observable and unobservable values which gives the system of linear measurement equations

$$\mathbf{m_k} = \begin{pmatrix} m_k^{\phi} \\ m_k^{\lambda} \\ m_k^{l} \end{pmatrix} =$$

$$I(t_k)\begin{pmatrix} x(t_k)\sin\phi(t_k) - y(t_k)\cos\phi(t_k) \\ +\varepsilon_k^{\phi}\cos\phi(t_k)(x(t_k)-X(t_k)) \\ \\ y(t_k)\sin\lambda(t_k) - z(t_k)\sin\phi(t_k)\cos\lambda(t_k) \\ +\varepsilon_k^{\lambda}\cos\lambda(t_k)(y(t_k)-Y(t_k)) \\ \\ z(t_k) - l(t_k)\sin\lambda(t_k) - \varepsilon_k^{l}\sin\lambda(t_k) \end{pmatrix}. \tag{7}$$

We assume that at the moment t_k we have unbiased estimates of target position

$$\hat{\mathbf{x}}(t_k), \hat{P}(t_k) = E(\mathbf{x}(t_k)-\hat{\mathbf{x}}(t_k))(\mathbf{x}(t_k)-\hat{\mathbf{x}}(t_k))^T,$$

where $\hat{x}(t_k)$ is such that

$$E(\hat{\mathbf{x}}(t_k)) = \mathbf{x}(t_k). \tag{8}$$

Using the prediction-correction filter algorithm [Amelin and Miller (2014)], [Miller (2015)], [Miller and Pankov (2007)] we receive estimates $(\hat{\mathbf{x}}(t_{k+1}), \hat{P}(t_{k+1}))$ at times t_{k+1}, which satisfy the unbiasedness condition (8), using the estimates received at time t_k, measurements m_k, known pursuer coordinates and target dynamics (2).

Thus the prediction is obtained by assuming that at the moment t_{k+1} the value of $m(t_{k+1})$ will be known

$$\begin{aligned} \tilde{\mathbf{x}}(t_{k+1}) &= Q\hat{\mathbf{x}}(t_k), \\ \tilde{\mathbf{P}}(t_{k+1}) &= Q\hat{P}(t_k)Q^T + E\mathbf{W}(\mathbf{t_k})\mathbf{W^T}(\mathbf{t_k}), \\ \tilde{\mathbf{m}}_{k+1} &= \begin{pmatrix} \tilde{m}_{k+1}^{\phi} \\ \tilde{m}_{k+1}^{\lambda} \\ \tilde{m}_{k+1}^{l} \end{pmatrix}. \end{aligned} \tag{9}$$

After getting the measurements at the moment t_{k+1} one can obtain the estimate of the target position and velocity at this moment and the matrix of the mean square errors:

$$\begin{aligned} &\hat{\mathbf{x}}(t_{k+1}) = \tilde{\mathbf{x}}(t_{k+1}) \\ &+\tilde{\mathbf{P}}^{\mathbf{xm}}(t_{k+1})(\tilde{P}^{mm}(t_{k+1}))^{-1}(\mathbf{m}_{k+1}-\tilde{\mathbf{m}}_{k+1}), \end{aligned} \tag{10}$$

$$\begin{aligned} &\hat{P}(t_{k+1}) = \tilde{P}(t_{k+1}) \\ &-\tilde{\mathbf{P}}^{\mathbf{xm}}(t_{k+1})(\tilde{P}^{mm}(t_{k+1}))^{-1}\tilde{\mathbf{P}}^{\mathbf{xm}}(t_{k+1})^T, \end{aligned} \tag{11}$$

where the innovation process has a form

$$(\mathbf{m}_{k+1}-\tilde{\mathbf{m}}_{k+1}) =$$

$$I(t_{k+1})\begin{pmatrix} (x(t_{k+1})-\tilde{x}(t_{k+1}))\sin\phi(t_{k+1}) \\ -(y(t_{k+1})-\tilde{y}(t_{k+1}))\cos\phi(t_{k+1}) \\ +\varepsilon_{k+1}^{\phi}\cos\phi(t_{k+1})(x(t_{k+1})-X(t_{k+1})) \\ \\ (y(t_{k+1})-\tilde{y}(t_{k+1}))\sin\lambda(t_{k+1}) \\ -(z(t_{k+1})-\tilde{z}(t_{k+1}))\sin\phi(t_{k+1})\cos\lambda(t_{k+1}) \\ +\varepsilon_{k+1}^{\lambda}\cos\lambda(t_{k+1})(y(t_{k+1})-Y(t_{k+1})) \\ \\ z(t_{k+1})-\tilde{z}(t_{k+1})-l(t_{k+1})\sin\lambda(t_{k+1}) \\ -\varepsilon_{k+1}^{l}\sin\lambda(t_{k+1}) \end{pmatrix}.$$

Quadratic characteristics of elements $x(t_{k+1}) - X(t_{k+1})$ may be evaluated via representation

$$\begin{aligned} &x(t_{k+1}) - X(t_{k+1}) \\ &= [x(t_{k+1})-\tilde{x}(t_{k+1})] + [\tilde{x}(t_{k+1})-X(t_{k+1})], \end{aligned}$$

where the first difference can be evaluated via $\tilde{P}^{xx}(t_{k+1})$ and the second one is known at time instant t_{k+1}. Here we use the orthogonality of the best linear estimate and the linear space spanned on observations $\mathbf{m}_i, i = 1..k$. Thereby the elements of matrix

$$\tilde{\mathbf{P}}^{\mathbf{xm}}(t_{k+1}) = (\tilde{P}^{xm}(t_{k+1}), \tilde{P}^{ym}(t_{k+1}), ..., \tilde{P}^{a_z m}(t_{k+1}))^T$$

are given by relations

$$\left[\tilde{P}^{xm}(t_{k+1})\right]^T =$$

$$E\left[(x(t_{k+1}) - \tilde{x}(t_{k+1}))(\mathbf{m}_{k+1} - \tilde{\mathbf{m}}_{k+1})\right]$$

$$= \begin{pmatrix} \tilde{P}^{xx}(t_{k+1}) \sin\phi(t_{k+1}) - \tilde{P}^{xy}(t_{k+1}) \cos\phi(t_{k+1}) \\ \\ \tilde{P}^{xy}(t_{k+1}) \sin\lambda(t_{k+1}) \\ -\tilde{P}^{xz}(t_{k+1}) \sin\phi(t_{k+1}) \cos\lambda(t_{k+1}) \\ \\ \tilde{P}^{xz}(t_{k+1}) \end{pmatrix} \tag{12}$$

and similarly for $\tilde{P}^{ym}(t_{k+1}), ..., \tilde{P}^{a_z m}(t_{k+1})$.

Matrix $\tilde{P}^{mm}(t_{k+1})$ has the following form

$$\tilde{P}^{mm}(t_{k+1}) = \begin{pmatrix} E[a^2] & E[ab] & E[ac] \\ E[ab] & E[b^2] & E[bc] \\ E[ac] & E[bc] & E[c^2] \end{pmatrix}, \tag{13}$$

where

$$\begin{array}{l} a = m^{\phi}_{k+1} - \tilde{m}^{\phi}_{k+1}, \\ b = m^{\lambda}_{k+1} - \tilde{m}^{\lambda}_{k+1}, \\ c = m^{l}_{k+1} - \tilde{m}^{l}_{k+1}. \end{array} \tag{14}$$

Method of the expectation calculation in (13) had been presented in [Amelin and Miller (2014)], [Miller and Miller (2014b)], [Miller (2015)], and as an example we give

$$\begin{aligned} &E[a^2](t_{k+1}) = I(t_{k+1})\Big[\tilde{P}_{yy}(t_{k+1})cos^2\phi(t_{k+1}) \\ &-\tilde{P}_{xy}(t_{k+1})sin2\phi(t_{k+1}) + \tilde{P}_{xx}(t_{k+1})sin^2\phi(t_{k+1}) \\ &+\sigma^2_{\phi}cos^2(t_{k+1}) \\ &\times\left(X(t_{k+1}) - \hat{x}((t_k) - \hat{v}_x(t_k)\Delta t - \hat{a}_x(t_k)\frac{(\Delta t)^2}{2}\right)^2 \\ &+\tilde{P}_{xx}(t_{k+1})\Big] \end{aligned} \tag{15}$$

other terms in (13) may be obtained similarly.

MODELLING OF THE FUSION OF BEARING-ONLY AND RANGE METERING

Generally equations (10), (11) give the dependence of current estimate accuracy in terms of the observer motion and the observation strategy, that is possible integration of bearing-only and range metering. In this article we wish to find the answer on the question how to combine passive and active metering. So we use three different strategies

- bearing-only, so the estimation is based on the angle measurements only,
- bearing-only and range metering in various combination, such as range measurements with fixed period which is greater or equal to the period of bearing measurement,
- bearing-only and range metering, where the range metering is made when the current quadratic error is greater than prescribed threshold.

For each type of strategies we are performing the Monte-Carlo simulation of random target motion under the same parameters and the same observer's motion. The observer motion and position are supposed to be known without errors, while the target motion is random with stationary accelerations satisfying equations (2) with perturbations having:

$$\begin{array}{l} \sigma_{\mathbf{x}} = 0.025 * 9.8, \\ \sigma_{\mathbf{y}} = 0.025 * 9.8, \\ \sigma_{\mathbf{z}} = 0.005 * 9.8. \end{array}$$

The following conditions were used for the duration of tracking and the accuracy of angle and range measurements, where we give the standard deviations:

- $T = 120$;
- $\sigma_{\phi} = \sigma_{\lambda} = 0.01 \approx 0.5^o$.
- $\sigma^2_l = 5.0$
- average velocity of observer ≈ 15, average velocity of target ≈ 10;
- the radius of observer's sonar sensitivity is 1600,

here the time and metric figures are in arbitrary units.

Examples of the observer motion and the target motions with and without the observer maneuver are shown on Fig. 1. The target position tracking are given on Fig. 2 for bearing-only, and on Fig. 3 and 4 for bearing and maneuvering observer, for the target position and velocities, respectively. Corresponding examples of the target tracking with range metering ($\Delta T = 10$) and with threshold strategy with the threshold $SD_{pos} = 10$ are given on Fig. 5 and Fig. 6. One can observe that in bearing-only case possible tracking loss may occur. The joint bearing and range metering provides much more reliable tracking but may be unacceptable due to the energy and obscurity constraints. So the possible solution may be the rare range measurement with either fixed period or with feedback law when error of tracking is less than desired. One should stress that PKF gives reliable means for such feedback since it provides the unbiased estimate of quadratic error.

Results of comparison of various observation strategies are summarized in the following table. These results are obtained with Monte-Carlo modelling by 100 samples of the target motions. where

$$SD_{pos} = \sqrt{\hat{P}^{xx} + \hat{P}^{yy} + \hat{P}^{zz}}$$

mean square error of the target position estimation,

$$SD_{vel} = \sqrt{\hat{P}^{V_x V_x} + \hat{P}^{V_y V_y} + \hat{P}^{V_z V_z}}$$

calculated in PKF during the tracking of target and averaged over Monte-Carlo modelling. SD_x, SD_y, SD_z are the standard deviations of the tracking errors calculated along with tracking of target with the aid of PKF, and $SD_{V_x}, SD_{V_y}, SD_{V_z}$ are the corresponding standard deviations of the velocity estimation errors. On each of

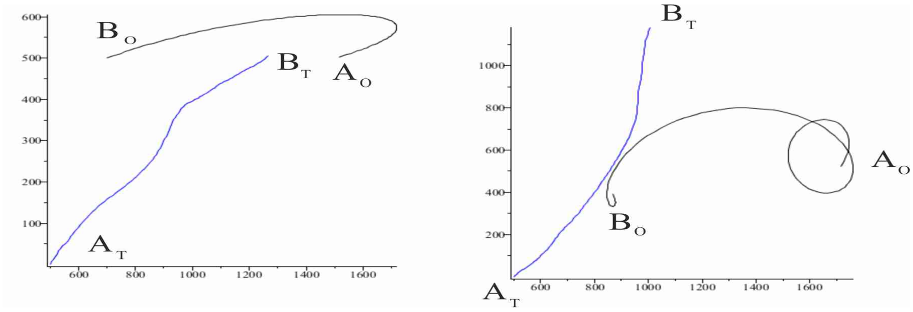


Figure 1: The Observer (black) and Target (blue) motion, A_O, B_O, A_T, B_T initial and final points of the observer and target positions. Left - without observer maneuvering, right - with observer maneuvering.

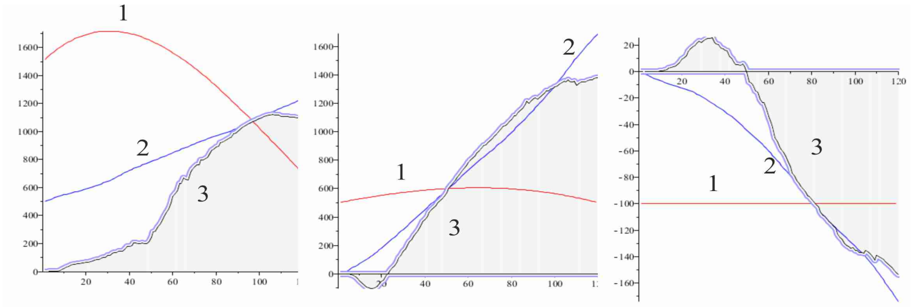


Figure 2: Tracking of the target position, bearing-only measurements without maneuvering. Left - X, center - Y, right - Z. 1 - observer position, 2 - (blue) real target position, 3 - target position estimation. There is a tendency to the tracking loss.

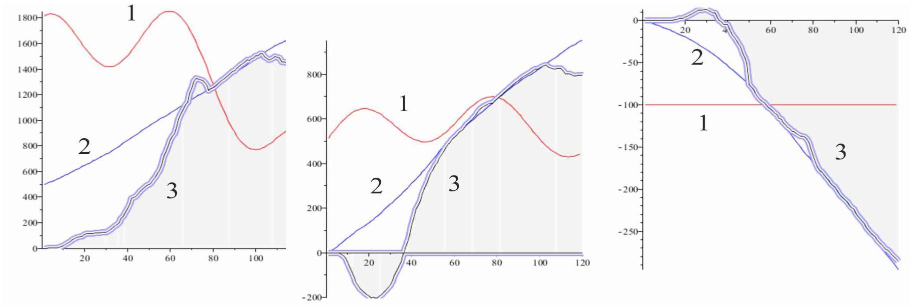


Figure 3: Tracking of the target position, bearing and maneuvering. Left - X, center - Y, right - Z. 1 - observer position, 2 - (blue) real target position, 3 - target position estimation. Even if in general the tracking is better than on Fig. 2, one can observe the tendency to the tracking loss.

the figures by gray color period of time when the target is in the scope of the pursuer DF is designated.

The main observation is that maneuvering target is rather difficult to track with typical hydro acoustic DF. But even if we assume less perturbations in the target motion the results of the estimation accuracy become worse. It is not surprising, but in accordance with the possibility of unobservability during the mission the tracking loss may occur. One can see this tendency to the tracking loss on Fig. 2 and on Fig. 6 at time instants $t = 50$ and $t = 100$. Meanwhile, the addition of range metering even rather rare (last line of the table and Fig. 7) could prevent the tendency to the tracking loss, which may be observed at time instants $t = 50$ and $t = 100$.

CONCLUSIONS

The modelling results show that PKF provides the reliable means for data fusion of bearing-only observations which are inherent to various passive tracking and measurement method with either active or passive measurements. In general:

- PKF has been developed for underwater targets tracking, the method may be used either in passive or active modes.

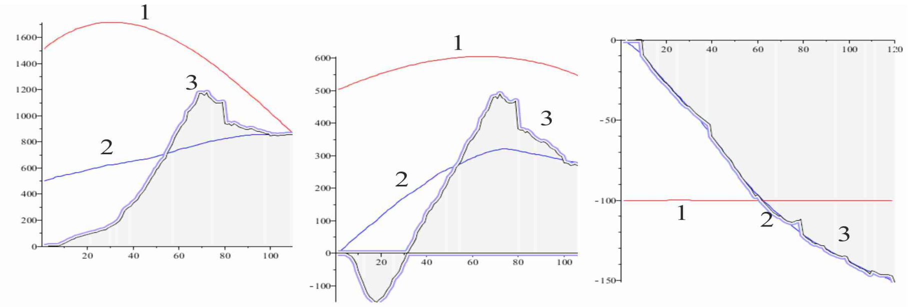


Figure 4: Tracking of the target position, bearing and range measurements with period $DT = 10$. Left - X, center - Y, right - Z. 1 - the observer position, 2 - (blue) the real target position, 3 - the target position estimation.

Mean squared error of tracking and velocities estimation

Method of measurements	SD_{pos}	SD_{vel}	SD_x	SD_y	SD_z	SD_{V_x}	SD_{V_y}	SD_{V_z}
Bearing-only (azimuth and elevation)	63.42	2.79	26.86	21.57	3.55	1.19	0.62	0.086
Bearing + the observer maneuvering	58.5	2.6	40	27	7.3	1.33	0.92	0.29
Bearing + range measurements	33.23	2.21	6.71	6.88	0.038	0.299	0.239	0.012
Bearing + range (threshold 20)	35.47	2.28	8.9	7.73	0.82	0.41	0.26	0.032
Bearing + range (threshold 10)	33.63	2.24	8.36	5.97	0.125	0.34	0.21	0.016
Bearing + range ($\Delta t = 10$)	34.63	2.23	9.12	5.54	0.73	0.35	0.19	0.018
Bearing + rare range ($\Delta t = 50$)	32.67	2.32	71.26	15.23	11.41	2.98	1.1	0.31

Figure 5: Mean squared error of tracking and velocities estimation in various modes of measurements

- Active modes provide essentially higher level of accuracy, however the combination of both modes may be rather effective in the case of obscurity constraints.
- Special maneuvering amends relatively to the bearing-only case, but not radically. Much better results give even very rare range metering which, however, must be coordinated with the obscurity constraints.

ACKNOWLEDGEMENTS

This research was conducted at the A. A. Kharkevish Institute for Information Transmission Problems, RAS, IITP RAS 127994, 19, Bolshoy Karetny per., Moscow, Russia E-mails: amiller@iitp.ru, bmiller@iitp.ru and partially supported by the Russian Basic Research Foundation grant 16-31-60049.

REFERENCES

Aidala, V. J. and S.C. Nardone. 1982. "Biased estimation properties of the pseudolinear tracking filter." *IEEE Transactions on Aerospace Electronic Systems*, 18(4):432–441.

Amelin, K. S. and A. B. Miller. 2014. "An algorithm for refinement of the position of a light uav on the basis of kalman filtering of bearing measurements." *Journal of Communications Technology and Electronics*, 59(6):622–631.

Andreev, K. V. and E. Ya. Rubinovich. 2016. "Moving observer trajectory control by angular measurements in tracking problem." *Automation and Remote Control*, 77(1):106–129.

Barisic, M.; A. Vasilijevic; and D. Nad. 2012. "Sigma-point unscented kalman filter used for auv navigatio." *20th Mediterranean Conference on Control & Automation (MED) Barcelona, Spain*, pages 1365–1372.

Chan, Eds. Y. T. 1988. "Underwater acoustic data processing nato asi series." 161.

Fei, Zhang; Zhou Xing-peng; Chen Xiao-hui; and Liu Rui-lan. 2008. "Particle filter for underwater bearings-only passive target tracking." *Proceedings of 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, pages 847–851.

Gupta, R.; A. Kumar; I. N. Kar; and R. Bahl. 2015. "Bearings-only tracking of non-maneuvering target with missing bearings data." *Proceedings of International Symposium on Underwater Technology 2015*, pages 1–7.

Helbling, L. F. 1988. "Bearing only target motion analysis." *Underwater Acoustic Data Processing*, 161:485–489.

Jauffret, C.; D. Pillon; and A.-C. Pignol. 2008. "Bearings-only tma without observer maneuver." *Proceedings of 11th International Conference on Information Fusion, June 2008*, pages 1–8.

Karpenko, S.; I. Konovalenko; A. Miller; B. Miller; and D. Nikolaev. 2015a. "Visual navigation of the uavs on the basis of 3d natural landmarks." *Proceedings of the 8th International Conference on Machine Vision (ICMV 2015), Barcelona, Spain.*

Karpenko, S.; I. Konovalenko; A. Miller; B. Miller; and D. Nikolaev. 2015b. "Visual navigation of the uavs on the basis of 3d natural landmarks." *Proceedings* *of the 8th International Conference on Machine Vision (ICMV 2015), Barcelona, Spain*, 9875:1–10.

Konovalenko, I.; A. Miller; B. Miller; and D. Nikolaev. 2015. "Uav navigation on the basis of feature points detection on underlying surface." *Proceedings of 29th EUROPEAN CONFERENCE ON MODELLING AND SIMULATION, ECMS 2015, Albena (Varna), Bulgaria.*

Leong, Pei H.; S. Arulampalam; T. A. Lamahewa; and T. D. Abhayapala. 2013. "A gaussian-sum based cubature kalman filter for bearings-only tracking." *IEEE Transaction on Aerospace and Electronic Systems*, AES-49(2):1161–1176.

Lin, X.; T. Kirubarajan; Y. Bar-Shalom; and S. Maskell. 2002. "Comparison of EKF, pseudomeasurement and particle filters for a bearing-only target tracking problem." *Proc. SPIE Int. Soc. Optic. Engin.*, 4728:240–250.

Miller, A. and B. Miller. 2014a. "Tracking of the uav trajectory on the basis of bearing-only observations." *Proceedings of 53rd IEEE Conference on Decision and Control Los Angeles, California, USA*, 59(6):622–631.

Miller, A. and B. Miller. 2014b. "UAV control on the basis of bearing-only observations." *Proceedings of Australian Control Conference, 17-18 November, 2014, Canberra*, pages 31–36.

Miller, A. and B. Miller. 2015. "Stochastic control of light uav at landing with the aid of bearing-only observations." *Proceedings of the 8th International Conference on Machine Vision (ICMV 2015), Barcelona, Spain.*

Miller, A. B. 2015. "Development of the motion control on the basis of kalman filtering of bearing-only measurements." *Automation and Remote Control*, 76(6):1018–1035.

Miller, B. M. and A. R. Pankov. 2007. "Theory of random processes [in russian]." *Moscow, Nauka, Fizmatlit.*

Nardone, S.; A. Lindgren; and Gong Kai. 1984. "Fundamental properties and performance of conventional bearings-only target motion analysis." *IEEE Trans. Automatic Control*, 29(9):775–787.

Nardone, S. C. and V. J. Aidala. 1981. "Observability criteria for bearings-only target motion analysis." *IEEE Transactions on Aerospace and Electronic Systems*, AES-17(2):162–166.

Pignol, A.-C.; C. Jauffret; and D. Pillon. 2010. "A statistical fusion for a leg-by-leg bearings-only tma without observer maneuver." *Proceedings of 13th International Conference on Information Fusion, July 2010*, pages 1–8.

Pugachev, V. S. and Sinitsyn I. N. 1987. "Stochastic differential systems. analysis and filtering." *Wiley 1987.*

Rubinovich, E. Ya. 2001. "Trajectory control over bearings-only observations in one r-encounter problem." *Proceedings of IFAC Symposium on Nonlinear Control Systems, St Petersburg, Russia, NOLCOS 2001.*

Sullivan, E. J. 2015. "Model-based processing for underwater acoustic arrays." *Springer.*

Wan, E. A. and R. van der Merwe. 2000. "The unscented kalman filter for nonlinear estimation." *Proceeding of IEEE Symposium 2000 (AS-SPCC), Lake Louise, Alberta, Canada*, pages 1–7.

Xin, G.; Yi Xiao; and H. You. 2004. "Bearings-only underwater track fusion solutions with feedback information." *Proceedings of ICSP'O4*, 3:2449–2452.

Xin-Chun, Zhang and Guo Cheng-Jun. 2013. "Cubature kalman filters: Derivation and extension." *Chin. Phys. B*, 22(12):128401.

Zhang, H.; D. Laneuville; B. de Saporta; A. Nègre; and F. Dufour. 2014. "Optimal trajectories for underwater vehicles by quantization and stochastic control." *Proceedings of 17th International Conference on Information Fusion (FUSION)*, pages 1–8.

AUTHOR BIOGRAPHIES

ALEXANDER B. MILLER was born in Krasnogorsk, Russia and went to the Moscow Institute of Physics and Technology (State University), where he obtained his degree in 2008. In 2012 he obtained PhD in the Institute for Information Transmission Problems RAS where he is working till now in the Laboratory of Image Analysis and Processing. His e-mail address is: amiller@iitp.ru and his Web-page can be found at http://iitp.ru/en/users/449.htm.

BORIS M. MILLER was born in Igevsk, Russia, graduated from Moscow Institute of Physics and Technology (State University) in 1974. He obtained PhD in 1978 and doctor of sciences (mathematics) in 1991. Now he is a principal research fellow in IITP RAS, professor. His e-mail address is: bmiller@iitp.ru.

APPROACHES TO STOCHASTIC MODELING OF WIND TURBINES

Migran N. Gevorkyan,
Anastasiya V. Demidova
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
Email: gevorkyan_mn@rudn.university,
demidova_av@rudn.university

Ivan S. Zaryadov
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
and Institute of Informatics Problems,
FRC CSC RAS, IPI FRC CSC RAS,
44-2 Vavilova str., Moscow, 119333, Russia
Email: zaryadov_is@rudn.university

Robert A. Sobolewski
Department of Power Engineering,
Fotonics and Lighting Technology,
Bialystok University of Technology,
45D Wiejska str., 15-351 Bialystok, Poland
Email: r.sobolewski@pb.edu.pl

Anna V. Korolkova
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
Email: korolkova_av@rudn.university

Dmitry S. Kulyabov
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
and Laboratory of Information Technologies
Joint Institute for Nuclear Research
6 Joliot-Curie, Dubna,
Moscow region, 141980, Russia
Email: kulyabov_ds@rudn.university

Leonid A. Sevastianov
Department of Applied Probability and Informatics,
RUDN University
(Peoples' Friendship University of Russia),
6 Miklukho-Maklaya str., Moscow, 117198, Russia
and Bogoliubov Laboratory of Theoretical Physics
Joint Institute for Nuclear Research
6 Joliot-Curie, Dubna,
Moscow region, 141980, Russia
Email: sevastianov_la@rudn.university

KEYWORDS

Weibull distribution, approximation, lognormal distribution, gamma distribution, beta distribution, wind speed, statistics

ABSTRACT

Background. This paper study statistical data gathered from wind turbines located on the territory of the Republic of Poland. The research is aimed to construct the stochastic model that predicts the change of wind speed with time. **Purpose.** The purpose of this work is to find the optimal distribution for the approximation of available statistical data on wind speed. **Methods.** We consider four distributions of a random variable: Log-Normal, Weibull, Gamma and Beta. In order to evaluate the parameters of distributions we use method of maximum likelihood. To assess the the results of approximation we use a quantile-quantile plot. **Results.** All the considered distributions properly approximate the available data. The Weibull distribution shows the best results for the extreme values of the wind speed. **Conclusions.** The results of the analysis are consistent with the common practice of using the Weibull distribution for wind speed modeling. In the future we plan to compare the results obtained with a much larger data set as well as to build a stochastic model of the evolution of the wind speed depending on time.

INTRODUCTION

This work is devoted to the problem of stochastic modeling of speed of wind, which is used to generate electrical power in wind plants located on the territory of the Republic of Poland. As a first step several distributions for accuracy of the wind speed approximation will be examined. For this purpose Log-normal, Weibull, Gamma and Beta are chosen. All these distributions have shape-location-scale parametrisation. For statistical data processing the authors used Python 3 with `numpy`, `scipy.stats` (see Jones et al. (2001)) and `matplotlib` (see Droettboom et al. (2017)) libraries and also `Jupyter` (see *Project Jupyter home* (2017))—an interactive shell. We used books (see Norman L. Johnson (1994, 1995); Nelson (1982)) as reference materials for distributions properties. Articles

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

(see Frchet (1927); Weibull (1951)) are the primary sources in which the Weibull distribution is presented for the first time. Articles (see Lun and Lam (2000); Seguro and Lambert (2000); Bowden et al. (1983); Yeh and Wang (2008); Islam et al. (2011); Garcia et al. (1998)) describe the use of the Weibull distribution in the modeling of wind turbines and wind speed.

THE DESCRIPTION OF THE STATISTICAL DATA STRUCTURE

The set of statistical data is stored in the file `csv` consisting of the following columns:

1) T — time of fixation of wind speed and direction by sensors installed on the wind power turbine (hh:mm format);
2) X_1 — output power of wind turbine [kW] (the negative values mean the power is consumed rather then generated);
3) X_2 — wind speed [m/s] (measured by anemometer installed at the top of wind turbine nacelle);
4) X_3 — wind direction [deg] (measured by anemometer installed at the top of wind turbine nacelle; measured clockwise, the value 0 to the N);
5) X_4 — wind speed 10 m [m/s] obtained at 10 m above the ground m;
6) X_5 — wind direction 10 m [deg] (obtained at 10 m above the ground; measured clockwise, the value from 0 to the N);
7) X_6 — wind speed 50 m [m/s] (obtained at 50 m above the ground);
8) X_7 — wind direction 50 m [deg] (obtained at 50 m above the ground; measured clockwise, the value ftom 0 to the N).

The indicators of wind speed and direction were read out from the sensors every 10 minutes for about 9 months. In total, the table contains 39606 entries.

To make an initial choice of distributions that may be suitable for wind speed approximation, the histograms of wind speed are drawn. Visual assessment of these histograms suggest that the adequate choice will be a "heavy-tailed" distribution. But for the wind direction approximation these distributions are not suitable, as can be seen from the figure 1.

To read out the data we used the function `genfromtxt` from `numpy` (see Jones et al. (2001)) lib:

```
ws1, ws2, ws3 =
    np.genfromtxt('data.csv',
    delimiter=';', skip_header=True,
    usecols=(2, 4, 6), unpack=True)
```

where `'data.csv'` is data file, `delimiter=';'` is columns separator, `skip_header = True` specifies ignoring of the first line as the names of the columns, `usecols=(2, 4, 6)` makes function to use only 2, 4, 6 columns (numbering begins with zero) and `unpack=True` — contents of each column should be written in separate arrays `ws1`, `ws2` and `ws3` for further analysis of the data separately.

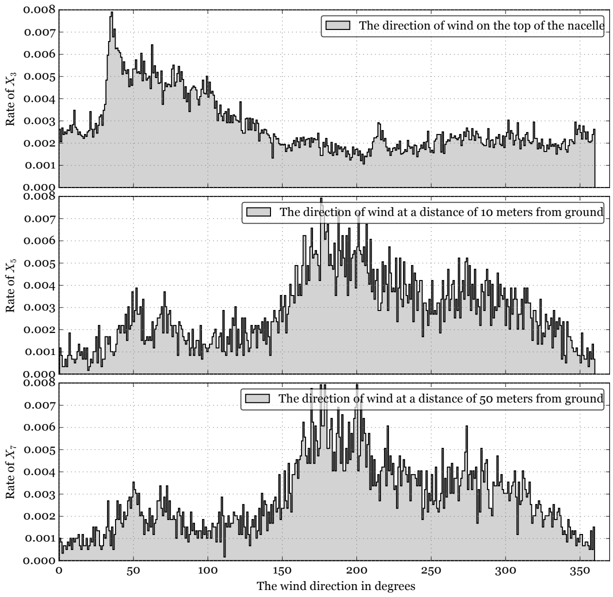


Fig. 1. Histogram of wind direction at three levels of height

PROBABILITY DISTRIBUTIONS

Each of distributions is parameterized by three parameters: α — shape factor, l — location factor and s — scale factor. In the case of the beta distribution the second scale factor is added, denoted by β-letter. All distributions parameters are positive real numbers: $\alpha, \beta, s, l \in \mathbb{R}$, $\alpha, \beta, s > 0$, $l \geqslant 0$.

The probability density function (PDF) of a Log-Normal random variable X is:

$$f_{LN}(x; \alpha, l, s) = \\ = \begin{cases} \dfrac{1}{(x-l)\alpha\sqrt{2\pi}} \cdot \\ \quad \cdot \exp\left(-\dfrac{1}{2}\left(\dfrac{\ln(x-l)-\ln s}{\alpha}\right)^2\right), & x \geqslant l. \\ 0, & x < l. \end{cases}$$

The probability density function of a Weibull (see Frchet (1927); Weibull (1951)) random variable X is:

$$f_W(x; \alpha, l, s) = \\ = \begin{cases} \dfrac{\alpha}{s}\left(\dfrac{x-l}{s}\right)^{\alpha-1} \exp\left[-\left(\dfrac{x-l}{s}\right)^{\alpha}\right], & x \geqslant l, \\ 0, & x < l. \end{cases}$$

The probability density function of a Gamma random variable X is:

$$f_\Gamma(x; \alpha, l, s) = \begin{cases} \dfrac{(x-l)^{\alpha-1}\exp\left(-\frac{(x-l)}{s}\right)}{s^\alpha\Gamma(\alpha)}, & x \geqslant l, \\ 0, & x < l. \end{cases}$$

where $\Gamma(\alpha)$ is gamma-function.

The probability density function of a Beta random variable X is:

$$f_{\mathcal{B}}(x;\alpha,\beta,l,s) =$$
$$= \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{s\Gamma(\alpha)\Gamma(\beta)}\left(\dfrac{x-l}{s}\right)^{\alpha-1}\left(1-\dfrac{x-l}{s}\right)^{\beta-1}, & x \geqslant l, \\ 0, & x < l. \end{cases}$$

If in PDF formulas of Log-Normal, Weibull and Gamma distributions let $l = 0$, and for Beta distribution let $s = 1$, we get the formulas of distributions most frequently used in Norman L. Johnson (1994); Nelson (1982).

DETERMINATION OF DISTRIBUTIONS PARAMETERS

In `scipy.stats` (see Jones et al. (2001)) following objects are defined: `lognorm`, `weibull_min`, `gamma` and `beta`. These objects implement distributions we work with. Every one of these objects has PDF function `pdf(x, a, [b,] loc, scale)` and CDF (cumulative distribution) function `cdf(x, a, [b,] loc, scale)`, where `x` — function argument, `a`, `b` — shape parameters α, (and β for Beta-distribution), `loc` and `scale` are location and scale parameters.

For parameters estimation of our distributions the library `scipy.stats` provides the function `fit(data)`, which calculates the parameters of distributions by maximum likelihood method and the empirical data. We used this function to calculate parameters of the considered distributions. Then we used `pdf` and `cdf` functions to compute values of the probability density function and cumulative distribution function.

There is the example of the code for the case of Log-Normal distribution:

```
s, loc, scale =
    scipy.stats.lognorm.fit(ws1)
xs = np.linspace(np.min(ws1),
    np.max(ws1), 1000)
logN_PDF =
    scipy.stats.lognorm.pdf(xs, s,
    loc, scale)
logN_CDF =
    scipy.stats.lognorm.cdf(xs, s,
    loc, scale)
```

The results are presented graphically on figures 2–9.

The figures were plotted for theoretical distributions, the parameters of which have been determined on the basis of the entire dataset. From the analysis of the quantile-quantile plots (Q-Q plots) we can conclude that the Weibull distribution is best suited for approximation of available data (although only slightly), outmatching them only in the approximation of extreme values of a random variable.

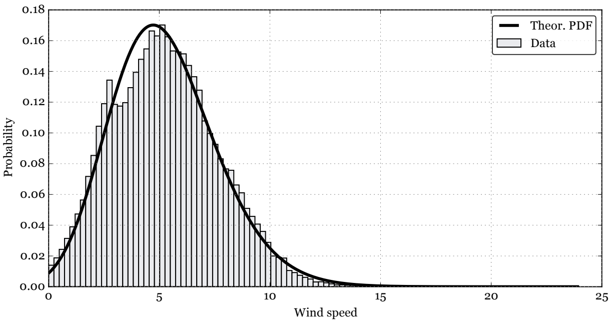


Fig. 2. PDF of **Log-Normal** distribution compared with data histogram

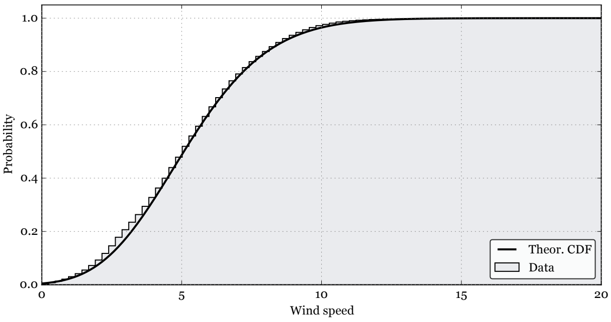


Fig. 3. CDF of **Log-Normal** distribution compared with empirical distribution function

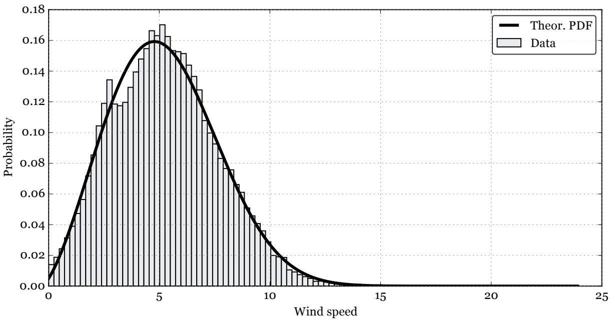


Fig. 4. PDF of **Weibull** distribution compared with data histogram

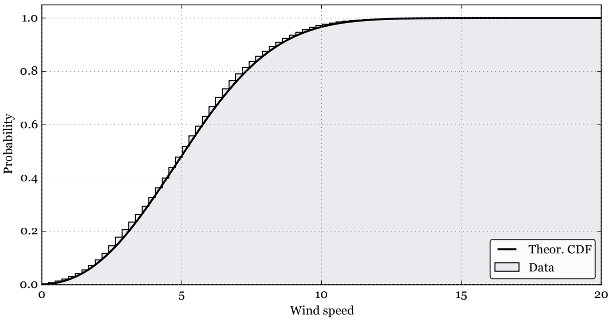


Fig. 5. CDF of **Weibull** distribution compared with empirical distribution function

We also performed computations with the considered distributions parameterized by only two parameters (let $l = 0$, and for Beta distribution an addition let $s = 1$). After plotting the results of calculations we found out that the two-parameter Weibull distribution has superiority over other two-parameters distributions (Log-Normal, Gamma and Beta), which is not true for

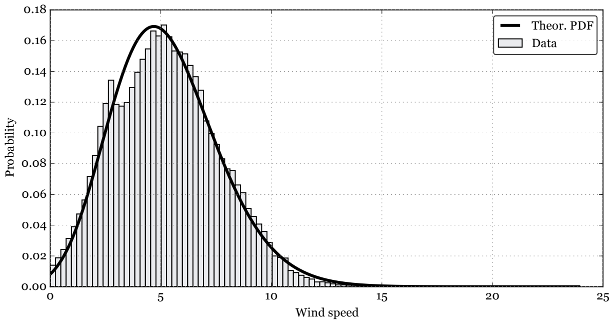


Fig. 6. PDF of **Gamma** distribution compared with data histogram

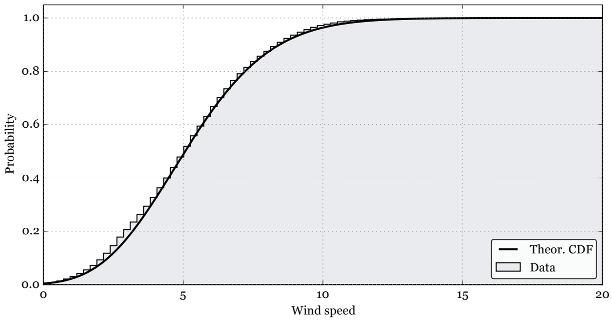


Fig. 7. CDF of **Gamma** distribution compared with empirical distribution function

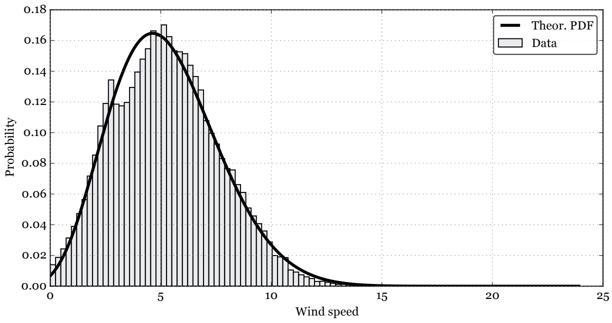


Fig. 8. PDF of **Beta** distribution compared with data histograms

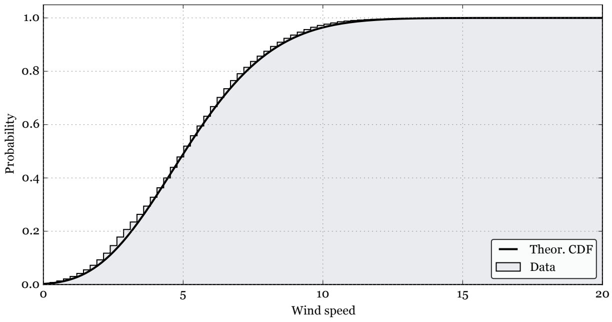


Fig. 9. CDF of **Beta** distribution compared with empirical distribution function

three-parameter case (Fig. 10–13).

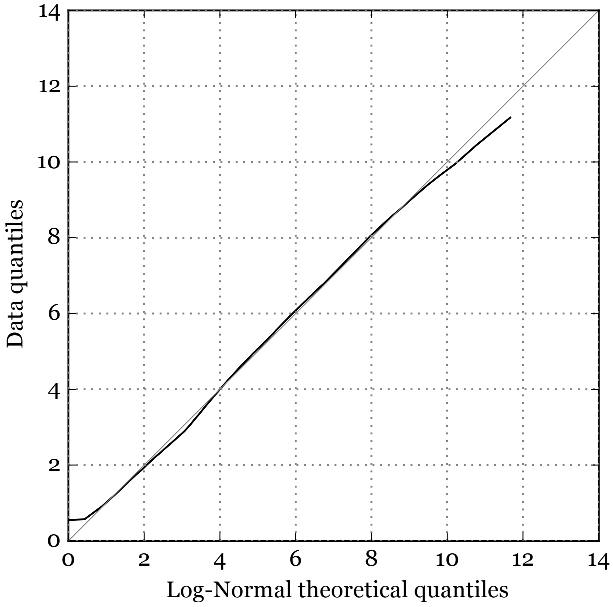


Fig. 10. Q-Q plot for LogNormal distribution

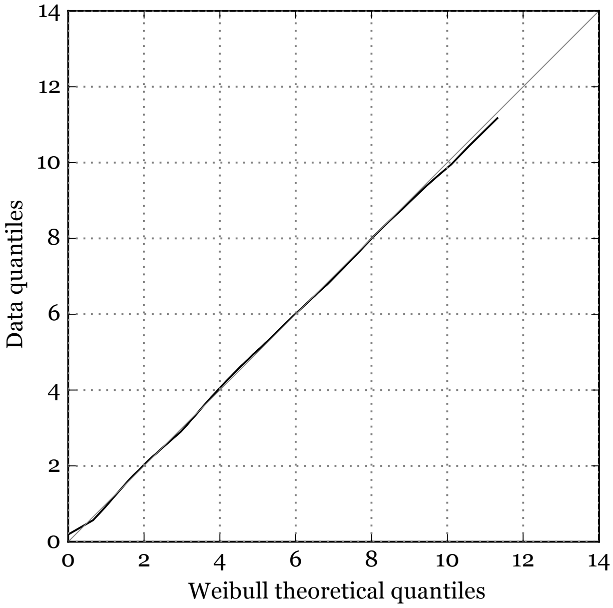


Fig. 11. Q-Q plot for Weibull distribution

CONCLUSIONS

The results of statistical data processing correspond to the results presented in the literature, where Weibull distribution is the most often used distribution for the wind speed approximation (see Lun and Lam (2000); Seguro and Lambert (2000); Bowden et al. (1983); Yeh and Wang (2008); Islam et al. (2011); Garcia et al. (1998)).

Our future work will be aimed at the construction of stochastic models that can approximate the wind speed depending on time (see Miano and Milano (2015)). On the other hand, we expect to verify the results of this work by using more dilated and large data array.

ACKNOWLEDGMENT

The work is partially supported by RFBR grants No's 15-07-08795 and 16-07-00556. Also the publication was supported by the Ministry of Education

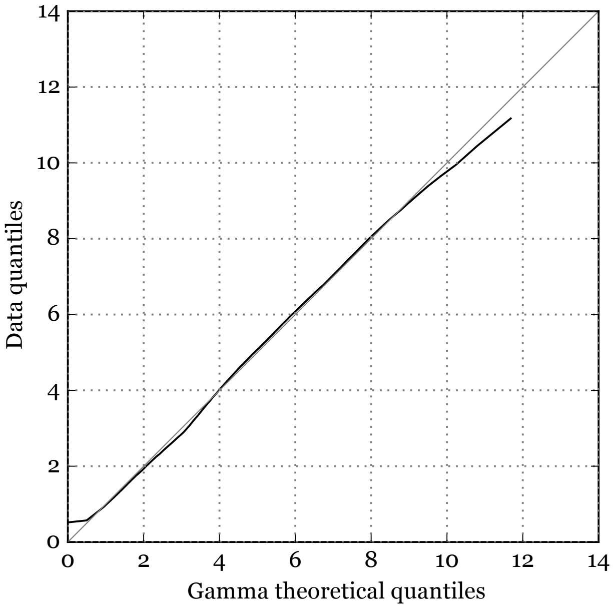


Fig. 12. Q-Q plot for Gamma distribution

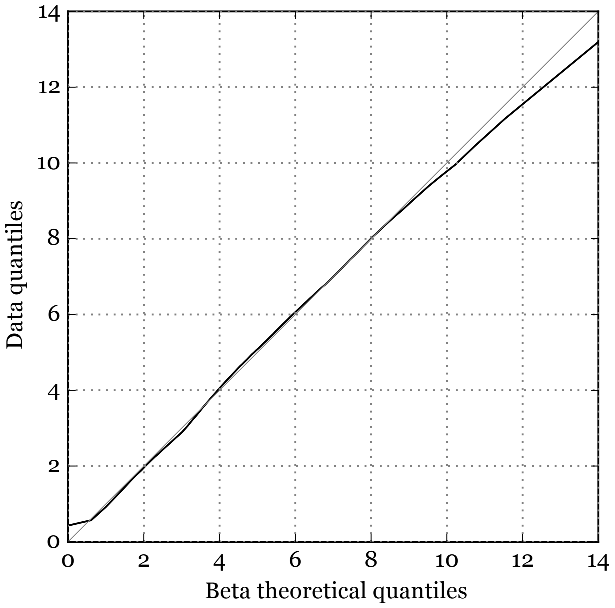


Fig. 13. Q-Q plot for Beta distribution

and Science of the Russian Federation (the Agreement No 02.a03.21.0008). The computations were carried out on the Felix computational cluster (RUDN University, Moscow, Russia) and on the HybriLIT heterogeneous cluster (Multifunctional center for data storage, processing, and analysis at the Joint Institute for Nuclear Research, Dubna, Russia).

REFERENCES

Bowden, G. J., Barker, P. R., Shestopal, V. O. and Twidell, J. W. (1983), The weibull distribution function and wind power statistics, *Wind Engineering* 7, 85–98. Provided by the SAO/NASA Astrophysics Data System.

Droettboom, M., Caswell, T. A., Hunter, J., Firing, E., Nielsen, J. H., Root, B., Elson, P., Dale, D., Lee, J.-J., Varoquaux, N., Seppnen, J. K., McDougall, D., May, R., Straw, A., de Andrade, E. S., Lee, A., Yu, T. S., Ma, E., Gohlke, C., Silvester, S., Moad, C., Hobson, P., Schulz, J., Wrtz, P., Ariza, F., Cimarron, Hisch, T., Kniazev, N., Vincent, A. F. and Thomas, I. (2017), matplotlib/matplotlib: v2.0.0. https://doi.org/10.5281/zenodo.248351

Frchet, M. R. (1927), Sur la loi de probabilit de l'cart maximum, *Annales de la Socit Polonaise de Mathematique* pp. 93–116.

Garcia, A., Torres, J., Prieto, E. and de Francisco, A. (1998), Fitting wind speed distributions: a case study, *Solar Energy* 62(2), 139–144.

Islam, M., Saidur, R. and Rahim, N. (2011), Assessment of wind energy potentiality at kudat and labuan, malaysia using weibull distribution function, *Energy* 36(2), 985–992.

Jones, E., Oliphant, T., Peterson, P. et al. (2001), SciPy: Open source scientific tools for Python. [Online; accessed 19.01.2017]. http://www.scipy.org/

Lun, I. Y. and Lam, J. C. (2000), A study of weibull parameters using long-term wind observations, *Renewable Energy* 20(2), 145–153.

Miano, R. Z. and Milano, F. (2015), Construction of sde-based wind speed models with exponential autocorrelation.

Nelson, W. B. (1982), *Applied Life Data Analysis (Wiley Series in Probability and Statistics)*.

Norman L. Johnson, Samuel Kotz, N. B. (1994), *Continuous Univariate Distributions*, Vol. 1 of *Wiley Series in Probability and Statistics*, Wiley-Interscience.

Norman L. Johnson, Samuel Kotz, N. B. (1995), *Continuous Univariate Distributions, Vol. 2*, Vol. 2 of *Wiley Series in Probability and Statistics*, Wiley-Interscience.

Project Jupyter home (2017). [Online; accessed 19.01.2017]. https://jupyter.org

Seguro, J. and Lambert, T. (2000), Modern estimation of the parameters of the weibull wind speed distribution for wind energy analysis, *Journal of Wind Engineering and Industrial Aerodynamics* 85(1), 75–84.

Weibull, W. (1951), A statistical distribution function of wide applicability, *Journal of Applied Mechanics* pp. 293–297.

Yeh, T. H. and Wang, L. (2008), A study on generator capacity for wind turbines under various tower heights and rated wind speeds using weibull distribution, *IEEE Transactions on Energy Conversion* 23(2), 592–602.

AUTHOR BIOGRAPHIES

MIGRAN N. GEVORKYAN received his Ph.D. in Mathematics in 2013. Since then, he has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). His current research activity focuses on mathematical modeling. His email address is gevorkyan_mn@rudn.university.

ANASTASIYA V. DEMIDOVA received her Ph.D. in Mathematics in 2014. Since then, she has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). Her email address is demidova_av@rudn.university.

IVAN S. ZARYADOV received his Ph.D. in Mathematics in 2010. Since then, he has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). His current research activity focuses on probability theory. His email address is zaryadov_is@rudn.university.

ROBERT A. SOBOLEWSKI He has worked in Department of Power Engineering, Fotonics and Lighting Technology in Bialystok University of Technology. His email address is r.sobolewski@pb.edu.pl.

ANNA V. KOROLKOVA received her Ph.D. in Mathematics in 2010. Since then, she has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). Her current research activity focuses on mathematical modeling. Her email address is korolkova_av@rudn.university.

DMITRY S. KULYABOV received his Ph.D. in Physics in 2000. Since then, he has worked as associate professor in RUDN University (Peoples' Friendship University of Russia). His current research activity focuses on mathematical modeling. His email address is kulyabov_ds@rudn.university.

LEONID A. SEVASTIANOV received his D.Sc. in Phys.-Math. in 1999. Since then, he has worked as full professor in RUDN University (Peoples' Friendship University of Russia). His current research activity focuses on mathematical modeling. His email address is sevastianov_la@rudn.university.

BOUNDS FOR MARKOVIAN QUEUES WITH POSSIBLE CATASTROPHES

Alexander Zeifman
Vologda State University,
Vologda, Russia
IPI FRC CSC RAS;
ISEDT RAS
Email: a_zeifman@mail.ru

Anna Korotysheva,
Yacov Satin
Vologda State University
Vologda, Russia

Ksenia Kiseleva
RUDN University
Moscow, Russia
Vologda State University
Vologda, Russia

Victor Korolev
Moscow State University,
Moscow, Russia
IPI FRC CSC RAS,
Moscow, Russia
Hangzhou Dianzi University,
Hangzhou, China

Sergey Shorgin
Institute of Informatics Problems
of the FRC CSC RAS,
Moscow, Russia

KEYWORDS

Inhomogeneous birth-death processes; queueing models; bounds on the rate of convergence.

ABSTRACT

We consider a general Markovian queueing model with possible catastrophes and obtain new and sharp bounds on the rate of convergence. Some special classes of such models are studied in details, namely, (a) the queueing system with S servers, batch arrivals and possible catastrophes and (b) the queueing model with "attracted" customers and possible catastrophes. A numerical example illustrates the calculations. Our approach can be used in modeling information flows related to high-performance computing.

INTRODUCTION

There is a large number of papers devoted to the research of Markovian queueing models with possible catastrophes, see for instance, [1], [3], [2], [10], [11], [17], [18], [19], [21], [24], [25] and the references therein. Such models are widely used in simulations for hight-performance computing. In particular, in some recent papers the authors deal with more or less special birth-death processes with additional transitions from and to origin [1], [2], [3], [10], [11], [21], [24], [25]. In the present paper we consider a more general class of Markovian queueing models with possible catastrophes and obtain key bounds on the rate of convergence, which allow us to compute the limiting characteristics of the corresponding processes.

Namely, we suppose that the queue-length process is an inhomogeneous continuous-time Markov chain $\{X(t),\ t \geq 0\}$ on the state space $E = \{0, 1, 2 \ldots\}$. All possible transition intensities are assumed to be non-random functions of time and may depend on the state of the process. From any state i the chain can jump to any another state $j > 0$ with transition intensity $q_{ij}(t)$. Moreover, the transition functions from state $i > 0$ to state 0 (catastrophe intensities) are $\beta_i(t)$. Denote by $p_{ij}(s,t) = \mathsf{P}\{X(t) = j \,| X(s) = i\}$, $i, j \geq 0,\ 0 \leq s \leq t$ the probability of transition $X(t)$, and by $p_i(t) = \mathsf{P}\{X(t) = i\}$ the corresponding state probability that $X(t)$ is in state i at the moment t. Let $\mathbf{p}(t) = (p_0(t), p_1(t), \ldots)^T$ be the vector of state probabilities at the moment t.

Throughout the paper we suppose that for any i, j

$$\mathsf{P}(X(t+h) = j | X(t) = i) =$$

$$= \begin{cases} q_{ij}(t)\,h + \alpha_{ij}(t,h), & \text{if } j \neq i, \\ \beta_i(t)\,h + \alpha_{i0}(t,h) = q_{i0}(t) + \alpha_{i0}(t,h), & \text{if } j = 0,\ i > 1, \\ 1 - \sum_{j \neq i} q_{ij}(t)h + \alpha_i(t,h), & \text{if } j = i, \end{cases} \tag{1}$$

where

$$\sup_i |\alpha_i(t,h)| = o(h). \tag{2}$$

Let $Q(t)$ be the corresponding intensity matrix. We suppose that all intensity functions are non-negative and locally integrable on $[0, \infty)$.
Put $a_{ij}(t) = q_{ji}(t)$ for $j \neq i$ and

$$a_{ii}(t) = -\sum_{j \neq i} a_{ji}(t) = -\sum_{j \neq i} q_{ij}(t). \tag{3}$$

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

As in our previous papers [9], [13], [16], [15], we suppose that

$$\sup_i |a_{ii}(t)| = L(t) < \infty, \tag{4}$$

for almost all $t \geq 0$.

Then one can write for $X(t)$ the forward Kolmogorov system

$$\frac{d\mathbf{p}(t)}{dt} = A(t)\,\mathbf{p}(t), \tag{5}$$

where $A(t) = Q^T(t)$ is a transposed intensity matrix.

Denote by $\|\cdot\|$ the l_1-norm of vector, $\|\mathbf{x}\| = \sum |x_i|$, $\|B\| = \sup_j \sum_i |b_{ij}|$, if $B = (b_{ij})_{i,j=0}^{\infty}$, and denote by Ω the set of all vectors from l_1 with nonnegative coordinates and unit norm.

We have$\|A(t)\| = 2\sup_k |a_{kk}(t)| \leq 2L(t)$ for almost all $t \geq 0$. Hence the operator function $A(t)$ from l_1 to itself is bounded for almost all $t \geq 0$ and locally integrable on interval $[0;\infty)$.

One can see that assumption (2) implies the equality $\mathbf{p}(t+h) = A(t)\mathbf{p}(t)h + \mathbf{p}(t) + o(h)$, hence the relation (5) can be considered as a differential equation in the space of sequences l_1 and one can apply to (5) the approach of [4].

Denote by $E(t,k) = E\{X(t)\,|X(0) = k\}$ the mathematical expectation (the mean) of $X(t)$ at the moment t if $X(0) = k$.

Definition. A Markov chain $X(t)$ is called *weakly ergodic*, if $\lim\limits_{t\to\infty} \left\|\mathbf{p}^1(t) - \mathbf{p}^2(t)\right\| = 0$ for any initial conditions $\mathbf{p}^1(0) = \mathbf{p}^1 \in \Omega$, $\mathbf{p}^2(0) = \mathbf{p}^2 \in \Omega$. In this situation one can consider *any* $\mathbf{p}^1(t)$ as a *quasi-stationary distribution* of the chain $X(t)$.

Definition. A Markov chain $X(t)$ has the limiting mean $\phi(t)$, if $|E(t;k) - \phi(t)| \to 0$ as $t \to \infty$ for any k.

There are two approaches to the study of the rate of convergence of continuous-time Markov chains.

The first approach is based on the notion of the logarithmic norm of a linear operator function and the respective bounds of Cauchy operator, the detailed discussion see in [6], [9], [14]. Namely, if $B(t)$, $t \geq 0$ is a one-parameter family of bounded linear operators on a Banach space $\mathcal{B}$, then

$$\gamma(B(t))_{\mathcal{B}} = \lim_{h\to+0} \frac{\|I + hB(t)\| - 1}{h} \tag{6}$$

is called the logarithmic norm of the operator $B(t)$. If $\mathcal{B} = l_1$ then the operator $B(t)$ is given by the matrix $B(t) = (b_{ij}(t))_{i,j=0}^{\infty}$, $t \geq 0$, and the logarithmic norm of $B(t)$ can be found explicitly:

$$\gamma(B(t)) = \sup_j \left(b_{jj}(t) + \sum_{i\neq j} |b_{ij}(t)| \right), \quad t \geq 0.$$

Here we apply **the second approach,** see detailed consideration for the situation of finite state space in our recent paper [22], see also [6], [8].

A matrix is called *essentially nonnegative*, if all off-diagonal elements of this matrix are nonnegative.

Let

$$\frac{d\mathbf{x}}{dt} = H(t)\mathbf{x}(t), \tag{7}$$

be a differential equation in the space of sequences l_1 with essentially nonnegative for all $t \geq 0$ countable matrix $H(t) = (h_{ij}(t))$ such that the corresponding operator function on l_1 is bounded for almost all $t \geq 0$ and locally integrable on $[0,\infty)$.

Therefore $\mathbf{x}(s) \geq 0$ implies $\mathbf{x}(t) \geq 0$ for any $t \geq s$.

Put

$$h^*(t) = \sup_j \sum_i h_{ij}(t),\; h_*(t) = \inf_j \sum_i h_{ij}(t). \tag{8}$$

Let $\mathbf{x}(0) \geq 0$. Then $\mathbf{x}(t) \geq 0$ if $t \geq 0$ and $\|\mathbf{x}(t)\| = \sum_i x_i(t)$. Hence (7) implies the inequality

$$\frac{d\mathbf{x}(t)}{dt} = \frac{d\sum_i x_i(t)}{dt} = \sum_i \left(\sum_j h_{ij} x_j \right) =$$

$$= \sum_j \left(\sum_i h_{ij} \right) x_j \leq h^*(t) \sum_j x_j = h^*(t)\|\mathbf{x}\|.$$

Then

$$\|\mathbf{x}(t)\| \leq e^{\int_0^t h^*(\tau)d\tau} \|\mathbf{x}(0)\|, \tag{9}$$

if $\mathbf{x}(0) \geq \mathbf{0}$. Let now $\mathbf{x}(0)$ be arbitrary vector from l_1. Put $x_i^+(0) = \sup(x_i(0), 0)$, $\mathbf{x}^+(0) = \left(x_1^+(0), x_2^+(0), \cdots\right)^T$ and $\mathbf{x}^-(0) = \mathbf{x}^+(0) - \mathbf{x}(0)$. Then $\mathbf{x}^+(0) \geq \mathbf{0}$, $\mathbf{x}^-(0) \geq \mathbf{0}$, $\mathbf{x}(0) = \mathbf{x}^+(0) - \mathbf{x}^-(0)$, hence $\|\mathbf{x}(0)\| = \|\mathbf{x}^+(0)\| + \|\mathbf{x}^-(0)\|$.

Finally we obtain the upper bound

$$\begin{aligned} &\|\mathbf{x}(t)\| = \\ &= \|\mathbf{x}^+(t) - \mathbf{x}^-(t)\| \leq \|\mathbf{x}^+(t)\| + \|\mathbf{x}^-(t)\| \leq \\ &\leq e^{\int_0^t h^*(\tau)d\tau} \left(\|\mathbf{x}^+(0)\| + \|\mathbf{x}^-(0)\|\right) = \\ &= e^{\int_0^t h^*(\tau)d\tau} \|\mathbf{x}(0)\|. \end{aligned} \tag{10}$$

for any initial condition.

On the other hand, if $\mathbf{x}(0) \geq 0$, then

$$\frac{d\mathbf{x}(t)}{dt} = \sum_j \left(\sum_i h_{ij} \right) x_j \geq h_*(t) \sum_j x_j = h_*(t)\|\mathbf{x}\|,$$

and we obtain the following lower bound

$$\|\mathbf{x}(t)\| \ge e^{\int_0^t h_*(\tau)d\tau}\|\mathbf{x}(0)\|, \tag{11}$$

for any nonnegative initial condition.

On sharpness of bounds.

Let us note that if the matrix of system (7) is essentially nonnegative for any t, then one can see that the logarithmic norm of this matrix is equal to our new characteristic, $\gamma\left(H(t)\right) = h^*(t)$.

Let $\{d_i\}$, $i \ge 0$ be a sequence of positive numbers such that $\inf_i d_i = d > 0$. Let $\mathsf{D} = diag\left(d_0, d_1, d_2, \dots\right)$ be the corresponding diagonal matrix and $l_{1\mathsf{D}}$ be a space of vectors

$$l_{1\mathsf{D}} = \left\{\mathbf{x} = (x_0, x_1, x_2, \ldots)/\|\mathbf{x}\|_{1\mathsf{D}} = \|\mathsf{D}\mathbf{x}\|_1 < \infty\right\}. \tag{12}$$

Put $\mathbf{z}(t) = \mathsf{D}\mathbf{x}(t)$, then (7) implies the equation

$$\frac{d\mathbf{z}}{dt} = H_{\mathsf{D}}(t)\mathbf{z}(t), \tag{13}$$

where $H_{\mathsf{D}}(t) = \mathsf{D}H(t)\mathsf{D}^{-1}$ with entries $h_{ij\mathsf{D}}(t) = \frac{d_i}{d_j}h_{ij}(t)$ is also essentially nonnegative for any $t \ge 0$. If one can find a sequence $\{d_i\}$ such that

$$h^*_{\mathsf{D}}(t) = \sup_j \sum_i \frac{d_i}{d_j} h_{ij}(t) = \inf_j \sum_i \frac{d_i}{d_j} h_{ij}(t), \tag{14}$$

then the following *equality* holds

$$\|\mathbf{x}(t)\|_{\mathsf{D}} = e^{\int_0^t h^*_{\mathsf{D}}(\tau)d\tau}\|\mathbf{x}(0)\|_{\mathsf{D}}, \tag{15}$$

for any nonnegative initial condition. Therefore, the bound

$$\|\mathbf{x}(t)\|_{\mathsf{D}} \le e^{\int_0^t h^*_{\mathsf{D}}(\tau)d\tau}\|\mathbf{x}(0)\|_{\mathsf{D}}, \tag{16}$$

which is correct for any initial condition, is *sharp*.

Note that the construction of such sequences for homogeneous birth-death processes has been studied previously in [5], [6], [7], [8], [12], [13].

ERGODICITY BOUNDS

Put

$$\beta_*\left(t\right) = \inf_i \beta_i\left(t\right), \quad \beta^*\left(t\right) = \sup_i \beta_i\left(t\right). \tag{17}$$

Theorem 1. Let catastrophe rates be essential, i. e. let

$$\int_0^\infty \beta_*\left(t\right)\, dt = +\infty. \tag{18}$$

Then the queue-length process $X\left(t\right)$ is weakly ergodic in the uniform operator topology and the following bound hold

$$\|\mathbf{p}^*\left(t\right) - \mathbf{p}^{**}\left(t\right)\| \le$$

$$\le e^{-\int\limits_0^t \beta_*(\tau)\,d\tau}\|\mathbf{p}^*\left(0\right) - \mathbf{p}^{**}\left(0\right)\| \le 2e^{-\int\limits_0^t \beta_*(\tau)\,d\tau}, \tag{19}$$

for any initial conditions $\mathbf{p}^*\left(0\right), \mathbf{p}^{**}\left(0\right)$ and any $t \ge 0$.

Moreover, if $\mathbf{p}^*\left(0\right) - \mathbf{p}^{**}\left(0\right) \ge \mathbf{0}$, then

$$\|\mathbf{p}^*\left(t\right) - \mathbf{p}^{**}\left(t\right)\| \ge$$

$$\ge e^{-\int\limits_0^t \beta_*(\tau)\,d\tau}\|\mathbf{p}^*\left(0\right) - \mathbf{p}^{**}\left(0\right)\|, \tag{20}$$

for any any $t \ge 0$.

Proof. Rewrite the forward Kolmogorov system (5) as

$$\frac{d\mathbf{p}}{dt} = A^*\left(t\right)\mathbf{p} + \mathbf{g}\left(t\right), \quad t \ge 0. \tag{21}$$

Here $\mathbf{g}\left(t\right) = \left(\beta_*\left(t\right), 0, 0, \ldots\right)^T$, $A^*\left(t\right) = \left(a^*_{ij}\left(t\right)\right)_{i,j=0}^{\infty}$, and

$$a^*_{ij}\left(t\right) = \left\{\begin{array}{cc} a_{0j}\left(t\right) - \beta_*\left(t\right), & \text{if } i = 0, \\ a_{ij}\left(t\right), & \text{otherwise .} \end{array}\right. \tag{22}$$

The solution of this equation can be written in the form

$$\mathbf{p}\left(t\right) = U^*\left(t, 0\right)\mathbf{p}\left(0\right) + \int_0^t U^*\left(t, \tau\right)\mathbf{g}\left(\tau\right)\, d\tau, \tag{23}$$

where $U^*\left(t, s\right)$ is the Cauchy operator of the corresponding homogeneous equation

$$\frac{d\mathbf{z}}{dt} = A^*\left(t\right)\mathbf{z}. \tag{24}$$

All off-diagonal elements of matrix $A^*\left(t\right)$ are nonnegative for any $t \ge 0$. Hence we can apply the approach of previous Section with $H(t) = A^*\left(t\right)$. Then we have

$$h^*(t) = \sup_i \left(a^*_{ii}\left(t\right) + \sum_{j \ne i} a^*_{ji}\left(t\right) \right) = -\beta_*\left(t\right), \tag{25}$$

hence, $\|U^*\left(t, s\right)\| \le e^{-\int_s^t \beta_*(\tau)\,d\tau}$, and we obtain

$$\|\mathbf{p}^*\left(t\right) - \mathbf{p}^{**}\left(t\right)\| \le \|U^*\left(t, 0\right)\| \, \|\mathbf{p}^*\left(0\right) - \mathbf{p}^{**}\left(0\right)\| \le \tag{26}$$

$$\le e^{-\int\limits_0^t \beta_*(\tau)\,d\tau}\|\mathbf{p}^*\left(0\right) - \mathbf{p}^{**}\left(0\right)\| \le 2e^{-\int\limits_0^t \beta_*(\tau)\,d\tau},$$

for any initial conditions $\mathbf{p}^*\left(0\right), \mathbf{p}^{**}\left(0\right)$ and any $t \ge 0$.

On the other hand, bound (20) follows from the inequality

$$h_*(t) = \inf_i \left(a^*_{ii}\left(t\right) + \sum_{j \ne i} a^*_{ji}\left(t\right) \right) = -\beta_*\left(t\right). \tag{27}$$

Now consider bounds in “weighted” norms. Let $\{d_i\}$, $1 = d_0 \le d_1 \le \ldots$ be a non-decreasing sequence, and $\mathsf{D} = diag\,(d_0, d_1, d_2, \dots)$ be the corresponding diagonal matrix. Let $l_{1\mathsf{D}}$ be the space of vectors such that (12) holds.

Put

$$\beta_{**}(t) = \inf_i \left(|a^*_{ii}(t)| - \sum_{j\neq i} \frac{d_j}{d_i} a^*_{ji}(t) \right), \tag{28}$$

and

$$\beta^{**}(t) = \sup_i \left(|a^*_{ii}(t)| - \sum_{j\neq i} \frac{d_j}{d_i} a^*_{ji}(t) \right). \tag{29}$$

Consider (21) as a differential equation in the space of sequences $l_{1\mathsf{D}}$. We have

$$\begin{gathered} \|A^*(t)\|_{1\mathsf{D}} = \|\mathsf{D}A^*(t)\mathsf{D}^{-1}\| = \\ = \sup_i \left(|a^*_{ii}(t)| + \sum_{j\neq i} \frac{d_j}{d_i} a^*_{ji}(t) \right) \le \\ \beta_{**}(t) + 2\sup_i |a^*_{ii}(t)| \le \beta_{**}(t) + 2L(t), \end{gathered} \tag{30}$$

and $\|\mathbf{g}(t)\|_{1\mathsf{D}} = \beta_*(t)$, hence we can apply the same approach to equation (21) in the space $l_{1\mathsf{D}}$, and the equality

$$\begin{gathered} \gamma\left(A^*(t)\right)_{1\mathsf{D}} = \gamma\left(\mathsf{D}A^*(t)\mathsf{D}^{-1}\right) = \\ = \sup_i \left(a^*_{ii}(t) + \sum_{j\neq i} \frac{d_j}{d_i} a^*_{ji}(t) \right) = -\beta_{**}(t), \end{gathered} \tag{31}$$

implies the following statement.

Theorem 2. Let $\{d_i\}$, $1 = d_0 \le d_1 \le \dots$ be a non-decreasing sequence such that,

$$\int_0^\infty \beta_{**}(t)\,dt = +\infty. \tag{32}$$

Then the following bound on the rate of convergence holds:

$$\begin{gathered} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1\mathsf{D}} \le \\ \le e^{-\int\limits_0^t \beta_{**}(\tau)\,d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1\mathsf{D}}, \end{gathered} \tag{33}$$

for any initial conditions $\mathbf{p}^*(0), \mathbf{p}^{**}(0)$ and any $t \ge 0$. Moreover, if $\mathbf{p}^*\,(0) - \mathbf{p}^{**}\,(0) \ge \mathbf{0}$, then

$$\begin{gathered} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1\mathsf{D}} \ge \\ \ge e^{-\int\limits_0^t \beta^{**}(\tau)\,d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1\mathsf{D}}, \end{gathered} \tag{34}$$

for any $t \ge 0$.

Let $l_{1E} = \{\mathbf{z} = (p_1, p_2, \dots)\}$ be a space of sequences such that $\|\mathbf{z}\|_{1E} = \sum_{k\ge1} k|p_k| < \infty$. Put $W = \inf_{k\ge1} \frac{d_k}{k}$. Then $W\|\mathbf{z}\|_{1E} \le \|\mathbf{z}\|_{1\mathsf{D}}$.

Corollary 1. Let a sequence $\{d_i\}$ be such that (32) holds, and, let moreover $W > 0$. Then $X(t)$ has the limiting mean, say $\phi(t) = E(t, 0)$, and the following bound holds:

$$|E(t,j) - E(t,0)| \le \frac{1+d_j}{W} e^{-\int\limits_0^t \beta_{**}(\tau)\,d\tau}, \tag{35}$$

for any j and any $t \ge 0$.

We can use this approach and formula (23) for obtaining the bounds of state probabilities in the following way. Consider again the space of sequences $l_{1\mathsf{D}}$, and put $X(0) = 0$. Then $\mathbf{p}(0) = \mathbf{0}$ and we obtain

$$\begin{gathered} \sum_i d_i p_i(t) = \|\mathbf{p}\,(t)\,\| \le \\ \le \int_0^t \|U^*\,(t,\tau)\,\mathbf{g}\,(\tau)\,\|\,d\tau \le \int_0^t \beta_*(\tau) e^{-\int\limits_\tau^t \beta_{**}(\tau)\,d\tau}, \end{gathered} \tag{36}$$

in the $1\mathsf{D}$-norm. Hence

$$d_N \sum_{i\ge N} p_i(t) \le \|\mathbf{p}\,(t)\,\| \le \int_0^t \beta_*(\tau) e^{-\int\limits_\tau^t \beta_{**}(\tau)\,d\tau}, \tag{37}$$

$$\sum_{i\ge N} p_i(t) \le d_N^{-1} \int_0^t \beta_*(\tau) e^{-\int\limits_\tau^t \beta_{**}(\tau)\,d\tau}, \tag{38}$$

and we obtain the following statement.

Corollary 2. Let sequence $\{d_i\}$ be such that (32) holds. Then the following bound holds:

$$\sum_{i<N} p_i(t) \ge 1 - d_N^{-1} \int_0^t \beta_*(\tau) e^{-\int\limits_\tau^t \beta_{**}(\tau)\,d\tau}, \tag{39}$$

if $X(0) = 0$ and any $t \ge 0$.

SPECIFIC QUEUEING SYSTEMS

1. Consider firstly the queueing system with S servers, batch arrivals and possible catastrophes, and suppose that the corresponding rate functions are the following:

$\lambda_k(t)$ is the intensity of arrival of a group of k customers to the queue,

$\mu_k(t) = \mu(t)\min(k, S)$ is the intensity of service of a customer if the current number of customers in the queue is k,

finally, $\beta_k(t)$ is the intensity of catastrophes if the current number of customers in the queue is k.

To simplify the formulas, we will assume all intensities $1-$periodic.

Firstly, if assumption (18) is fulfilled, then the queue-length process $X(t)$ is weakly ergodic in the uniform operator topology and bound of the rate of convergence (19) holds.

Bounds in weighted norms seem essentially more interesting.

1a. Let arrival rates be exponentially decreasing in k, namely, let there exist $r > 1$ such that $\lambda_k(t) = r^{-k}\lambda(t)$. Put $d_k = \delta^k$, where $1 < \delta < r$.

Then $\beta_{**}(t) = \inf \alpha_i(t)$, where $\alpha_i(t) = |a^*_{ii}(t)| - \sum_{j\neq i} \frac{d_j}{d_i} a^*_{ji}(t)$. We have

$$\alpha_0(t) = \beta_*(t) - \lambda(t)\frac{r(\delta-1)}{(r-1)(r-\delta)},$$

and

$$\alpha_i(t) = \beta_*(t) + \left(1-\delta^{-k}\right)\left(\beta_k(t) - \beta_*(t)\right) +$$

$$+\left(1-\delta^{-1}\right)\mu_k(t) - \lambda(t)\frac{r(\delta-1)}{(r-1)(r-\delta)} \geq \alpha_0(t),$$

hence

$$\beta_{**}(t) = \alpha_0(t) = \beta_*(t) - \lambda(t)\frac{r(\delta-1)}{(r-1)(r-\delta)},$$

and $\int_0^1 \beta_{**}(t)\,dt > 0$ for sufficiently small $0 < \delta - 1$.

Finally, in this situation $X(t)$ is weakly ergodic in the corresponding $l_{1\mathrm{D}}$-norm and has the limiting mean for any service rate $\mu(t)$ and any S.

1b. Let arrival rates be decreasing in k more slowly, and let, however,

$$\sum_k k\lambda_k(t) \leq Q < \infty, \tag{40}$$

for any $t \in [0,1]$.

Put $d_k = \frac{N+k}{N}$, $k \geq 0$, where N is sufficiently large.

Then also $\beta_{**}(t) = \inf \alpha_i(t)$, where $\alpha_i(t) = |a^*_{ii}(t)| - \sum_{j\neq i} \frac{d_j}{d_i} a^*_{ji}(t)$. We have

$$\alpha_0(t) = \beta_*(t) - \lambda(t)\sum_k \left(\frac{N+k}{N} - 1\right) =$$

$$= \beta_*(t) - \sum_k \frac{k}{N}\lambda(t) \geq \beta_*(t) - \frac{Q}{N},$$

and

$$\alpha_i(t) \geq \alpha_0(t).$$

Therefore

$$\beta_{**}(t) = \alpha_0(t) \geq \beta_*(t) - \frac{Q}{N},$$

and $\int_0^1 \beta_{**}(t)\,dt > 0$ for sufficiently large N.

Finally, in this situation $X(t)$ is weakly ergodic in the corresponding $l_{1\mathrm{D}}$-norm and has the limiting mean for any service rate $\mu(t)$ and any S.

2. Consider now the queueing model with "attracted" customers and possible catastrophes.

Namely, we consider an analog of an inhomogeneous $M|M|S$ queue with catastrophes where customers may arrive to the queue only in groups of $k+1$ customers with intensity $a_{k,2k+1} = \lambda(t)$, if the length of the queue at this moment equals k,
$\mu_k(t) = \mu(t)\min(k, S)$ is the intensity of service of a customer, if the current number of customers in the queue is k,
finally, $\beta_k(t)$ is the intensity of catastrophes, if the current number of customers in the queue is k.

To simplify the formulas we will suppose all intensities 1- periodic.

Certainly, if assumption (18) is fulfilled, then the queue-length process $X(t)$ is weakly ergodic in the uniform operator topology and bound (19) of the rate of convergence holds.

Moreover, the limiting mean for this queue-length process exists under the simple additional assumption

$$\int_0^1 \left(\beta_*(t) - \lambda(t)\right)\,dt > 0. \tag{41}$$

To check this claim put $d_k = k+1$, $k \geq 0$.

Then also $\beta_{**}(t) = \inf \alpha_i(t)$, where $\alpha_i(t) = |a^*_{ii}(t)| - \sum_{j\neq i} \frac{d_j}{d_i} a^*_{ji}(t)$. We have $\alpha_0(t) = \beta_*(t) - \lambda(t)$, and $\alpha_i(t) \geq \alpha_0(t)$, for any $i \geq 1$.

Therefore $\beta_{**}(t) = \alpha_0(t) = \beta_*(t) - \lambda(t)$, and $\int_0^1 \beta_{**}(t)\,dt > 0$ if (41) holds. Finally, in this situation $X(t)$ is weakly ergodic in the respective $l_{1\mathrm{D}}$-norm and has the limiting mean for any service rate $\mu(t)$ and any S.

EXAMPLE

Consider now a simple special model. Let $\lambda_k(t) = \frac{(3+\sin 2\pi t)}{4^k}$, $\mu_k(t) = (1+\cos 2\pi t)\min(k,5)$, $\beta_k(t) = \beta_*(t) = 2 - \sin 2\pi t$.

Put $\delta = 4/3$ and $d_k = \delta^k$. Then we have

$$\beta_{**}(t) = \alpha_0(t) = \beta_*(t) - \lambda(t)/6 = 1.5 - \frac{7}{6}\sin 2\pi t,$$

and $\int_0^1 \beta_{**}(t)\,dt = \beta^0_{**} = 1.5$.

Hence Theorem 2 implies the bound

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1\mathrm{D}} \leq Me^{-1.5t}\|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1\mathrm{D}}, \tag{42}$$

where $M \leq e^{\int_0^1 \frac{7}{6}|\sin 2\pi t|\,dt} \leq 2$.

Hence there exists a limiting 1-periodic regime, say $\pi(t)$. We can apply inequality (36) and obtain the following bounds for the solution of the system (21) with zero initial condition:

$$\|\pi\,(t)\,\|_{1\mathrm{D}} \leq \int_0^t \beta_*(\tau)e^{-\int\limits_\tau^t \beta_{**}(\tau)\,d\tau} \leq$$

$$\leq 6\int_0^t e^{-1.5(t-\tau)\,d\tau} \leq 4, \tag{43}$$

for any $t \geq 0$.

Therefore, $\limsup_{t\to\infty} \|\pi(t)\|_{1\mathsf{D}} \leq 4$, and the $1-$periodicity of the limit regime implies the inequality $\|\pi(t)\|_{1\mathsf{D}} \leq 4$ for any t and *any* initial condition.

Now from (42) we have

$$\|\mathbf{p}^*(t) - \pi(t)\|_{1\mathsf{D}} \leq 2e^{-1.5t}\left(\|\mathbf{p}^*(0)\|_{1\mathsf{D}} + 4\right), \quad (44)$$

and particularly

$$\|\mathbf{p}^*(t) - \pi(t)\|_{1\mathsf{D}} \leq 2e^{-1.5t}\left(\left(\frac{4}{3}\right)^k + 4\right), \quad (45)$$

if $X(0) = k$.

Finally we can apply the approach of [20], [23] and find the appropriate truncations for $X(t)$. The corresponding plots of the limiting characteristics for the queue-length process are shown here.

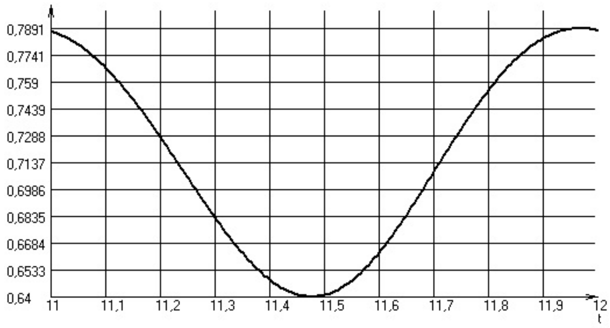


Fig. 1. Approximation of the limiting probability of empty queue $\mathsf{P}\{X(t) = 0|X(0) = 0\}$ on $[11, 12]$.

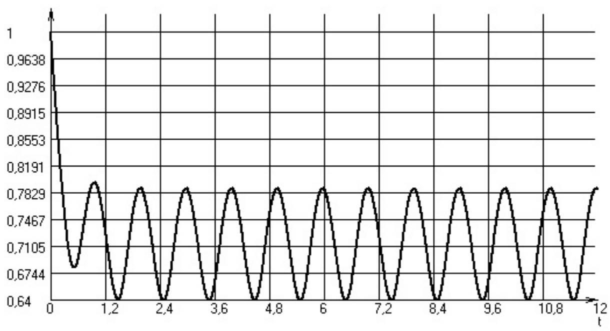


Fig. 2. Approximation of the probability of empty queue $\mathsf{P}\{X(t) = 0|X(0) = 0\}$ on $[0, 12]$.

CONCLUSION

We consider a general Markovian queueing model with possible catastrophes and obtain new and sharp bounds on the rate of convergence. Some special classes of such models are studied in details, namely, (a) the queueing system with S servers, batch arrivals and possible catastrophes and (b) the queueing model with "attracted" customers and possible catastrophes. A numerical example illustrates the calculations. Our approach can be used in modeling information flows related to high-performance computing. Perturbation bounds and estimation of error of truncations will be studied in the next paper.

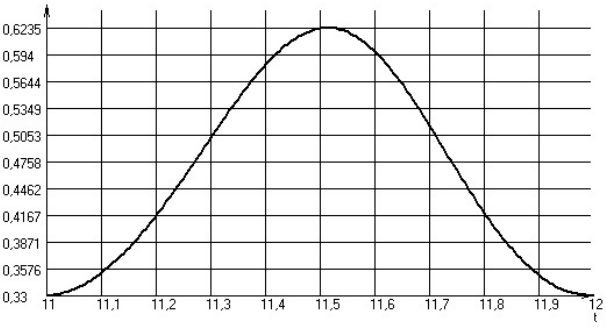


Fig. 3. Approximation of the limiting mean $E(t, 0)$ on $[11, 12]$.

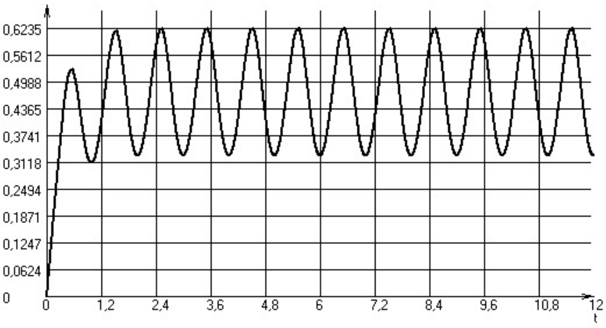


Fig. 4. Approximation of the mean $E(t, 0)$ on $[0, 12]$.

ACKNOWLEDGEMENTS

The work was supported by the Ministry of Education of the Russian Federation (the Agreement number 02.a03.21.0008 of 24 June 2016), by the Russian Foundation for Basic Research, project no. 15-01-01698, and by the RAS Presidium Program No. I.33P (project 0063-2016-0015).

REFERENCES

[1] A. Y. Chen, E. Renshaw. The $M|M|1$ queue with mass exodus and mass arrives when empty // J. Appl. Prob., 1997, 34, 192–207 .

[2] Chen, A. Y. and Renshaw, E. Markov bulk-arriving queues with state-dependent control at idle time // Adv. Appl. Prob. 2004, 36, 499–524.

[3] A. Y. Chen, P. Pollett, J. P. Li, H. J. Zhang. Markovian bulk-arrival and bulk-service queues with state-dependent control // 2010, Queueing Syst., 64, 267–304.

[4] Daleckij, Ju.L., Krein, M.G. (1974). Stability of solutions of differential equations in Banach space. Amer. Math. Soc. Transl. **43**.

[5] *Van Doorn E.* Conditions for exponential ergodicity and bounds for the decay parameter of a birth-death process // Adv. Appl. Probab., 17, 1985, p. 514–530.

[6] E. A. Van Doorn, A. I. Zeifman, T. L. Panfilova (2010). Bounds and asymptotics for the rate of convergence of birth-death processes // Th. Prob. Appl., **54**, 97–113.

[7] *Granovsky B. L., Zeifman A. I.* The decay function of nonhomogeneous birth-death processes, with application to mean-field models // Stoch. Proc. Appl., 72, 1997, p. 105–120.

[8] *Granovsky B. L., Zeifman A. I.* The N-limit of spectral gap of a class of birth-death Markov chains // Appl. Stoch. Models in Business and Industry, 16, 2000, p. 235–248.

[9] *Granovsky B. L., Zeifman A. I.* Nonstationary Queues: Estimation of the Rate of Convergence // Queueing Systems, 2004, 46, p. 363–388.

[10] Junping Li, Anyue Chen. The Decay Parameter and Invariant Measures for Markovian Bulk-Arrival Queues with Control at Idle Time // Methodology and Computing in Applied Probability, 2013, 15, 467–484.

[11] Zhang, L., Li, J.. The M—M—c queue with mass exodus and mass arrivals when empty //Journal of Applied Probability, 2015, 52, 990–1002.

[12] *Zeifman A. I.* Some estimates of the rate of convergence for birth and death processes // J. Appl. Probab. 28, 1991, p. 268–277.

[13] A. I. Zeifman, Upper and lower bounds on the rate of convergence for nonhomogeneous birth and death processes // Stoch. Proc. Appl., 1995, 59, 157–173.

[14] *Zeifman A. I.* On the estimation of probabilities for birth and death processes // J. Appl. Probab., 32, 1995, p. 623–634.

[15] *Zeifman A., Leorato S., Orsingher E., Satin Ya., Shilova G.* Some universal limits for nonhomogeneous birth and death processes // Queueing systems, 52, 2006, p. 139–151.

[16] Zeifman A. I., Bening V. E., Sokolov I. A. Continuous-time Markov chains and models. Elex-KM, Moscow. 2008.

[17] Zeifman, A. I., Korotysheva, A. V., Satin, Y. A., Shorgin, S. Y. (2010). On stability for nonstationary queueing systemswith catastrophes. Informatika i Ee Primeneniya [Informatics and its Applications], 4(3), 9–15.

[18] Zeifman, A. I., Korotysheva, A. V., Panfilova, T. Y. L., Shorgin, S. Y. (2011). Stability bounds for some queueing systems with catastrophes. Informatics and its Applications, 5(3), 27–33 (in Russian).

[19] A. Zeifman, A. Korotysheva Perturbation Bounds for $M_t|M_t||N$ Queue with Catastrophes // Stochastic Models, 28:1, 2012. – 49–62.

[20] A. Zeifman, Ya. Satin, V. Korolev, S. Shorgin. On truncations for weakly ergodic inhomogeneous birth and death processes // International Journal of Applied Mathematics and Computer Science, 2014, 24, 503–518.

[21] A. Zeifman, Y. Satin, A. Korotysheva, V. Korolev, S. Shorgin, R. Razumchik. Ergodicity and perturbation bounds for inhomogeneous birth and death processes with additional transitions from and to origin // Int. J. Appl. Math. Comput. Sci, 2015, 25(4), 503–518.

[22] Zeifman, A. I., Korolev, V. Y. Two-sided bounds on the rate of convergence for continuous-time finite inhomogeneous Markov chains. Statistics & Probability Letters, 2015, 103, 30–36.

[23] Zeifman, A. I.; Korotysheva, A. V.; Korolev, V. Yu.; Satin Ya. A. Truncation bounds for approximations of inhomogeneous continuous-time Markov chains. Th. Prob. Appl. **2016**, 61, 563–569.

[24] Zeifman, A. I.; Satin, Ya. A.; Korotysheva, A. V.; Korolev, V. Y.; Bening, V. E. On a class of Markovian queuing systems described by inhomogeneous birth-and-death processes with additional transitions. Doklady Mathematics. **2016**, 94, 502–505.

[25] Zeifman A., Korotysheva A., Satin Y., Razumchik R., Korolev V., Shorgin, S. Ergodicity and uniform in time truncation bounds for inhomogeneous birth and death processes with additional transitions from and to origin. arXiv preprint arXiv:1604.02294, 2016.

AUTHOR BIOGRAPHIES

ALEXANDER ZEIFMAN is Doctor of Science in physics and mathematics; professor, Heard of Department of Applied Mathematics, Vologda State University; senior scientist, Institute of Informatics Problems, Federal Research Center ”Computer Science and Control” of the Russian Academy of Sciences; principal scientist, Institute of Socio-Economic Development of Territories, Russian Academy of Sciences. His email is `a_zeifman@mail.ru`.

ANNA KOROTYSHEVA is Candidate of Science (PhD) in physics and mathematics, associate professor, Vologda State University. Her email is `a_korotysheva@mail.ru`.

YACOV SATIN is Candidate of Science (PhD) in physics and mathematics, associate professor, Vologda State University. His email is `yacovi@mail.ru`.

KSENIA KISELEVA is scientific researcher, RUDN University, Moscow, Russia; PhD student, Vologda State University. Her email is `ksushakiseleva@mail.ru`.

VICTOR KOROLEV is Doctor of Science in physics and mathematics, professor, Head of Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; professor, Hangzhou Dianzi University, Hangzhou, China. His email is `victoryukorolev@yandex.ru`.

SERGEY SHORGIN is Doctor of Science in physics and mathematics, professor, Deputy Director of the Institute of Informatics Problems of the Federal Research Center ”Computer Science and Control” of the Russian Academy of Sciences. His email is `sshorgin@ipiran.ru`.

TWO-SIDED TRUNCATIONS FOR THE $M_t|M_t|S$ QUEUEING MODEL

Yacov Satin, Anna Korotysheva,
Galina Shilova, Alexander Sipin and Elena Fokicheva
Vologda State University
Vologda, Russia

Ksenia Kiseleva
RUDN University
Moscow, Russia
Vologda State University
Vologda, Russia

Alexander Zeifman
Vologda State University,
Vologda, Russia
IPI FRC CSC RAS;
ISEDT RAS
Email: a_zeifman@mail.ru

Victor Korolev
Moscow State University,
Moscow, Russia
IPI FRC CSC RAS,
Moscow, Russia
Hangzhou Dianzi University,
Hangzhou, China

Sergey Shorgin
Institute of Informatics Problems
of the FRC CSC RAS
Moscow, Russia

KEYWORDS

Inhomogeneous birth-death processes; queueing models; two-sided uniform approximation bounds

ABSTRACT

The paper deals with the problem of existence and construction of limiting characteristics for time-inhomogeneous birth and death processes which is important for queueing applications. For this purpose we calculate the limiting characteristics for the process via the construction of two-sided uniform in time truncation bounds. We consider the $M_t|M_t|S$ queueing model and obtain uniform approximation bounds of two-sided truncations. A numerical example is considered. Our approach to truncations of the state space can be used in modeling information flows related to high-performance computing.

INTRODUCTION

Explicit expressions for the probability characteristics of stochastic birth-death queueing models can be found in a few special cases. One of main problems for obtaining the limiting behavior of the process is studying of the rate of convergence as time $t \to \infty$ to the steady state of a process. The problem of existence and construction of limiting characteristics for time-inhomogeneous birth and death processes is important for queueing applications, see [2], [4], [5], [9]. Calculation of the limiting characteristics for the process via "north-west" truncations was firstly mentioned in [7] and was considered in details in [9]. Uniform in time north-west truncation bounds have been obtained in [10], [11] for birth-death processes and general Markov chains respectively.

Two-sided uniform in time truncation bounds were firstly studied in our previous paper [6]. This paper is the continuation of [6]. Namely, we apply this approach for a specific class of queueing models.

Let $X = X(t)$, $t \geq 0$ be a birth and death process (BDP) with birth and death rates $\lambda_n(t)$, $\mu_n(t)$ respectively.

Let $p_{ij}(s,t) = Pr\{X(t) = j\,|X(s) = i\}$ for $i, j \geq 0$, $0 \leq s \leq t$ be the transition probability functions of $X = X(t)$ and $p_i(t) = Pr\{X(t) = i\}$ - the state probabilities.

Also we assume that

$$\mathsf{P}\left(X\left(t+h\right)=j|X\left(t\right)=i\right)=$$

$$=\begin{cases} q_{ij}\left(t\right)h+\alpha_{ij}\left(t,h\right) & \text{if } j\neq i,\\ 1-\sum\limits_{k\neq i} q_{ik}\left(t\right)h+\alpha_i\left(t,h\right) & \text{if } j=i, \end{cases} \qquad (1)$$

where all $\alpha_i(t,h)$ are $o(h)$ uniformly in i, i. e. $\sup_i |\alpha_i(t,h)| = o(h)$. Here all $q_{i,i+1}(t) = \lambda_i(t)$, $i \geq 0$, $q_{i,i-1}(t) = \mu_i(t)$ $i \geq 1$, and all other $q_{ij}(t) \equiv 0$.

The probabilistic dynamics of the process is represented by the forward Kolmogorov system of differential equations:

$$\begin{cases} \frac{dp_0}{dt} &= -\lambda_0(t)p_0+\mu_1(t)p_1,\\ \frac{dp_k}{dt} &= \lambda_{k-1}(t)p_{k-1}-\left(\lambda_k(t)+\mu_k(t)\right)p_k+\\ & \quad +\mu_{k+1}(t)p_{k+1}, \quad k\geq 1. \end{cases} \qquad (2)$$

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

By $\mathbf{p}(t) = (p_0(t), p_1(t), \dots)^\top$, $t \geq 0$, we denote the column vector of state probabilities and by $A(t) = (a_{ij}(t))$, $t \geq 0$ the matrix related to (2). Moreover, $A(t) = Q^\top(t)$, where $Q(t)$ - the intensity (or infinitesimal) matrix for $X(t)$.

We assume that all birth and death intensity functions $\lambda_i(t)$ and $\mu_i(t)$ are locally integrable on $[0, \infty)$. We suppose that

$$\lambda_n(t) \leq \Lambda_n \leq L < \infty, \quad \mu_n(t) \leq \Delta_n \leq L < \infty, \tag{3}$$

for almost all $t \geq 0$. By $\|\cdot\|$ we denote the l_1-norm, i. e. $\|\mathbf{x}\| = \sum |x_i|$, and $\|B\| = \sup_j \sum_i |b_{ij}|$ for $B = (b_{ij})_{i,j=0}^{\infty}$.

Let Ω be a set all stochastic vectors, i. e. l_1 vectors with nonnegative coordinates and unit norm.

We have

$$\|A(t)\| \leq 2\sup(\lambda_k(t) + \mu_k(t)) \leq 4L,$$

for almost all $t \geq 0$. Hence the operator function $A(t)$ from l_1 into itself is bounded for almost all $t \geq 0$ and locally integrable on $[0; \infty)$.

We consider the system (2) as a differential equation

$$\frac{d\mathbf{p}}{dt} = A(t)\,\mathbf{p}, \quad \mathbf{p} = \mathbf{p}(t), \quad t \geq 0, \tag{4}$$

in the space l_1 with bounded operator function $A(t)$.

The Cauchy problem for differential equation (1) has unique solutions for arbitrary initial condition (see, for instance, [1]), and moreover $\mathbf{p}(s) \in \Omega$ implies $\mathbf{p}(t) \in \Omega$ for $t \geq s \geq 0$.

We apply the general approach to employ the logarithmic norm of a matrix for the study of the problem of stability of Kolmogorov system of differential equations associated with nonhomogeneous Markov chains, see the corresponding definitions, bounds, references and other details in [3], [4], [8], [9], [10].

Definition. A Markov chain $X(t)$ is called weakly ergodic, if $\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \to 0$ as $t \to \infty$ for any initial conditions $\mathbf{p}^*(0), \mathbf{p}^{**}(0)$, where $\mathbf{p}^*(t)$ and $\mathbf{p}^{**}(t)$ are the corresponding solutions of (4).

Put $E_k(t) = E\{X(t)\,|X(0) = k\}$ (then the corresponding initial condition of system (4) is the $k - th$ unit vector $\mathbf{e_k}$).

Definition. Let $X(t)$ be a Markov chain. Then $\varphi(t)$ is called the *limiting mean* of $X(t)$ if

$$\lim_{t\to\infty} (\varphi(t) - E_k(t)) = 0$$

for any k.

TWO-SIDED TRUNCATIONS OF INHOMOGENEOUS BIRTH-DEATH PROCESSES

By introducing $p_i(t) = 1 - \sum_{j \neq i} p_j(t)$, (for arbitrary fixed i and $\mathbf{p}(t) \in \Omega$, $t \geq 0$) we have the following system from (4)

$$\frac{d\mathbf{z}(t)}{dt} = B(t)\mathbf{z}(t) + \mathbf{f}(t), \tag{5}$$

where $\mathbf{z}(t)$ is $\mathbf{p}(t)$ without coordinate p_i, namely, $\mathbf{z}(t) = (p_0, p_1, \dots, p_{i-1}, p_{i+1}, \dots)$. Hence we obtain $f(t) = (0, 0, \dots, \mu_i, \lambda_i, 0, \dots)$, and the corresponding $B(t)$.

Let D^* be a matrix

$$D^* = \begin{array}{c} \\ 0 \\ \\ i-2 \\ i-1 \\ i+1 \\ i+2 \\ i+3 \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccccc} 0 & & i-2 & i-1 & i+1 & i+2 & i+3 & \end{array} \\ \begin{pmatrix} -1 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & \cdots & -1 & 0 & 0 & 0 & 0 & \cdots \\ -1 & \cdots & -1 & -1 & 0 & 0 & 0 & \cdots \\ 0 & \cdots & 0 & 0 & 1 & 1 & 1 & \cdots \\ 0 & \cdots & 0 & 0 & 0 & 1 & 1 & \cdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & 1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \end{array},$$

then $D^*BD^{*-1} =$

$$\begin{pmatrix} -\mu_1 - \lambda_0 & \mu_1 & 0 & 0 & \cdots \\ \lambda_1 & -\mu_2 - \lambda_1 & \mu_2 & 0 & \cdots \\ 0 & \lambda_2 & -\mu_3 - \lambda_2 & \mu_3 & \cdots \\ 0 & 0 & \lambda_3 & -\mu_4 - \lambda_3 & \cdots \\ 0 & 0 & 0 & \lambda_4 & \cdots \\ \cdots & & & & \end{pmatrix}$$

Let now $\{d_k\}$ be a sequence of positive numbers, and $D^{**} = diag\,(d_0, d_1, \dots, d_{i-1}, d_{i+1}, d_{i+2}, \dots)$. Put $D = D^{**}D^*$,

$$D = \begin{pmatrix} -d_0 & 0 & 0 & 0 & 0 & \cdots \\ -d_1 & -d_1 & 0 & 0 & 0 & \cdots \\ \cdots & & & & & \\ -d_{i-1} & -d_{i-1} & \cdots & -d_{i-1} & 0 & \cdots \\ 0 & 0 & \cdots & 0 & d_{i+1} & d_{i+1} & \cdots \\ 0 & 0 & \cdots & 0 & 0 & d_{i+2} & \cdots \\ \cdots & & & & & \end{pmatrix}.$$

Let l_{1D} be the space of sequences: $l_{1D} = \{\mathbf{z} = (p_0, p_1, ..., p_{i-1}, p_{i+1}, ...)^\top : \|\mathbf{z}\|_{1D} \equiv \|D\mathbf{z}\| < \infty\}$. We introduce the auxiliary space of sequences l_{1E} as $l_{1E} = \{\mathbf{z} = (p_0, p_1, ..., p_{i-1}, p_{i+1}, ...)^\top : \|\mathbf{z}\|_{1E} \equiv \sum_{k \neq i} k|p_k| < \infty\}$.

Consider the expressions:

$$\alpha_k(t) = \begin{cases} \lambda_k(t) + \mu_{k+1}(t) - \frac{d_{k+1}}{d_k}\lambda_{k+1}(t) - \frac{d_{k-1}}{d_k}\mu_k(t), & k < i-1 \\ \lambda_{i-1}(t) + \mu_i(t) - \frac{d_{i+1}}{d_{i-1}}\lambda_i(t) - \frac{d_{i-2}}{d_{i-1}}\mu_{i-1}(t), & k = i-1 \\ \lambda_i(t) + \mu_{i+1}(t) - \frac{d_{i+2}}{d_{i+1}}\lambda_{i+1}(t) - \frac{d_{i-1}}{d_{i+1}}\mu_i(t), & k = i \\ \lambda_k(t) + \mu_{k+1}(t) - \frac{d_{k+2}}{d_{k+1}}\lambda_{k+1}(t) - \frac{d_k}{d_{k+1}}\mu_k(t), & k > i \end{cases} \tag{6}$$

and

$$\alpha(t) = \inf_{k \geq 0} \alpha_k(t). \tag{7}$$

Considering (5) as a differential equation in the space l_{1D}, we have its solution:

$$\mathbf{z}(t) \;=\; V(t,0)\mathbf{z}(0) + \int_0^t V(t,\tau)\mathbf{f}(\tau)\,d\tau, \quad (8)$$

where $V(t,z)$ is the Cauchy operator of (5), see [8].

We obtain $\|\mathbf{f}(t)\|_{1D} = d_{i-1}\mu_i(t) + d_{i+1}\lambda_i(t) \leq d_{i-1}\Delta_i + d_{i+1}\Lambda_i$ for almost all $t \geq 0$. On the other hand, putting

$$\beta_k\left(t\right) =$$

$$\begin{cases} \lambda_k\left(t\right) + \mu_{k+1}\left(t\right) + \frac{d_{k+1}}{d_k}\lambda_{k+1}\left(t\right) + \frac{d_{k-1}}{d_k}\mu_k\left(t\right), \; k < i-1 \\ \lambda_{i-1}\left(t\right) + \mu_i\left(t\right) + \frac{d_{i+1}}{d_{i-1}}\lambda_i\left(t\right) + \frac{d_{i-2}}{d_{i-1}}\mu_{i-1}\left(t\right), \; k = i-1 \\ \lambda_i\left(t\right) + \mu_{i+1}\left(t\right) + \frac{d_{i+2}}{d_{i+1}}\lambda_{i+1}\left(t\right) + \frac{d_{i-1}}{d_{i+1}}\mu_i\left(t\right), \; k = i \\ \lambda_k\left(t\right) + \mu_{k+1}\left(t\right) + \frac{d_{k+2}}{d_{k+1}}\lambda_{k+1}\left(t\right) + \frac{d_k}{d_{k+1}}\mu_k\left(t\right), \; k > i. \end{cases} \quad (9)$$

one has

$$\|B(t)\|_{1D} = \sup_{k\geq 0}\beta_k(t) \leq 4L - \alpha(t),$$

for almost all $t \geq 0$.

Then the following bound for the logarithmic norm $\gamma\left(B(t)\right)$ in l_{1D} is correct:

$$\gamma\left(B\right)_{1D} = \gamma\left(DB(t)D^{-1}\right)_1 = \\ -\inf_{k\geq 0}\left(\alpha_k\left(t\right)\right) = -\alpha(t), \quad (10)$$

in accordance with (7), see detailed discussion in [3], [4], [9], [10].

Therefore,

$$\|V(t,s)\|_{1D} \leq e^{-\int\limits_s^t \alpha(\tau)\,d\tau}. \quad (11)$$

Suppose now that there exist positive M and α such that

$$e^{-\int\limits_s^t \alpha(\tau)\,d\tau} \leq Me^{-\alpha(t-s)}, \quad (12)$$

for any $0 \leq s \leq t$. Then $X(t)$ is exponentially weakly ergodic in $1D$ norm.

Put now $\mathbf{z}\left(0\right) = 0$ (i. e., $\mathbf{p}\left(0\right) = \mathbf{e}_i$) and

$$g_k = \sum_{j=k}^{i-1} d_j, \quad G_k = \sum_{j=i+1}^{k} d_j.$$

Then we can obtain

$$\|\mathbf{z}\left(t\right)\|_{1D} = \|D\,\mathbf{z}\left(t\right)\| = \sum_{k<i} p_k\left(t\right)g_k + \\ \sum_{k>i} p_k\left(t\right)G_k \geq \begin{cases} p_k\left(t\right)g_k, \; k < i \\ p_k\left(t\right)G_k, \; k > i \end{cases} \quad (13)$$

Therefore $p_k\left(t\right) \leq \begin{cases} \frac{\|\mathbf{z}(t)\|_{1D}}{g_k}, \; k < i \\ \frac{\|\mathbf{z}(t)\|_{1D}}{G_k}, \; k > i \end{cases}$ and the following statement is correct.

Theorem 1. Let a BDP $X(t)$ with rates $\lambda_k(t)$ and $\mu_k(t)$ be given. Assume that there exists a sequence $\{d_k\}$ such that (12) is fulfilled. Then $X(t)$ is exponentially weakly ergodic in $1D$ norm and the following bound holds

$$p_k\left(t\right) \leq \begin{cases} \frac{M(d_{i-1}\Delta_i + d_{i+1}\Lambda_i)}{\alpha\, g_k}, \; k < i \\ \frac{M(d_{i-1}\Delta_i + d_{i+1}\Lambda_i)}{\alpha\, G_k}, \; k > i \end{cases}, \quad (14)$$

for any k.

Two-sided truncations was firstly mentioned in our previous work [6]. We considered truncated BDP on state space $N_1, N_1+1, \ldots, N_2$ with intensities $\lambda_k^*(t) = \lambda_k(t)$, $N_1 \leq k < N_2$, and $\mu_k^*(t) = \mu_k(t)$, $N_1 < k \leq N_2$ and supposed other birth and death rates equal to zero. We denoted by $A^*(t)$, $\mathbf{p}^*(t)$ and so on the correspondent characteristics of truncated BDP. Then using our general approach and considering corresponding differential equations, we obtained

$$e^{-\int\limits_s^t \alpha^*(\tau)\,d\tau} \leq M^* e^{-\alpha^*(t-s)}, \quad (15)$$

for any $0 \leq s \leq t$, instead of (12), see the bounds and other details in [6], Put $d = \min\left(d_{i-1}, d_{i+1}\right)$ and $W = \inf\limits_k\left(\frac{g_k}{k}, \frac{d}{i}, \frac{G_k}{k}\right)$.

Therefore, the following statements are correct.

Theorem 2. Let birth-death processes $X(t)$ and $X^*(t)$ be such that (12) and (15) hold. Let $\mathbf{p}\left(0\right) = \mathbf{p}^*\left(0\right) = e_i$ (i. e., $X(0) = X^*(0) = i$). Then the following bounds hold:

$$\|\mathbf{p}\left(t\right) - \mathbf{p}^*\left(t\right)\| \leq \\ \frac{4M\,M^*\left(\Delta_i\, d_{i-1}^* + \Lambda_i\, d_{i+1}^*\right)}{d\alpha\,\alpha^*} \\ \cdot\left(\frac{g_{N_1-1}\Delta_{N_1}}{g_{N_1}^*} + \frac{G_{N_2+1}\Lambda_{N_2}}{G_{N_2}^*}\right), \quad (16)$$

and

$$\|\mathbf{p}\left(t\right) - \mathbf{p}^*\left(t\right)\|_{1E} \leq \\ \frac{4M\,M^*\left(\Delta_i\, d_{i-1}^* + \Lambda_i\, d_{i+1}^*\right)}{W\alpha\,\alpha^*} \\ \cdot\left(\frac{g_{N_1-1}\Delta_{N_1}}{g_{N_1}^*} + \frac{G_{N_2+1}\Lambda_{N_2}}{G_{N_2}^*}\right). \quad (17)$$

Corollary 1. Let under assumptions of Theorem 2 $N_2 = \infty$. Then the following bounds hold

$$\|\mathbf{p}\left(t\right) - \mathbf{p}^*\left(t\right)\| \leq \\ \frac{4M\,M^*\left(\Delta_i\, d_{i-1}^* + \Lambda_i\, d_{i+1}^*\right)g_{N_1-1}\Delta_{N_1}}{d\alpha\,\alpha^* g_{N_1}^*}, \quad (18)$$

$$\|\mathbf{p}(t) - \mathbf{p}^*(t)\|_{1E} \leq \frac{4M\,M^*\left(\Delta_i\, d^*_{i-1} + \Lambda_i\, d^*_{i+1}\right) g_{N_1-1}\Delta_{N_1}}{W\alpha\,\alpha^* g^*_{N_1}}, \tag{19}$$

for any $i > N_1$.

Corollary 2. Let under assumptions of Theorem 2 $N_1 = 0$. Then the following bounds hold

$$\|\mathbf{p}(t) - \mathbf{p}^*(t)\| \leq \frac{4M\,M^*\left(\Delta_i\, d^*_{i-1} + \Lambda_i\, d^*_{i+1}\right) G_{N_2+1}\Lambda_{N_2}}{d\alpha\,\alpha^* G^*_{N_2}}, \tag{20}$$

$$\|\mathbf{p}(t) - \mathbf{p}^*(t)\|_{1E} \leq \frac{4M\,M^*\left(\Delta_i\, d^*_{i-1} + \Lambda_i\, d^*_{i+1}\right) G_{N_2+1}\Lambda_{N_2}}{W\alpha\,\alpha^* G^*_{N_2}}, \tag{21}$$

for any $i < N_2$.

THE $M_t|M_t|S$ QUEUEING MODEL

Now we consider non-stationary $M_t|M_t|S$ queuing model with S servers, and intensities of arrival and service of a customer $\lambda_k(t) = \lambda(t)$ and $\mu_k(t) = \mu(t)\min(k,S)$ respectively if there is k customers in the queue.

Let $X = X(t),\ t \geq 0$ be a queue-length process for the $M_t|M_t|S$ queuing system. This is a BDP with the birth and death rates $\lambda_k(t) = \lambda$, if $k \geq 0$ and $\mu_k(t) = k\mu(t)$, if $k \leq S$ or $\mu_k(t) = S\mu(t)$, if $k > S$ respectively.

Let $i_1 < i < S-1 < S$. Put $d_k = 1$, if $i_1 \leq k \leq S$, $d_k = \zeta_k d_{k-1}$, if $k > S$ and $d_{k-1} = \xi_k d_k$, if $k < i_1$.

Rewrite (6) in the form

$$\alpha_k(t) = \begin{cases} \lambda + S\mu - \zeta_{k+2}\lambda - \frac{1}{\zeta}_{k+1} S\mu,\ k \geq S \\ \lambda + S\mu - \zeta_{S+1}\lambda - (S-1)\mu,\ k = S-1 \\ \mu,\ i_1 < k < S \\ \lambda + (i_1+1)\mu - \lambda - i_1\xi_{i_1-1}\mu,\ k = i_1 \\ \lambda + (k+1)\mu - \frac{1}{\xi}_k\lambda - k\xi_{k-1}\mu,\ k < i_1 \end{cases} \tag{22}$$

Then we have

$$\alpha_k(t) \geq \begin{cases} (1-\zeta_{k+2})\Lambda + (1-\frac{1}{\zeta}_{k+1})S\,m,\ k \geq S \\ m + (1-\zeta_{S+1})\Lambda,\ k = S-1 \\ m,\ i_1 < k < S-1 \\ (i_1 + 1 - i_1\xi_{i_1-1})\,m,\ k = i_1 \\ l(1-\frac{1}{\xi}_k) + (k+1-k\xi_{k-1})\Delta,\ k < i_1\, and\, k+1-k\xi_{k-1} < 0 \\ l(1-\frac{1}{\xi}_k) + (k+1-k\xi_{k-1})m,\ k < i_1\, and\, k+1-k\xi_{k-1} > 0 \end{cases}, \tag{23}$$

where $l \leq \lambda(t) \leq \Lambda, \quad m \leq \mu(t) \leq \Delta.$

Let now $\{d_k\}$ be a sequence of positive numbers such that

$$\alpha_k = c_1,\ k \geq S,$$

$$\alpha_k = c_2,\ k \leq i_1,$$

where

$$c_1 = m + (1-\zeta_{S+1})\Lambda,$$

$$\zeta_k = 1 + \frac{mS - c_1}{\Lambda} - \frac{mS}{\zeta_{k-1}\Lambda}, \tag{24}$$

$$c_2 = (i_1 + 1 - i_1\xi_{i_1-1})m,$$

and

$$\xi_{k-1} = \frac{(k+1)\Delta + l - c_2}{k\Delta} - \frac{l}{k\Delta\xi_k}, \tag{25}$$

if $k+1-k\xi_{k-1} < 0$, or

$$\xi_{k-1} = \frac{(k+1)m + l - c_2}{km} - \frac{l}{km\xi_k},. \tag{26}$$

if $k+1-k\xi_{k-1} > 0$.

We can rewrite (25) and (26) in the form

$$\xi_{k-1} = 1 + \frac{1}{k} + \frac{l(\xi_k - 1) - c_2\xi_k}{km\xi_k}, \tag{27}$$

if $\xi_{k-1} < 1 + \frac{1}{k}$, i. e. $l(\xi_k - 1) - c_2\xi_k < 0$,

$$\xi_{k-1} = 1 + \frac{1}{k} + \frac{l(\xi_k - 1) - c_2\xi_k}{kM\xi_k}, \tag{28}$$

if $\xi_{k-1} > 1 + \frac{1}{k}$, i. e. $l(\xi_k - 1) - c_2\xi_k > 0$.

Then we obtain

$$\alpha_k \geq \min(c_1, c_2) \text{ for } S - i_1 = 3,$$

and

$$\alpha_k \geq \min(c_1, m, c_2) \text{ for } S - i_1 > 3.$$

Hence Theorem 2 implies the following statement.

Theorem 3. Let there exist a sequence $\{d_k\}$ of positive numbers such that the corresponding c_1, c_2 be positive. Then queue-length process $X(t)$ is exponentially ergodic and bounds of Theorem 2 hold.

Remark. Consider $M_t|M_t|S|S+K$, this is a Markovian queueing model with S servers and finite number K of waiting rooms. The corresponding queue-length process $X(t)$ is BDP with finite state space $\{0, 1, \ldots, S+K\}$ and arrival and service rates $\lambda_k(t) = \lambda(t)$, $\mu_k(t) = \mu(t)\min(k,S)$ respectively. Then the same bounds for two-sided truncations hold for sufficiently large K.

EXAMPLE

Let $X = X(t),\ t \geq 0$ be a queue-length process for the $M_t|M_t|S$ with $S = 200$ and periodical rates:

$$\lambda(t) = 250 + 50 \sin 2\pi t,$$

$$\mu(t) = 2.5 + 0.5 \cos 2\pi t,$$

$$\mu_k(t) = k\mu(t), \quad k \leq 200,$$

$$\mu_k(t) = 200\mu(t), \quad k > 200.$$

Let $i = 198$.

Then we have $i_1 = 197;\ m = 2;\ \Delta = 3;\ l = 200;\ \Lambda = 300$.

And

$$\ldots, d_{196} = 1.0048,\ d_{197} = 1,$$

$$d_{199} = 1, d_{200} = 1.0063,\ \ldots$$

We obtain our sequence $\{d_k\}$ using formulas (24),(27) and (28):

$c_1 \approx 0.11; c_2 \approx 0.1088$, i.e. $\alpha_k = 0.1$.

We have

$$\ldots,\ d^*_{197} = 1, d^*_{196} = 1.0048, \ldots$$

$$\ldots,\ d^*_{199} = 1,\ d^*_{200} = 1.0063, \ldots,$$

where $\{d^*\}$ can be chosen as a geometric progression with $q_1 = 1.0048$ and $q_2 = 1.0063$ respectively.

Then $N_1 = 30, N_2 = 330, W = 0.005$.

Hence Theorems 2,3 imply the following bounds:

$$\|\mathbf{p}\,(t) - \mathbf{p}^*\,(t)\,\| \leq 10^{-5}, \tag{29}$$

$$\|\mathbf{p}\,(t) - \mathbf{p}^*\,(t)\,\|_{1E} \leq 0.002. \tag{30}$$

Figures 1 and 2 show the limiting mean $E(t, 100)$ on the ”final” interval [29,30] and the behaviour of the mean $E(t, 100)$ on the whole time interval [0,30].

Figures 3–6 show the probabilities of some ’essential’ states of the queue-length process $p_k = \mathsf{P}(X(t) = k$ with the initial condition $X(0) = 100$ on the ”final” interval [29,30] and the behaviour of the mean $E(t, 30)$ on the whole time interval [0,30].

CONCLUSIONS

The paper deals with the problem of existence and construction of limiting characteristics for time-inhomogeneous birth and death processes which is important for queueing applications. For this purpose we calculate the limiting characteristics for the process via the construction of two-sided uniform in time truncation bounds. We consider the $M_t|M_t|S$ queueing model and obtain uniform approximation bounds of two-sided truncations in the case of sufficiently large traffic intensity. A numerical example is considered. Our approach to truncations of the state space can be used in modeling information flows related to high-performance computing. The development of methodology for other classes of inhomogeneous Markovian queueing models seems to be a promising direction of research.

ACKNOWLEDGEMENTS

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008) (work of Ksenia Kiseleva), by the Russian Foundation for Basic Research, projects no. 15-01-01698, 15-07-05316, and by the RAS Presidium Program No. I.33P (project 0063-2016-0015).

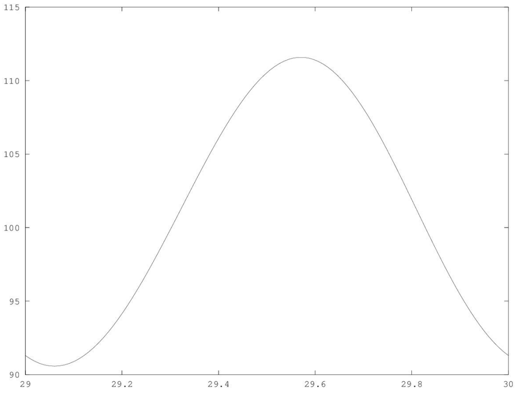


Fig. 1. Approximation of the limiting mean $E(t, 100)$ on $[29, 30]$.

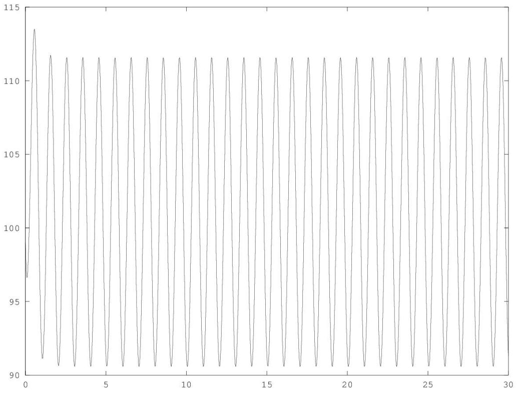


Fig. 2. Approximation of the mathematical expectation of the length of queue $E(t, 100)$ on $[0, 30]$.

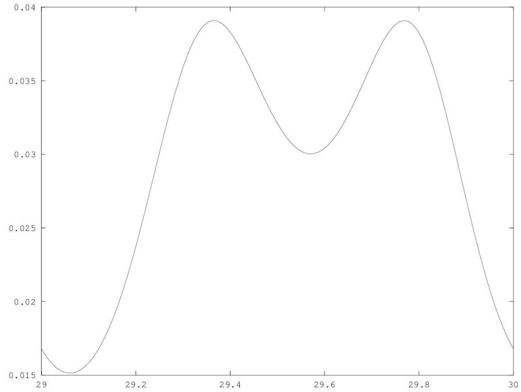


Fig. 3. Approximation of the limiting probability p_{105} on $[29, 30]$.

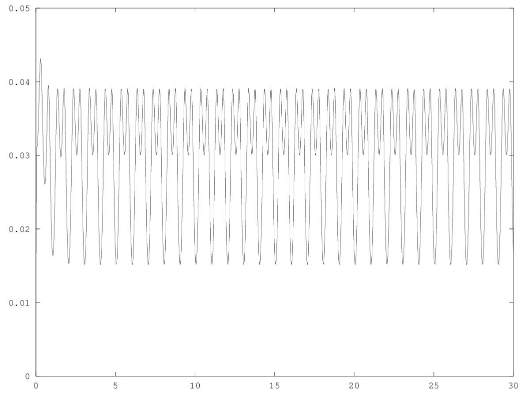


Fig. 4. Approximation of the probability p_{105} on $[0, 30]$.

REFERENCES

[1] Daleckij, Ju.L., Krein, M.G. (1974). Stability of solutions of differential equations in Banach space. Amer. Math. Soc. Transl. **43**.

[2] Di Crescenzo A., Nobile, A. G. (1995). Diffusion approximation to a queueing system with time dependent arrival and service rates // Queueing Syst., **19**, 41–62.

[3] E. A. Van Doorn, A. I. Zeifman, T. L. Panfilova (2010). Bounds and asymptotics for the rate of convergence of birth-death processes // Th. Prob. Appl., **54**, 97–113.

[4] B. L. Granovsky, A. I. Zeifman (2004). Nonstationary Queues: Estimation of the Rate of Convergence // Queueing Syst. **46**, 363–388.

[5] Mandelbaum A., Massey W. (1995). Strong approximations for time-dependent queues // Math. Oper. Research, **20**, 33–64.

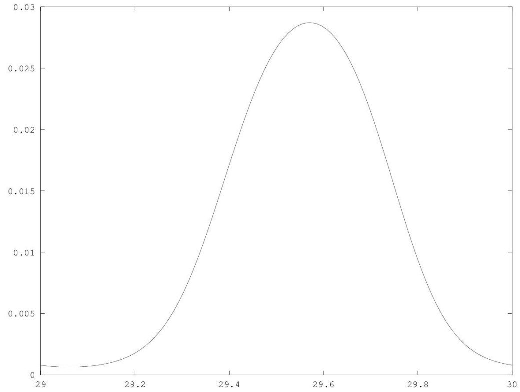


Fig. 5. Approximation of the limiting probability p_{120} on $[29, 30]$.

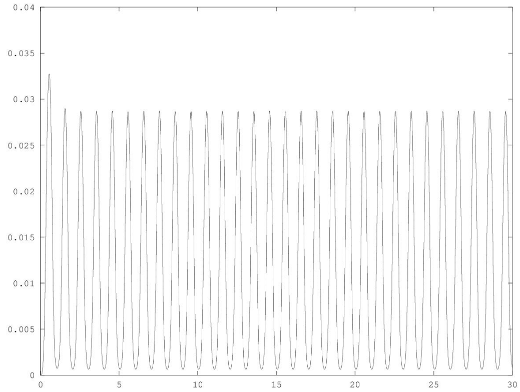


Fig. 6. Approximation of the probability p_{120} on $[0, 30]$.

[6] Satin Y, Korotysheva A., Kiseleva K., Shilova G., Fokicheva E. Zeifman A., Korolev V. (2016). Two-Sided Truncations Of Inhomogeneous Birth-Death Processes // ECMS 2016 Proceedings edited by: Thorsten Claus, Frank Herrmann, Michael Manitz, Oliver Rose European Council for Modeling and Simulation. doi:10.7148/2016–0663.

[7] Zeifman, A.I. (1988). Truncation error in a birth and death system // USSR Computational Mathematics and Mathematical Physics, **28(6)**, 210–211.

[8] A. I. Zeifman (1995). Upper and lower bounds on the rate of convergence for nonhomogeneous birth and death processes // Stoch. Proc. Appl., **59**, 157–173.

[9] A. Zeifman, S. Leorato, E. Orsingher, Ya. Satin, G. Shilova (2006). Some universal limits for nonhomogeneous birth and death processes // Queueing Syst., **52**, 139–151.

[10] A. Zeifman, Ya. Satin, V. Korolev, S. Shorgin (2014). On truncations for weakly ergodic inhomogeneous birth and death processes // International Journal of Applied Mathematics and Computer Science, **24**, 503–518.

[11] Zeifman, A. I.; Korotysheva, A. V.; Korolev, V. Yu.; Satin Ya. A. Truncation bounds for approximations of inhomogeneous continuous-time Markov chains. Th. Prob. Appl. **2016**, 61, 563–569.

AUTHOR BIOGRAPHIES

YACOV SATIN is Candidate of Science (PhD) in physics and mathematics, associate professor, Vologda State University. His email is yacovi@mail.ru.

ANNA KOROTYSHEVA is Candidate of Science (PhD) in physics and mathematics, associate professor, Vologda State University. Her email is a_korotysheva@mail.ru.

GALINA SHILOVA is Candidate of Science (PhD) in physics and mathematics, associate professor, Heard of Department of Mathematics, Vologda State University. Her email is shgn@mail.ru.

ALEXANDER SIPIN is Doctor of Science in physics and mathematics; professor, Vologda State University His email is cac@uni-vologda.ac.ru.

ELENA FOKICHEVA is Candidate of Science (PhD) in physics and mathematics, associate professor, Vologda State University. Her email is eafokicheva2007@yandex.ru.

KSENIA KISELEVA is scientific researcher, RUDN University, Moscow, Russia; PhD student, Vologda State University. Her email is ksushakiseleva@mail.ru.

ALEXANDER ZEIFMAN is Doctor of Science in physics and mathematics; professor, Heard of Department of Applied Mathematics, Vologda State University; senior scientist, Institute of Informatics Problems, Federal Research Center ”Computer Science and Control” of the Russian Academy of Sciences; principal scientist, Institute of Socio-Economic Development of Territories, Russian Academy of Sciences. His email is a_zeifman@mail.ru.

VICTOR KOROLEV is Doctor of Science in physics and mathematics, professor, Head of Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; professor, Hangzhou Dianzi University, Hangzhou, China. His email is victoryukorolev@yandex.ru.

SERGEY SHORGIN is Doctor of Science in physics and mathematics, professor, Deputy Director of the Institute of Informatics Problems of the Federal Research Center ”Computer Science and Control” of the Russian Academy of Sciences. His email is sshorgin@ipiran.ru.

GENERALIZED GAMMA DISTRIBUTIONS AS MIXED EXPONENTIAL LAWS AND RELATED LIMIT THEOREMS

Victor Korolev
Moscow State University,
Moscow, Russia
IPI FRC CSC RAS,
Moscow, Russia
Hangzhou Dianzi University,
Hangzhou, China

Andrey Gorshenin
IPI FRC CSC RAS
Moscow, Russia
Alexander Korchagin
Moscow State University
Moscow, Russia

Alexander Zeifman
Vologda State University,
Vologda, Russia
IPI FRC CSC RAS;
ISEDT RAS
Email: a_zeifman@mail.ru

KEYWORDS

Generalized gamma distribution, Weibull distribution, mixed exponential distribution, mixed Poisson distribution, mixed geometric distribution, strictly stable distribution, normal mixture, random sum, random sample size

ABSTRACT

A theorem due to L. J. Gleser stating that a gamma distribution with shape parameter no greater than one is a mixed exponential distribution is extended to generalized gamma distributions introduced by E. W. Stacy as a special family of lifetime distributions containing both gamma distributions, exponential power and Weibull distributions. It is shown that the mixing distribution is a scale mixture of strictly stable laws concentrated on the nonnegative halfline. As a corollary, the representation is obtained for the mixed Poisson distribution with the generalized gamma mixing law as a mixed geometric distribution. Limit theorems are proved establishing the convergence of the distributions of statistics constructed from samples with random sizes obeying the mixed Poisson distribution with the generalized gamma mixing law including random sums to special normal mixtures.

INTRODUCTION

A. *Motivation*

In most papers dealing with the statistical analysis of meteorological data available to the authors, the suggested analytical models for the observed statistical regularities in precipitation are rather ideal and far from being adequate. For example, it is traditionally assumed that the duration of a wet period (the number of subsequent wet days) follows the geometric distribution (for example, see [1]). Perhaps, this prejudice is based on the conventional interpretation of the geometric distribution in terms of the Bernoulli trials as the distribution of the number of subsequent wet days ("successes") till the first dry day ("failure"). But the framework of Bernoulli trials assumes that the trials are independent whereas a thorough statistical analysis of precipitation data registered in different points demonstrates that the sequence of dry and wet days is not only independent, but it is also devoid of the Markov property so that the framework of Bernoulli trials is absolutely inadequate for analyzing meteorological data.

It turned out that the statistical regularities of the number of subsequent wet days can be very reliably modeled by the negative binomial distribution with the shape parameter less than one. For example, in [2] the data registered in so climatically different points as Potsdam (Brandenburg, Germany) and Elista (Kalmykia, Russia) was analyzed and it was demonstrated that the fluctuations of the numbers of successive wet days with very high confidence fit the negative binomial distribution with shape parameter $r \approx 0.8$. In the same paper a schematic attempt was undertaken to explain this phenomenon by the fact that negative binomial distributions can be represented as mixed Poisson laws with mixing gamma-distributions whereas the Poisson distribution is the best model for the discrete stochastic chaos (see, e. g., [3], [4])

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

and the mixing distribution accumulates the stochastic influence of factors that can be assumed exogenous with respect to the local system under consideration.

In the present paper we try to give further theoretic explanation of the high adequacy of the negative binomial model. For this purpose we use the concept of a mixed geometric law introduced in [5] (also see [6], [7]). Having first proved that any generalized gamma distribution (*GG-distribution*) with shape parameter less than one is mixed exponential and thus generalizing Gleser's similar theorem on gamma-distributions [8], we then prove that any mixed Poisson distribution with the generalized gamma mixing law (*GG-mixed Poisson distribution*) is actually mixed geometric. The mixed geometric distribution can be interpreted in terms of the Bernoulli trials as follows. First, as a result of some "preliminary" experiment the value of some random variables taking values in $[0,1]$ is determined which is then used as the probability of success in the sequence of Bernoulli trials in which the original "unconditional" mixed Poisson random variable is nothing else than the "conditionally" geometrically distributed random variable having the sense of the number of trials up to the first failure. This makes it possible to assume that the sequence of wet/dry days is not independent, but is conditionally independent and the random probability of success is determined by some outer stochastic factors. As such, we can consider the seasonality or the type of the cause of a rainy period.

The obtained results can serve as a theoretical explanation of some mixed models used within the popular Bayesian approach to the statistical analysis of lifetime data related to high performance information systems.

B. Notation and definitions

In the paper, conventional notation is used. The symbols $\stackrel{d}{=}$ and $\Longrightarrow$ denote the coincidence of distributions and convergence in distribution, respectively. The integer and fractional parts of a number z will be respectively denoted $[z]$ and $\{z\}$.

In what follows, for brevity and convenience, the results will be presented in terms of random variables (r.v:s) with the corresponding distributions. It will be assumed that all the r.v:s are defined on the same probability space $(\Omega,\ \mathfrak{F},\ \mathsf{P})$.

A r.v. having the gamma distribution with shape parameter $r>0$ and scale parameter $\lambda>0$ will be denoted $G_{r,\lambda}$,

$$\mathsf{P}(G_{r,\lambda}<x)=\int_0^x g(z;r,\lambda)dz,$$

with

$$g(x;r,\lambda)=\frac{\lambda^r}{\Gamma(r)}x^{r-1}e^{-\lambda x},\ \ x\geq 0,$$

where $\Gamma(r)$ is Euler's gamma-function, $\Gamma(r)=\int_0^\infty x^{r-1}e^{-x}dx$, $r>0$.

In these notation, obviously, $G_{1,1}$ is a r.v. with the standard exponential distribution: $\mathsf{P}(G_{1,1}<x)=\big[1-e^{-x}\big]\mathbf{1}(x\geq 0)$ (here and in what follows $\mathbf{1}(A)$ is the indicator function of a set A).

The gamma distribution is a particular representative of the class of generalized gamma distributions (GG-distributions), which were first described in [9] as a special family of lifetime distributions containing both gamma distributions and Weibull distributions.

Definition 1. A *generalized gamma distribution* (*GG-distribution*) is the absolutely continuous distribution defined by the density

$$g^*(x;r,\gamma,\lambda)=\frac{|\gamma|\lambda^r}{\Gamma(r)}x^{\gamma r-1}e^{-\lambda x^\gamma},\qquad x\geq 0,$$

with $\gamma\in\mathbb{R}$, $\lambda>0$, $r>0$.

The properties of GG-distributions are described in [9], [10]. In what follows we will be interested only in GG-distributions with $\gamma\in(0,1]$. A r.v. with the density $g^*(x;r,\gamma,\lambda)$ will be denoted $G^*_{r,\gamma,\lambda}$.

For a r.v. with the Weibull distribution, a particular case of GG-distributions corresponding to the density $g^*(x;1,\gamma,1)$ and the distribution function (d.f.) $\big[1-e^{-x^\gamma}\big]\mathbf{1}(x\geq 0)$, we will use a special notation W_γ. Thus, $G_{1,1}\stackrel{d}{=}W_1$. It is easy to see that $W_1^{1/\gamma}\stackrel{d}{=}W_\gamma$.

A r.v. with the standard normal d.f. $\Phi(x)$ will be denoted X,

$$\mathsf{P}(X<x)=\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-z^2/2}dz,\qquad x\in\mathbb{R}.$$

A r.v. having the Laplace distribution corresponding to the density $f^\Lambda(x)=\frac{1}{2}e^{-|x|}$, $x\in\mathbb{R}$, will be denoted Λ.

The d.f. and the density of a strictly stable distribution with the characteristic exponent α and shape parameter θ defined by the characteristic function (ch.f.)

$$\mathfrak{f}_{\alpha,\theta}(t)=\exp\big\{-|t|^\alpha\exp\{-\tfrac{1}{2}i\pi\theta\alpha\mathrm{sign}t\}\big\},\qquad t\in\mathbb{R},$$

wheter $0<\alpha\leq 2$, $|\theta|\leq\min\{1,\frac{2}{\alpha}-1\}$, will be respectively denoted $F_{\alpha,\theta}(x)$ and $f_{\alpha,\theta}(x)$ (see, e. g., [11]). A r.v. with the d.f. $F_{\alpha,\theta}(x)$ will be denoted $S_{\alpha,\theta}$. To symmetric strictly stable distributions there correspond the value $\theta=0$ and the ch.f. $\mathfrak{f}_{\alpha,0}(t)=e^{-|t|^\alpha}$, $t\in\mathbb{R}$. Hence, it is easy to see that $S_{2,0}\stackrel{d}{=}\sqrt{2}X$.

To one-sided strictly stable distributions concentrated on the nonnegative halfline there correspond the values $\theta = 1$ and $0 < \alpha \leq 1$. The pairs $\alpha = 1$, $\theta = \pm 1$ correspond to the distributions degenerate in ± 1, respectively. All the rest strictly stable distributions are absolutely continuous. Stable densities cannot be explicitly represented via elementary functions with four exceptions: the normal distribution ($\alpha = 2$, $\theta = 0$), the Cauchy distribution ($\alpha = 1$, $\theta = 0$), the Lévy distribution ($\alpha = \frac{1}{2}$, $\theta = 1$) and the distribution symmetric to the Lévy law ($\alpha = \frac{1}{2}$, $\theta = -1$).

According to the ¡¡multiplication theorem¿¿ (see, e. g., theorem 3.3.1 in [11]) for any admissible pair of parameters (α, θ) and any $\alpha' \in (0, 1]$ the multiplicative representation $S_{\alpha\alpha',\theta} \stackrel{d}{=} S_{\alpha,\theta} \cdot S_{\alpha',1}^{1/\alpha}$ holds, in which the multipliers on the right-hand side are independent. In particular, for any $\alpha \in (0, 2]$

$$S_{\alpha,0} \stackrel{d}{=} X\sqrt{2S_{\alpha/2,1}}, \tag{1}$$

that is, any symmetric strictly stable distribution is a normal scale mixture.

Let $p \in (0, 1)$. By V_p we denote a r.v. having the *geometric distribution* with parameter p: $\mathsf{P}(V_p = k) = p(1-p)^k$, $k = 0, 1, 2, ...$ This means that for any $m \in \mathbb{N}$

$$\mathsf{P}(V_p \geq m) = \sum\nolimits_{k=m}^{\infty} p(1-p)^k = (1-p)^m.$$

DEFINITION 2. Let Y be a r.v. taking values in the interval $(0, 1)$. Moreover, let for all $p \in (0, 1)$ the r.v:s Y and V_p are independent. Let $N = V_Y$, that is,

$$\mathsf{P}(N \geq m) = \int_0^1 (1-y)^m d\mathsf{P}(Y < y)$$

for any $m \in \mathbb{N}$. The distribution of the r.v. N will be called *Y-mixed geometric*.

MAIN RESULTS

In the paper [8] it was shown that any gamma distribution with shape parameter no greater than one is mixed exponential. For convenience, formulate this result as the following lemma.

LEMMA 1 [8]. *The density of a gamma distribution $g(x; r, \mu)$ with $0 < r < 1$ can be represented as*

$$g(x; r, \mu) = \int_0^\infty ze^{-zx} p(z; r, \mu)dz,$$

where

$$p(z; r, \mu) = \frac{\mu^r}{\Gamma(1-r)\Gamma(r)} \cdot \frac{\mathbf{1}(z \geq \mu)}{(z-\mu)^r z}.$$

Moreover, a gamma distribution with shape parameter $r > 1$ cannot be represented as a mixed exponential distribution.

LEMMA 2 [12]. *For $r \in (0, 1)$ let $G_{r,\,1/2}$ and $G_{1-r,\,1/2}$ be independent gamma-distributed r.v:s. Let $\mu > 0$, $p \in (0, 1)$. Then the density $p(z; r, \mu)$ in lemma 1 corresponds to the r.v.*

$$Z_{r,\mu} = \mu(G_{r,\,1/2} + G_{1-r,\,1/2})/G_{r,\,1/2}.$$

LEMMA 3 [13]. *Let $\alpha \in (0, 1]$. Then $W_\alpha \stackrel{d}{=} W_1 \cdot S_{\alpha,1}^{-1}$ with the r.v:s on the right-hand side being independent.*

LEMMA 4. *A d. f. $F(x)$ with $F(0) = 0$ corresponds to a mixed exponential distribution if and only if the function $1 - F(x)$ is completely monotone: $F \in C_\infty$ and $(-1)^{n+1}F^{(n)}(x) \geq 0$ for all $x > 0$.*

This statement immediately follows from the Bernstein theorem [14].

THEOREM 1. *Let $\alpha \in (0, 1]$, $r \in (0, 1)$, $\mu > 0$. Then the GG-distribution with parameters r, α, μ is a mixed exponential distribution: $G^*_{r,\alpha,\mu} \stackrel{d}{=} W_1 \cdot \big(S_{\alpha,1} Z_{r,\mu}^{1/\alpha}\big)^{-1}$ with the r.v:s on the right-hand side being independent. Moreover, a GG-distribution with $\alpha r > 1$ cannot be represented as mixed exponential.*

PROOF. Prove the first assertion of the theorem. First, note that $\mathsf{P}(G_{r,\mu}^{1/\alpha} > x) = \mathsf{P}(G_{r,\mu} > x^\alpha)$. Hence, according to lemma 1 for $x \geq 0$ we have

$$\mathsf{P}(G_{r,\mu}^{1/\alpha} > x) = \mathsf{P}(W_1 > Z_{r,\mu}x^\alpha) =$$

$$= \int_0^\infty e^{-zx^\alpha} p(z; r, \mu)dz = \int_0^\infty \mathsf{P}(W_\alpha > xz^{1/\alpha})p(z; r, \mu)dz,$$

that is, $G_{r,\mu}^{1/\alpha} \stackrel{d}{=} W_\alpha \cdot Z_{r,\mu}^{-1/\alpha}$. Now apply lemma 3 and obtain

$$G_{r,\mu}^{1/\alpha} \stackrel{d}{=} W_1 \cdot \big(S_{\alpha,1} Z_{r,\mu}^{1/\alpha}\big)^{-1}. \tag{2}$$

Second, it is easy to see that

$$G_{r,\mu}^{1/\alpha} \stackrel{d}{=} G^*_{r,\alpha,\mu} \tag{3}$$

for any $r > 0$, $\mu > 0$ and $\alpha > 0$. Now the desired assertion follows from (2) and (3).

To prove the second assertion, assume that $\alpha r > 1$ and the r.v. $G^*_{r,\alpha,\mu}$ has a mixed exponential distribution. By lemma 4 this means that the function $\psi(s) = \mathsf{P}(G^*_{r,\alpha,\mu} > s)$, $s \geq 0$, is completely monotone. But $\psi'(s) = g^*(s; r, \alpha, \mu) \geq 0$ for all $s \geq 0$, whereas $\psi''(s) = (g^*)'(s; r, \alpha, \mu) = \frac{\alpha\mu^r}{\Gamma(r)} s^{\alpha r - 2} e^{-\mu s^\alpha}\big((\alpha r - 1) - \mu\alpha s^\alpha\big) \leq 0$, only if $(\alpha r - 1) - \mu\alpha s^\alpha \leq 0$, that is, $s \geq s_0 \equiv \big[(\alpha r - 1)/\mu\alpha\big]^{1/\alpha} > 0$, and $\psi''(s) \geq 0$ for $s \in (0, s_0) \neq \varnothing$ contradicting the complete monotonicity of $\psi(s)$ and thus proving the second assertion. The theorem is proved.

DEFINITION 3. For $r > 0$, $\alpha \in \mathbb{R}$ and $\mu > 0$ let $\Pi_{r,\alpha,\mu}$ be a r.v. with the *GG-mixed Poisson distribution*

$$\mathsf{P}(\Pi_{r,\alpha,\mu} = k) = \frac{1}{k!}\int_0^\infty e^{-z}z^k g^*(z; r, \alpha, \mu)dz,$$

$k = 0, 1, 2...$

Since negative binomial distributions are mixed Poisson laws with gamma-mixing distributions [15], [4] the class of GG-mixed Poisson laws contains negative binomial distributions ($\alpha = 1$). Moreover, it also contains Poisson-Weibull distributions ($r = 1$) [6].

THEOREM 2. *If* $r \in (0, 1]$, $\alpha \in (0, 1]$ *and* $\mu > 0$, *then a GG-mixed Poisson distribution is a* $Y_{r,\alpha,\mu}$-*mixed geometric distribution*:

$$\mathsf{P}(\Pi_{r,\alpha,\mu} = k) = \int_0^1 y(1-y)^k d\mathsf{P}(Y_{r,\alpha,\mu} < y),$$

$k = 0, 1, 2...$, *where*

$$Y_{r,\alpha,\mu} \overset{d}{=} \frac{S_{\alpha,1}Z_{r,\mu}^{1/\alpha}}{1 + S_{\alpha,1}Z_{r,\mu}^{1/\alpha}} \overset{d}{=}$$

$$\overset{d}{=} \frac{\mu^{1/\alpha}S_{\alpha,1}(G_{r,1/2} + G_{1-r,1/2})^{1/\alpha}}{G_{r,1/2}^{1/\alpha} + \mu^{1/\alpha}S_{\alpha,1}(G_{r,1/2} + G_{1-r,1/2})^{1/\alpha}}, \quad (4)$$

where the r.v:s $S_{\alpha,1}$ *and* $Z_{\mu,r}$ *or* $S_{\alpha,1}$, $G_{r,1/2}$ *and* $G_{1-r,1/2}$ *are independent.*

PROOF. Using theorem 1 we have

$$\mathsf{P}(\Pi_{r,\alpha,\mu} = k) = -\frac{1}{k!}\int_0^\infty e^{-z}z^k d\mathsf{P}(G^*_{r,\alpha,\mu} > z) =$$

$$= -\frac{1}{k!}\int_0^\infty e^{-z}z^k d\mathsf{P}(W_1 > S_{\alpha,1}Z_{r,\mu}^{1/\alpha}z) =$$

$$= \frac{1}{k!}\int_0^\infty x\biggl(\int_0^\infty e^{-z(1+x)}z^k dz\biggr)d\mathsf{P}(S_{\alpha,1}Z_{r,\mu}^{1/\alpha} < x) =$$

$$= \frac{\Gamma(k+1)}{k!}\int_0^\infty \frac{x}{(1+x)^{k+1}}d\mathsf{P}(S_{\alpha,1}Z_{r,\mu}^{1/\alpha} < x) =$$

$$= \int_0^\infty \frac{x}{1+x}\Bigl(1 - \frac{x}{1+x}\Bigr)^k d\mathsf{P}(S_{\alpha,1}Z_{r,\mu}^{1/\alpha} < x).$$

Changing the variables $\frac{x}{1+x} \longmapsto y$, we finally obtain

$$\mathsf{P}(\Pi_{r,\alpha,\mu} = k) = \int_0^1 y(1-y)^k d\mathsf{P}\bigl(S_{\alpha,1}Z_{r,\mu}^{1/\alpha} < \tfrac{y}{1-y}\bigr) =$$

$$= \int_0^1 y(1-y)^k d\mathsf{P}\biggl(\frac{S_{\alpha,1}Z_{r,\mu}^{1/\alpha}}{1 + S_{\alpha,1}Z_{r,\mu}^{1/\alpha}} < y\biggr). \quad (5)$$

Moreover, (5) and lemma 2 yield representation (4). The theorem is proved.

REMARK 1. With the account of lemma 1 it is easy to verify that the density $q(y; r, \alpha, \mu)$ of the r.v. $Y_{r,\alpha,\mu}$ admits the following integral representation via the strictly stable density $f_{\alpha,1}(x)$:

$$q(y; r, \alpha, \mu) = \frac{\mu^r}{\Gamma(1-r)\Gamma(r)} \cdot \frac{\mathbf{1}(0 \le y \le 1)}{(1-y)^2} \times$$

$$\times \int_\mu^\infty f_{\alpha,1}\biggl(\frac{yz^{-1/\alpha}}{1-y}\biggr)\frac{dz}{(z-\mu)^r z^{1+2/\alpha}}.$$

REMARK 2. It is easily seen that the sum $G_{r,1/2} + G_{1-r,1/2}$ in (4) has the exponential distribution with parameter $\frac{1}{2}$. However, the numerator and denominator of the expression on the right-hand side of (4) are not independent.

From (4) we easily obtain the following asymptotic assertion.

COROLLARY 1. *As* $\mu \to 0$, *the r.v.* $Y_{r,\alpha,\mu}$ *is the quantity of order* $\mu^{1/\alpha}$ *in the sense that*

$$\mu^{-1/\alpha}Y_{r,\alpha,\mu} \Longrightarrow S_{\alpha,1}Z_{r,1}^{1/\alpha} \overset{d}{=}$$

$$\overset{d}{=} S_{\alpha,1} \cdot \biggl(\frac{G_{r,1/2} + G_{1-r,1/2}}{G_{r,1/2}}\biggr)^{1/\alpha},$$

where the r.v:s $S_{\alpha,1}$ *and* $Z_{\mu,r}$ *or* $S_{\alpha,1}$, $G_{r,1/2}$ *and* $G_{1-r,1/2}$ *are independent.*

Theorem 1, corollary 1, lemma 3 and theorem 1 of [12] yield the following statement.

THEOREM 3. *If* $r \in (0, 1]$, $\alpha \in (0, 1]$ *and* $\mu > 0$, *then*

$$\mu^{1/\alpha}\Pi_{r,\alpha,\mu} \Longrightarrow \frac{W_1}{S_{\alpha,1}Z_{r,1}^{1/\alpha}} \overset{d}{=}$$

$$\overset{d}{=} W_\alpha \cdot \biggl(\frac{G_{r,1/2}}{G_{r,1/2} + G_{1-r,1/2}}\biggr)^{1/\alpha} \overset{d}{=} G^*_{r,\alpha,\mu}$$

as $\mu \to 0$, *where the r.v:s* W_1, S_α *and* $Z_{r,1}$ *are independent as well as the r.v:s* W_α, $G_{r,1/2}$ *and* $G_{1-r,1/2}$.

LIMIT THEOREMS FOR SUMS OF INDEPENDENT RANDOM VARIABLES IN WHICH THE NUMBER OF SUMMANDS HAS THE GG-MIXED POISSON DISTRIBUTION

Consider a sequence of independent identically distributed (i.i.d.) r.v:s $X_1, X_2, \ldots$ defined on a probability space $(\Omega, \mathfrak{F}, \mathsf{P})$. Assume that $\mathsf{E}X_1 = 0$, $0 < \sigma^2 = \mathsf{D}X_1 < \infty$. For a natural $n \ge 1$ let $S_n = X_1 + \ldots + X_n$. Let $N_1, N_2, \ldots$ be a sequence of nonnegative integer random variables defined on the same probability space so that for each $n \ge 1$ the random variable N_n is independent of the sequence $X_1, X_2, \ldots$ A random sequence $N_1, N_2, \ldots$ is said to be infinitely increasing ($N_n \longrightarrow \infty$) in probability, if $\mathsf{P}(N_n \le m) \longrightarrow 0$ as $n \to \infty$ for any $m \in (0, \infty)$.

LEMMA 5. *Assume that the r.v:s $X_1, X_2, \ldots$ and $N_1, N_2, \ldots$ satisfy the conditions specified above and $N_n \longrightarrow \infty$ in probability as $n \to \infty$. A d.f. $F(x)$ such that*

$$\mathsf{P}\big(S_{N_n} < x\sigma\sqrt{n}\big) \Longrightarrow F(x) \quad (n \to \infty),$$

exists if and only if there exists a d.f. $Q(x)$ satisfying the conditions $Q(0) = 0$,

$$F(x) = \int_0^\infty \Phi\big(x/\sqrt{y}\big)dQ(y), \quad x \in \mathbb{R},$$

$$\mathsf{P}(N_n < nx) \Longrightarrow Q(x) \ \ (n \to \infty).$$

PROOF. This lemma is a particular case of a result proved in [16], the proof of which is, in turn, based on general theorems on convergence of superpositions of independent random sequences [18]. Also see [19], theorem 3.3.2.

Re-denote $n = \mu^{-1/\alpha}$. Then $\mu = 1/n^\alpha$. Consider the r.v. $\Pi_{r,\alpha,1/n^\alpha}$. From theorem 3 it follows that $\Pi_{r,\alpha,1/n^\alpha} \to \infty$ in probability and

$$\frac{\Pi_{r,\alpha,1/n^\alpha}}{n} \Longrightarrow \frac{W_1}{S_{\alpha,1}Z_{r,1}^{1/\alpha}} \overset{d}{=}$$

$$\overset{d}{=} W_\alpha \cdot \left(\frac{G_{r,1/2}}{G_{r,1/2} + G_{1-r,1/2}}\right)^{1/\alpha} \qquad (6)$$

as $n \to \infty$, where in each term the involved r.v:s are independent.

Now from (6), lemma 5 with $N_n = \Pi_{r,\alpha,1/n^\alpha}$, (1) and the well-known relation $\Lambda \overset{d}{=} X\sqrt{2W_1}$ we directly obtain

THEOREM 4. *Assume that the random variables $X_1, X_2, \ldots$ and $N_1, N_2, \ldots$ satisfy the conditions specified above. Let $r \in (0, 1]$, $\alpha \in (0, 1]$. Then*

$$\frac{S_{\Pi_{r,\alpha,1/n^\alpha}}}{\sigma\sqrt{n}} \Longrightarrow X \cdot \sqrt{G^*_{r,\alpha,\mu}} \overset{d}{=}$$

$$\overset{d}{=} X \cdot \sqrt{\frac{W_1}{S_{\alpha,1}Z_{r,1}^{1/\alpha}}} \overset{d}{=} \frac{\Lambda}{\sqrt{2S_{\alpha,1}Z_{r,1}^{1/\alpha}}}$$

as $n \to \infty$, where in each term the involved r.v:s are independent.

LIMIT THEOREMS FOR STATISTICS CONSTRUCTED FROM SAMPLES WITH RANDOM SIZES HAVING THE GG-MIXED POISSON DISTRIBUTIONS

Consider a sequence of i.i.d. r.v:s $X_1, X_2, \ldots$ defined on a probability space $(\Omega, \mathfrak{F}, \mathsf{P})$. Let $N_1, N_2, \ldots$ be a sequence of nonnegative integer random variables defined on the same probability space so that for each $n \ge 1$ the random variable N_n is independent of the sequence $X_1, X_2, \ldots$ A random sequence $N_1, N_2, \ldots$ is said to be infinitely increasing ($N_n \longrightarrow \infty$) in probability, if $\mathsf{P}(N_n \le m) \longrightarrow 0$ as $n \to \infty$ for any $m \in (0, \infty)$.

For $n \ge 1$ let $T_n = T_n(X_1, \ldots, X_n)$ be a statistic, that is, a measurable function of the random variables $X_1, \ldots, X_n$. For each $n \ge 1$ define the random variable T_{N_n} by letting $T_{N_n}(\omega) = T_{N_n(\omega)}\big(X_1(\omega), \ldots, X_{N_n(\omega)}(\omega)\big)$ for every elementary outcome $\omega \in \Omega$. We will say that the statistic T_n is asymptotically normal, if there exists $\vartheta \in \mathbb{R}$ such that

$$\mathsf{P}\left(\sqrt{n}(T_n - \vartheta) < x\right) \Longrightarrow \Phi(x) \quad (n \to \infty). \qquad (7)$$

LEMMA 6. *Assume that $N_n \longrightarrow \infty$ in probability as $n \to \infty$. Let the statistic T_n be asymptotically normal in the sense of* (7). *Then a distribution function $F(x)$ such that*

$$\mathsf{P}\left(\sqrt{n}(T_{N_n} - \vartheta) < x\right) \Longrightarrow F(x) \quad (n \to \infty),$$

exists if and only if there exists a distribution function $Q(x)$ satisfying the conditions $Q(0) = 0$,

$$F(x) = \int_0^\infty \Phi\big(x\sqrt{y}\big)dQ(y), \quad x \in \mathbb{R},$$

$$\mathsf{P}(N_n < nx) \Longrightarrow Q(x) \ \ (n \to \infty).$$

This lemma is a particular case of theorem 3 in [17], the proof of which is, in turn, based on general theorems on convergence of superpositions of independent random sequences [18]. Also see [19], theorem 3.3.2.

From (6), lemma 5 with $N_n = \Pi_{r,\alpha,1/n^\alpha}$ and (1) with the account of the easily verified property of GG-distributions $(G^*_{r,\alpha,\mu})^{-1} \overset{d}{=} G^*_{r,-\alpha,\mu}$ we directly obtain

THEOREM 5. *Let $r \in (0, 1]$, $\alpha \in (0, 1]$. Let the statistic T_n be asymptotically normal in the sense of* (7). *Then*

$$\sqrt{n}\big(T_{\Pi_{r,\alpha,1/n^\alpha}} - \vartheta\big) \Longrightarrow X \cdot \sqrt{G^*_{r,-\alpha,\mu}} \overset{d}{=}$$

$$\overset{d}{=} X \cdot \sqrt{\frac{S_{\alpha,1}Z_{r,1}^{1/\alpha}}{W_1}} \overset{d}{=} S_{2\alpha,0} \cdot \sqrt{\frac{Z_{r,1}^{1/\alpha}}{2W_1}}$$

as $n \to \infty$, where in each term the involved r.v:s are independent.

Remark 3. The distribution of the limit r.v. in theorem 4 is a special case of the so-called *generalized variance gamma distributions*, see [10]. If $\alpha = 1$, then $S_{\alpha,1} \equiv 1$ and according to lemma 1 the limit law in theorem 4 turns into that of the r.v. $X\sqrt{Z_{r,1}W_1^{-1}} \stackrel{d}{=} XG_{r,1}^{-1/2}$, that is, the Student distribution with $2r$ degrees of freedom (see [20], [4]).

Remark 4. It is worth noting that the mixing GG-distributions in the limit normal scale mixtures in theorems 4 and 5 differ only by the sign of the parameter α.

CONCLUSION

In the paper, a theorem due to L. J. Gleser stating that a gamma distribution with shape parameter no greater than one is a mixed exponential distribution was extended to generalized gamma distributions introduced by E. W. Stacy as a special family of lifetime distributions containing both gamma distributions, exponential power and Weibull distributions. It was shown that the mixing distribution is a scale mixture of strictly stable laws concentrated on the nonnegative halfline. As a corollary, the representation was obtained for the mixed Poisson distribution with the generalized gamma mixing law as a mixed geometric distribution. Limit theorems were proved establishing the convergence of the distributions of statistics constructed from samples with random sizes obeying the mixed Poisson distribution with the generalized gamma mixing law including random sums to special normal mixtures.

The obtained results can serve as a theoretical explanation of some mixed models used within the popular Bayesian approach to the statistical analysis of lifetime data related to high performance information systems.

Acknowledgement

This research was partly supported by the RAS Presidium Program No. I.33P (project 0063-2016-0015) and by the Russian Foundation for Basic Research (project 15-07-04040).

REFERENCES

[1] O. Zolina, C. Simmer, K. Belyaev, S. Gulev, and P. Koltermann. Changes in the duration of European wet and dry spells during the last 60 years // Journal of Climate, 2013. Vol. 26. P. 2022–2047.

[2] V. Yu. Korolev, A. K. Gorshenin, S. K. Gulev, K. P. Belyaev, A. A. Grusho. Statistical analysis of precipitation events // Proceedings of the 14th International Conference of Numerical Analysis and Applied Mathematics ICNAAM 2016, 19-25 September, 2016, Rhodes, Greece. – American Institute of Physics Proceedings, 2017. To appear.

[3] J. F. C. Kingman. Poisson processes. – Oxford: Clarendon Press, 1993.

[4] V. Yu. Korolev, V. E. Bening, S. Ya. Shorgin. Mathematical Foundations of Risk Theory. 2nd ed. – Moscow: FIZMATLIT, 2011 (in Russian).

[5] V. Yu. Korolev. Limit distributions for doubly stochastically rarefied renewal processes and their properties // Theory of Probability and Its Applications, 2016. Vol. 61. No. 4. P. 1–22.

[6] V. Yu. Korolev, A. Yu. Kporchagin, A. I. Zeifman. The Poisson theorem for Bernoulli trials with a random probability of success and a discrete analog of the Weibull distribution // Informatics and its Applications, 2016. Vol. 10. No. 4. P. 11–20.

[7] V. Yu. Korolev, A. Yu. Korchagin, A. I. Zeifman. On doubly stochastic rarefaction of renewal processes // Proceedings of the 14th International Conference of Numerical Analysis and Applied Mathematics ICNAAM 2016, 19-25 September, 2016, Rhodes, Greece. – American Institute of Physics Proceedings, 2017. To appear.

[8] L. J. Gleser. The gamma distribution as a mixture of exponential distributions // American Statistician, 1989. Vol. 43. P. 115–117.

[9] E. W. Stacy. A generalization of the gamma distribution // Annals of Mathematical Statistics, 1962. Vol. 33. P. 1187–1192.

[10] L. M. Zaks, V. Yu. Korolev. Generalized variance gamma distributions as limit laws for random sums // Informatics and Its Applications, 2013. Vol. 7. No. 1. P. 105–115.

[11] *Zolotarev V. M.* One-Dimensional Stable Distributions. – Providence, R.I.: American Mathematical Society, 1986.

[12] V. Yu. Korolev. Analogs of Gleser's theorem for negative binomial and generalized gamma distributions and some their applications // Informatics and Its Applications, 2017. Vol. 11. To appear.

[13] V. Yu. Korolev. Product representations for random variables with the Weibull distributions and their applications // Journal of Mathematical Sciences, 2016. Vol. 218. No. 3. P. 298–313.

[14] S. N. Bernstein. Sur les fonctions absolument monotones // Acta Mathematica, 1928. Vol. 52. P. 1–66. doi:10.1007/BF02592679.

[15] M. Greenwood and G. U. Yule. An inquiry into the nature of frequency-distributions of multiple happenings, etc. // J. Roy. Statist. Soc., 1920. Vol. 83. P. 255–279.

[16] *Korolev V. Yu.* Convergence of random sequences with independent random indices. I // Theory Probab. Appl., 1994. Vol. 39, No. 2. P. 313–333.

[17] *Korolev V. Yu.* Convergence of random sequences with independent random indices. II // Theory Probab. Appl., 1995. Vol. 40, No. 4. P. 907–910.

[18] *Korolev V. Yu.* A general theorem on the limit behavior of superpositions of independent random processes with applications to Cox processes // Journal of Mathematical Sciences, 1996. Vol. 81, No. 5. P. 2951–2956.

[19] *Gnedenko B. V., Korolev V. Yu.* Random Summation: Limit Theorems and Applications // Boca Raton: CRC Press, 1996.

[20] V. E. Bening, V. Yu. Korolev. On an application of the Student distribution in the theory of probability and mathematical statistics // Theory of Probability and Its Applications, 2005. Vol. 49. No. 3. P. 377–391.

AUTHOR BIOGRAPHIES

VICTOR KOROLEV is Doctor of Science in physics and mathematics, professor, Head of Department of Mathematical Statistics, Faculty

of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; professor, Hangzhou Dianzi University, Hangzhou, China. His email is victoryukorolev@yandex.ru.

ANDREY GORSHENIN is Candidate of Science (PhD) in physics and mathematics, associate professor, leading scientist, Institute of Informatics Problems, Federal Research Center ”Computer Science and Control” of the Russian Academy of Sciences. His email is agorshenin@frccsc.ru.

ALEXANDER KORCHAGIN is junior scientist, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University. His email is sasha.korchagin@gmail.com.

ALEXANDER ZEIFMAN is Doctor of Science in physics and mathematics; professor, Heard of Department of Applied Mathematics, Vologda State University; senior scientist, Institute of Informatics Problems, Federal Research Center ”Computer Science and Control” of the Russian Academy of Sciences; principal scientist, Institute of Socio-Economic Development of Territories, Russian Academy of Sciences. His email is a_zeifman@mail.ru.

SYSTEM PERFORMANCE OF A VARIABLE-CAPACITY BATCH-SERVICE QUEUE WITH GEOMETRIC SERVICE TIMES AND CUSTOMER-BASED CORRELATION

Jens Baetens[1]
Bart Steyaert[1]
Dieter Claeys[1,2,3]
Herwig Bruneel[1]
[1]SMACS Research Group
Department of Telecommunications and Information Processing
Ghent University
St-Pietersnieuwstraat 41, B 9000 Gent, Belgium
[2]Lean Enterprise Research Center
Department of Industrial Systems Engineering and Product Design
Ghent University
Technologiepark 903, 9052 Zwijnaarde (Gent), Belgium
[3]Department of Agile and Human Centered Production and Robotic Systems
Flanders Make
E-mail:{jens.baetens, bart.steyaert,dieter.claeys,herwig.bruneel}@ugent.be

KEYWORDS

Queueing Performance; Batch Service; Variable Server Capacity; Two Classes; Geometric Service Times; Customer-based Correlation

ABSTRACT

In many queueing systems the server processes several customers simultaneously. Although the capacity of a batch, that is the number of customers that can be processed simultaneously, is often variable in practice, nearly all batch-service queueing models in literature consider a constant capacity. In this paper, we extend previous work on a batch-service queueing model with variable server capacity, where customers of two classes are accommodated in a common first-come-first-served single-server queue. We include correlation between the classes of consecutive customers, and the service times are geometrically distributed. We establish the equations that govern the system behaviour, the stability condition, and an expression for the steady-state probability generating function of the system occupancy at random slot boundaries. In addition, some numerical results are shown to study the impact of the mean service times and of the customer-based correlation in the arrival process on the performance of the queueing system.

INTRODUCTION

In manufacturing environments, a single machine often has the capability to process multiple products simultaneously in a single batch. The maximum number of products that can be processed at the same time, also called the service capacity, is often assumed to be a constant and bounded value, e.g. Banerjee and Gupta 2012; Claeys et al. 2012, 2013; Goswami et al. 2006; Weng and Leachman 1993. The difference with multi-server systems is that a newly arrived customer cannot join an ongoing service, even if the service is not completely full. In manufacturing, this can be found in for instance a furnace where a heating phase cannot be interrupted. In real systems, the maximum batch size or service capacity is often variable and stochastic, which has been studied in only a few papers. In many of these papers, the service capacity is generically distributed and does not depend on any parameter of the system, like the number of customers waiting in the queue, see the papers of Chaudhry and Chang 2004; Germs and Foreest 2010; Pradhan et al. 2015; Sikdar and Samanta 2016. A more detailed description of these papers can be found in our paper (Baetens et al. 2016). Also, Germs and Foreest (2013) have developed an algorithmic method to analyse the performance of continuous-time batch service queueing systems with arrival process, service time distribution and variable service capacity that depend on the number of customers in the queue.

None of the previous papers on batch service considered multiple customer classes. Differentiated service is common in priority queueing and polling systems, where the system can use different scheduling algorithms to optimize the performance of the system (e.g. Reddy et al. (1993); Boxma et al. (2008); Dorsman et al. (2012)).In contrast with polling systems and priority queueing, that use a unique queue for each possible class of customers, we use a common queue with a global First-Come-First-Served service discipline, because it is not always feasible to install a multi-queue system due to certain constraints like the increased cost of a more complicated system. An example of such a

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

system is a furnace in a production line. A furnace can handle multiple customers simultaneously as long as the products must be heated to the same temperature and for the same duration.

In this paper, we analyse the performance of a two-class discrete-time batch-service queueing model, with a variable service capacity that depends on the queue size and on the specific classes of the successive customers. This kind of queueing systems can be found in many telecommunication technologies like optical burst switching networks (Chen et al. 2004) and wireless local area networks (Lu et al. 2005). To the best of our knowledge, the combination of batch service with variable capacity and multiple customer classes has only appeared in our previous paper Baetens et al. (2016). The difference with that paper is that we relax the assumption of fixed single-slot service times by considering geometric service times instead. We also introduce correlation between the classes of consecutive customers in order to model the tendency for same-class customers to arrive in clusters.

The paper is structured as follows. In the next section, we describe the discrete-time two-class queueing model with batch service in detail. In the third section we establish the system equations, from which we deduce the stability condition, and derive a closed-form expression for the steady-state probability generating function (pgf) of the system occupancy at random slot boundaries. Next, using the expressions obtained in this part, we evaluate the performance of the system through some numerical examples and study the impact of the mean service time and the degree of correlation between the classes of consecutive customers on the system performance. Finally, we draw some conclusions about the obtained results.

MODEL DESCRIPTION

In this paper we study a discrete-time two-class queueing system with infinite queue size. This system uses a batch server with a stochastic service capacity based on the class of the customer at the head of the queue. We distinguish two different customer classes in the arrival process, which we will call class A and B. When the server is idle or becomes available at the start of a new slot and finds a non-empty queue, a new service is initiated immediately. The capacity of this service is determined by the number of consecutive customers at the head of the queue that are of the same class which means it depends on the class of the batch being processed. More specifically, the server starts serving a batch of n customers if and only if one of the following two cases occurs:

- Exactly n customers are present and they are all of the same class.
- More than n customers are present, the n customers at the front of the queue are of the same class and the $(n+1)$-th customer is of the other class.

We define the class of a batch as the class of the customers within it. The aggregated numbers of customer arrivals in consecutive slots are modelled as a sequence of independent and identically distributed (i.i.d.) random variables, with common probability mass function (pmf) $e(n)$ and pgf $E(z)$. The mean of these i.i.d. random variables is denoted as λ. The class of a newly arrived customer depends on the class of the previous customer. If the previous customer was of class A (B), then the newly arrived customer will also be of class A (B) with probability (w.p.) α (β). If $\alpha + \beta > 1$, same-class customers will have a tendency to arrive in clusters. The service time of a random batch follows a geometric distribution with mean μ, and does not depend on the class of the processed batch and its size. The pgf of the service time distribution is defined as

$$S(z) = \frac{z}{z + \mu(1-z)} \ .$$

ANALYSIS

In this section, we first determine the system equations that capture the system behaviour. In the next part we analyse the conditions under which the system is stable. In the main part, we establish the steady-state pgf of the system occupancy, that is the number of customers in the system at the beginning of a random slot, including those in the ongoing service if the server is not idle.

System Equations

We start by defining the variables we use in the system equations that capture the behaviour of the system at consecutive slot boundaries. The number of customers in the system, or the system occupancy at random slot boundaries, is denoted by u_k. The number of customers in the system while the server is idle and the previously processed batch is of class A or B is respectively denoted as $u_{I,A,k}$ and $u_{I,B,k}$. We also define the random variables $u_{A,k}$ and $u_{B,k}$ as the system occupancy at random slot boundaries if the server is processing a class A or B batch during slot k.

If the server is idle in slot k, then the server will remain idle if there are no new arrivals. On the other hand, when the server finds at least one newly arrived customer, then a new batch is started immediately. The probability that the class of the first arrival in slot k is of class A or B depends on the class of the most recently processed batch which is why we distinguish between the two cases. The system equations if the previous batch was a class A batch are given by

$$\begin{aligned} &u_{I,A,k+1} = 0 \ \text{ if } e_k = 0 \\ &u_{A,k+1} = e_k \ \text{ if } e_k > 0 \ \& \ \text{ first arrival of class } A \\ &u_{B,k+1} = e_k \ \text{ if } e_k > 0 \ \& \ \text{ first arrival of class } B, \end{aligned} \quad (1)$$

where e_k represents the number of arrivals during slot k. The analogous equations if the previous batch was of class B are

$$\begin{aligned}
u_{I,B,k+1} &= 0 \text{ if } e_k = 0 \\
u_{A,k+1} &= e_k \text{ if } e_k > 0 \text{ \& first arrival of class } A \\
u_{B,k+1} &= e_k \text{ if } e_k > 0 \text{ \& first arrival of class } B. \quad (2)
\end{aligned}$$

On the other hand, if the server is processing a class A batch during slot k and the service period does not end in slot $k+1$, then the service continues and the newly arrived customers are added to the tail of the queue. However, if the service ends, then the behaviour will be determined by the probability that all customers in the system at service initiation of the batch, processed during slot k, were of the same class. In the case that all waiting customers were of the same class, then the behaviour will be as if the server was idle in the previous slot. On the other hand, if at least one customer in the queue was a class B customer, then the next batch will always be a class B batch. The resulting system equations in the case of an ongoing class A batch are

$$\begin{aligned}
u_{A,k+1} =& u_{A,k} + e_k \text{ if Service not done} \\
u_{I,A,k+1} =& 0 \text{ if Service done} \\
& \text{\& } c_{A,k} = u_{A,\text{ini},k} \text{ \& } se_k = 0 \\
u_{A,k+1} =& se_k \text{ if Service done \& } c_{A,k} = u_{A,\text{ini},k} \\
& \text{\& } se_k > 0 \text{ \& first arrival of class } A \\
u_{B,k+1} =& se_k \text{ if Service done \& } c_{A,k} = u_{A,\text{ini},k} \\
& \text{\& } se_k > 0 \text{ \& first arrival of class } B \\
u_{B,k+1} =& u_{A,\text{ini},k} - c_{A,k} + se_k \\
& \text{if Service done \& } c_{A,k} < u_{A,\text{ini},k} \ , \quad (3)
\end{aligned}$$

where se_k, $c_{A,k}$ ($c_{B,k}$) and $u_{A,ini,k}$ ($u_{B,ini,k}$) represent respectively the number of arrivals, the service capacity and the system occupancy at initiation of the ongoing service during slot k which is of class A (B). The analogous system equations for the case that the ongoing service is of class B are given by

$$\begin{aligned}
u_{B,k+1} =& u_{B,k} + e_k \text{ if Service not done} \\
u_{I,B,k+1} =& 0 \text{ if Service done} \\
& \text{\& } c_{B,k} = u_{B,\text{ini},k} \text{ \& } se_k = 0 \\
u_{B,k+1} =& se_k \text{ if Service done \& } c_{B,k} = u_{B,\text{ini},k} \\
& \text{\& } se_k > 0 \text{ \& first arrival of class } B \\
u_{A,k+1} =& se_k \text{ if Service done \& } c_{B,k} = u_{B,\text{ini},k} \\
& \text{\& } se_k > 0 \text{ \& first arrival of class } A \\
u_{A,k+1} =& u_{B,\text{ini},k} - c_{B,k} + se_k \\
& \text{if Service done \& } c_{B,k} < u_{B,\text{ini},k} \ . \quad (4)
\end{aligned}$$

Stability Condition

In this part we analyse a system in which the server is always busy and the variable server capacity is always smaller than the number of waiting customers, also called a saturated system. In such a system, the size of the processed batches is geometrically distributed and a class A and B batch are processed alternately. The system is stable when the mean number of customer arrivals during two consecutive service periods, which is equal to $2\mu\lambda$, is less than the mean number of customers that leave the system during the same two service periods. Since a class A and B batch leave the system during this time period, the mean number of customers that leave the system is the sum of the mean batch size of a class A and B batch. The batch size of a class A and B batch follow a geometric distribution with parameter α or β respectively. The stability condition is then given by

$$2\mu\lambda < \frac{1}{1-\alpha} + \frac{1}{1-\beta} \ .$$

If α or β is equal to 1, then the system will always be stable, since all customers that arrive are of the same class. The server can then group all waiting customers, which leaves an empty queue after service initiation. This also follows from the stability condition, which is reduced to $\lambda < \infty$ under the restriction that α or β is equal to 1. Another element of interest in the stability condition, is the maximum allowed arrival rate, which reaches a minimum value for $\alpha = \beta = 0.5$. Finally, the load ρ of the system is defined as the fraction of λ versus the maximum allowed arrival rate, which leads to

$$\rho := \frac{2\mu\lambda}{\frac{1}{1-\alpha} + \frac{1}{1-\beta}} = 2\mu\lambda \frac{(1-\alpha)(1-\beta)}{2-\alpha-\beta} < 1 \ .$$

The stability condition implies that the load is smaller than 1.

System Occupancy

Assuming the stability condition is met, we can define the steady-state pmf of the system occupancy at random slot boundaries, as

$$u(i) := \lim_{k\to\infty} \Pr[u_k = i] \ ,$$

with corresponding pgf

$$U(z) := \sum_{i=0}^{\infty} u(i) z^i \ .$$

We can split the generating function of the system occupancy $U(z)$ in two parts based on the class of the most recently initiated service. This leads to

$$U(z) = u_{I,A} + u_{I,B} + U_A(z) + U_B(z) \ ,$$

where we introduced the following definitions

$$\begin{aligned}
u_I, A &:= \lim_{k\to\infty} \Pr[u_{I,A,k} = 0] \ , \\
u_I, B &:= \lim_{k\to\infty} \Pr[u_{I,B,k} = 0] \ , \\
U_A(z) &:= \sum_{i=1}^{\infty} \lim_{k\to\infty} \Pr[u_{A,k} = i] z^i \ , \\
U_B(z) &:= \sum_{i=1}^{\infty} \lim_{k\to\infty} \Pr[u_{B,k} = i] z^i \ .
\end{aligned}$$

The probability that the server is idle and the previously initiated service contained class A customers, denoted by $u_{I,A}$, is found by invoking system equations

Eq. (1) and Eq. (3).

$$u_{I,A} = \frac{S(E(0))}{1-E(0)} \frac{U_A(\alpha)}{\mu\alpha} \quad . \tag{5}$$

Using Eq. (2) and Eq. (4), we find the analogous equation if the previous batch is of class B

$$u_{I,B} = \frac{S(E(0))}{1-E(0)} \frac{U_B(\beta)}{\mu\beta} \quad . \tag{6}$$

The partial pgf $U_A(z)$ of the system occupancy when the server is processing a class A batch can be split based on the state of the server in the previous slot. This leads to

$$\begin{aligned} U_A(z) &= E[z^{u_{A,k+1}}] \\ &= E[z^{u_{A,k+1}} I_{\{u_{I,A,k}=0\}}] + E[z^{u_{A,k+1}} I_{\{u_{I,B,k}=0\}}] \\ &\quad + E[z^{u_{A,k+1}} I_{\{u_{A,k}>0\}}] + E[z^{u_{A,k+1}} I_{\{u_{B,k}>0\}}] \ , \end{aligned} \tag{7}$$

where $I_{\{C\}}$ are indicator functions which are equal to 1 if event C occurs and zero otherwise. Invoking the system equations in Eq. (1), we can write the first term of the right-hand side of Eq. (7) as

$$E[z^{u_{A,k+1}} I_{\{u_{I,A,k}=0\}}] = \alpha u_{I,A}(E(z)-E(0)) \quad . \tag{8}$$

Analogously, by using Eq. (2), we obtain

$$\begin{aligned} &E[z^{u_{A,k+1}} I_{\{u_{I,B,k}=0\}}] \\ &\qquad = (1-\beta) u_{I,B}(E(z)-E(0)) \quad . \end{aligned} \tag{9}$$

Using Eq. (3), we obtain the following equation for the third term of the right hand side of Eq. (7)

$$\begin{aligned} E[z^{u_{A,k+1}} I_{\{u_{A,k}>0\}}] &= (1-\frac{1}{\mu}) U_A(z) E(z) \\ &\quad + \frac{\alpha}{\mu}\Big(\frac{U_A(\alpha)E(\alpha)}{\alpha S(E(\alpha))} - \frac{U_A(\alpha)}{\alpha} S(E(0))\Big) \\ &\quad \cdot \frac{S(E(z))-S(E(0))}{1-S(E(0))} \quad . \end{aligned} \tag{10}$$

The last term of $U_A(z)$ results in

$$\begin{aligned} &E[z^{u_{A,k+1}} I_{\{u_{B,k}>0\}}] \\ &= \frac{1-\beta}{\mu}\Big(\frac{U_B(\beta)E(\beta)}{\beta S(E(\beta))} - \frac{U_B(\beta)}{\beta} S(E(0))\Big) \\ &\quad \cdot \frac{S(E(z))-S(E(0))}{1-S(E(0))} + \frac{1-\beta}{\mu(z-\beta)} \\ &\quad \cdot \Big(\frac{U_B(z)E(z)}{S(E(z))} - \frac{z}{\beta}\frac{U_B(\beta)E(\beta)}{S(E(\beta))}\Big) S(E(z)) \quad . \end{aligned} \tag{11}$$

By combining Eqs. (8-11), we obtain

$$\begin{aligned} U_A(z)\Big(\mu - (\mu-1)E(z)\Big) &= \frac{1-\beta}{z-\beta} U_B(z) E(z) \\ &\quad + S(E(0))\Big(\frac{E(z)-E(0)}{1-E(0)} - \frac{S(E(z))-S(E(0))}{1-S(E(0))}\Big) \\ &\quad \cdot \Big(U_A(\alpha) + \frac{1-\beta}{\beta} U_B(\beta)\Big) + \frac{S(E(z))-S(E(0))}{1-S(E(0))} \\ &\quad \cdot \Big(\frac{U_A(\alpha)E(\alpha)}{S(E(\alpha))} + \frac{1-\beta}{\beta}\frac{U_B(\beta)E(\beta)}{S(E(\beta))}\Big) \\ &\quad - \frac{(1-\beta)z}{\beta(z-\beta)} \frac{U_B(\beta)E(\beta)}{S(E(\beta))} S(E(z)) \quad . \end{aligned} \tag{12}$$

Multiplying both sides of Eq. (12) by $\frac{S(E(z))}{E(z)}$ results in

$$\begin{aligned} U_A(z) &= \frac{1-\beta}{z-\beta} U_B(z) S(E(z)) + S(E(0)) \frac{S(E(z))}{E(z)} \\ &\quad \cdot \Big(\frac{E(z)-E(0)}{1-E(0)} - \frac{S(E(z))-S(E(0))}{1-S(E(0))}\Big) \\ &\quad \cdot \Big(U_A(\alpha) + \frac{1-\beta}{\beta} U_B(\beta)\Big) + \frac{S(E(z))-S(E(0))}{1-S(E(0))} \\ &\quad \cdot \frac{S(E(z))}{E(z)}\Big(\frac{U_A(\alpha)E(\alpha)}{S(E(\alpha))} + \frac{1-\beta}{\beta}\frac{U_B(\beta)E(\beta)}{S(E(\beta))}\Big) \\ &\quad - \frac{(1-\beta)z}{\beta(z-\beta)} \frac{U_B(\beta)E(\beta)}{S(E(\beta))} \frac{S(E(z))^2}{E(z)} \quad . \end{aligned} \tag{13}$$

A similar analysis leads to an equation for the partial pgf of the system occupancy if the customer is processing a class B batch

$$\begin{aligned} U_B(z) &= \frac{1-\alpha}{z-\alpha} U_A(z) S(E(z)) + S(E(0)) \frac{S(E(z))}{E(z)} \\ &\quad \cdot \Big(\frac{E(z)-E(0)}{1-E(0)} - \frac{S(E(z))-S(E(0))}{1-S(E(0))}\Big) \\ &\quad \cdot \Big(U_B(\beta) + \frac{1-\alpha}{\alpha} U_A(\alpha)\Big) + \frac{S(E(z))-S(E(0))}{1-S(E(0))} \\ &\quad \cdot \frac{S(E(z))}{E(z)}\Big(\frac{U_B(\beta)E(\beta)}{S(E(\beta))} + \frac{1-\alpha}{\alpha}\frac{U_A(\alpha)E(\alpha)}{S(E(\alpha))}\Big) \\ &\quad - \frac{(1-\alpha)z}{\alpha(z-\alpha)} \frac{U_A(\alpha)E(\alpha)}{S(E(\alpha))} \frac{S(E(z))^2}{E(z)} \quad . \end{aligned} \tag{14}$$

Using Eqs. (5), (6), (13) and (14), we obtain

$$\begin{aligned} &U(z)\Big((z-\alpha)(z-\beta) - (1-\alpha)(1-\beta)S(E(z))^2\Big) \\ &= \frac{S(E(0))}{1-E(0)}\frac{U_A(\alpha)}{\mu\alpha} + \frac{S(E(0))}{1-E(0)}\frac{U_B(\beta)}{\mu\beta} \\ &\quad + \Big(\frac{E(z)-E(0)}{1-E(0)} - \frac{S(E(z))-S(E(0))}{1-S(E(0))}\Big) \\ &\quad \cdot S(E(0))\frac{S(E(z))}{E(z)}\Bigg[(z-\alpha)(z-\beta)\Big(\frac{U_A(\alpha)}{\alpha} + \frac{U_B(\beta)}{\beta}\Big) \\ &\quad + (1-\alpha)S(E(z))(z-\alpha-\beta z+\alpha z)\frac{U_A(\alpha)}{\alpha} \\ &\quad + (1-\beta)S(E(z))(z-\beta-\alpha z+\beta z)\frac{U_B(\beta)}{\beta}\Bigg] \\ &\quad + \frac{S(E(z))-S(E(0))}{1-S(E(0))}\frac{S(E(z))}{E(z)}\Bigg[\Big((z-\alpha)(z-\beta) \\ &\quad + (1-\alpha)S(E(z))(z-\alpha-\beta z+\alpha z)\Big)\frac{U_A(\alpha)E(\alpha)}{\alpha S(E(\alpha))} \\ &\quad + \Big((z-\alpha)(z-\beta) + (1-\beta)S(E(z)) \\ &\quad \cdot (z-\beta-\alpha z+\beta z)\Big)\frac{U_B(\beta)E(\beta)}{\beta S(E(\beta))}\Bigg] \\ &\quad - (1-\beta)z\Big(z-\alpha+(1-\alpha)S(E(z))\Big) \\ &\quad \cdot \frac{U_B(\beta)E(\beta)}{\beta S(E(\beta))}\frac{S(E(z))^2}{E(z)} - \frac{U_A(\alpha)E(\alpha)}{\alpha S(E(\alpha))}\frac{S(E(z))^2}{E(z)} \\ &\quad \cdot (1-\alpha)z\Big(z-\beta+(1-\beta)S(E(z))\Big) \quad . \end{aligned} \tag{15}$$

In Eq. (15), the two remaining unknowns $U_A(\alpha)$ and $U_B(\beta)$ still have to be calculated. With the theorem of Rouché, we can easily prove that the common denominator of these partial pgf's, given by the left hand side of Eq. (15), has two zeros inside or on the unit circle. Each zero of the denominator must also be a zero of the numerator since generating functions are analytical functions inside the complex unit disk and bounded for $|z| \leq 1$. We can easily see that $z = 1$ is a zero of the denominator, which leads to the same condition as the normalisation condition. The other zero can be calculated numerically. The condition that the numerator of $U_A(z)$ is equal to zero for the second zero of the denominator, combined with the condition from the normalisation condition, constitutes a set of two linear equations that leads to a unique solution for the two remaining unknowns.

With these results we can also obtain the result of our previous paper, see Baetens et al. 2016, by using a mean service time $\mu = 1$, which corresponds to single-slot service times and $\beta = 1 - \alpha$. Substituting these assumptions in Eq. (13) and Eq. (14) results in

$$\begin{aligned}
&U_A(z)\big((z-\alpha)(z-1+\alpha) - \alpha(1-\alpha)E(z)^2\big)\\
&\quad=\alpha(z-\alpha)\frac{E(z)-E(0)}{1-E(0)}\Big(z-1+\alpha+(1-\alpha)E(z)\Big)\\
&\qquad\cdot\Big(\frac{U_A(\alpha)}{\alpha}+\frac{U_B(1-\alpha)}{1-\alpha}\Big) - (1-\alpha)zE(z)^2U_A(\alpha)\\
&\qquad - \alpha z(z-\alpha)E(z)\frac{U_B(1-\alpha)}{1-\alpha}\ .
\end{aligned}$$

NUMERICAL RESULTS

In this section we will study the impact of different parameters on the probability u_I that the server is idle, which is given by the sum of $u_{I,A}$ and $u_{I,B}$, and the mean system occupancy. The number of arrivals in each slot follows a geometric distribution with mean arrival rate λ. In Fig. 1, we show the impact of the mean service time μ on the probability that the server is idle. In this figure, the probabilities α and β are both equal to 0.5 and results are obtained for a number of different mean arrival rates. We observe that for all arrival rates, the probability that the server is idle decreases when the mean service time increases, until it reaches the point that the probability is equal to 0 and the system becomes unstable. A higher value for μ results in, on average, longer service periods, which means that the probability that there are no arrivals during a service period decreases. The two requirements for the system to become idle after a service is finished are that all customers at service initiation must be of the same class and there are no arrivals during the service period. Because of its impact on the probability of this second requirement, it is clear that an increase in the mean arrival rate leads to a decrease of the probability that the server is idle.

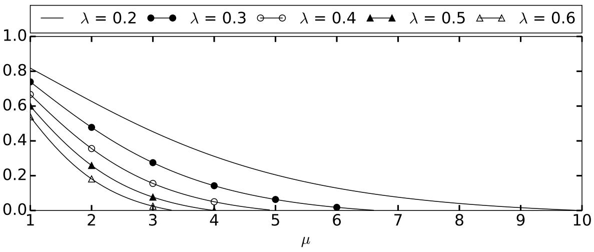


Fig. 1: Impact of Mean Service Time on the Idle Probability using $\alpha = \beta = 0.5$.

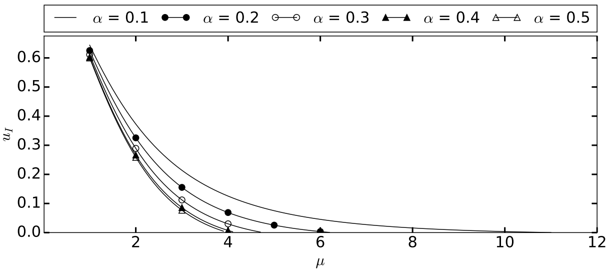


Fig. 2: Impact of Asymmetry in the Arrival Process on the Idle Probability

In Figure 1, we used a symmetric arrival process, that is the probability for a class A and B customer are equal. We can introduce asymmetry in the arrival process by using $\alpha < 0.5$ while keeping $\alpha + \beta = 1$. The impact of a bigger difference between α and β, or an increasing degree of asymmetry in the arrival process, on the idle probability, is shown in Fig. 2 using the mean arrival rate $\lambda = 0.5$. We note that an increasing degree of asymmetry in the arrival process leads to an increased idle probability for the same value of μ. This is because the mean length of a sequence of class A and B customers increases for values of α closer to 0, which in turn increases the probability that all waiting customers are of the same class. This corresponds with the first requirement for the server to jump from a busy state to an idle state. We also note that more asymmetry in the arrival process allows using a slower server, that is a server with a higher mean service time.

In Fig. 3, we analyse the impact of asymmetry in the arrival process on the system occupancy, for an arrival process with mean arrival rate $\lambda = 0.5$ and $\alpha + \beta = 1$. We clearly see that increasing the degree of asymmetry significantly reduces the number of customers in the system, and allows the server to work more slowly while still being stable. The reason for this is that values of α closer to 0 lead to a higher mean length of a sequence of same-class customers, thus allowing the server to process higher service capacity. If the server can process larger batches, the system occupancy will be reduced and the service time must be higher to have the same mean number of customers in the system. We note that the point at which the system becomes unstable is inversely proportional to the parameter α and β as can be seen in the deduction of the stability condition.

In the previous figures, we assumed there was no ten-

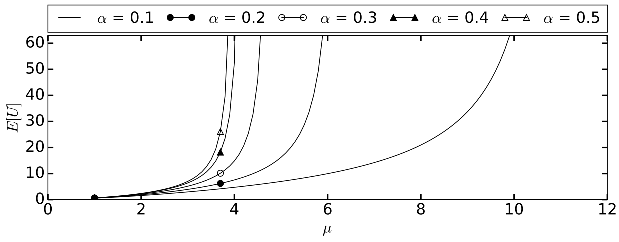


Fig. 3: Impact of Asymmetry in the Arrival Process on the System Occupancy

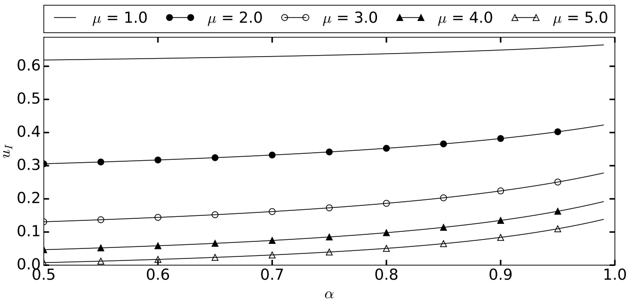


Fig. 4: Impact of Clustering on the Idle Probability

dency for same class customers to arrive in clusters or that $\beta = 1 - \alpha$. In the following results, we study the impact of this tendency for clustering by using values of α and β so that $\alpha + \beta > 1$. In Fig. 4, we first analyse the influence of this tendency on the idle probability, by using $\beta = 0.7$, α varying between 0.5 and 1, and a mean arrival rate $\lambda = 0.5$. We clearly see that increasing α, or the tendency for clustering, also leads to an increasing idle probability. This occurs because using more clustering in the arrival process means the expected length of sequences of same-class customers increases. This leads to a higher probability that all waiting customers are of the same class, which in turn results in a higher probability that the queue is empty after service initiation.

The influence of this tendency for clustering on the mean system occupancy, for the same system configuration as in the previous figure, is shown in Fig. 5. In case of a small mean service time, e.g. $\mu = 4$, we see that increasing the degree of clustering only has a very small influence on the average number of customers in the queue. On the other hand, for larger values of μ, the system is stable only for a certain degree of clustering, and the more clustering in the system, the lower the mean system occupancy. An increased degree of clustering in the arrival process leads to a higher mean length of a sequence of same-class customers, which means that on average more customers can be processed. This increase in the mean service capacity means that the server processes larger batches, which leads to a lower mean system occupancy.

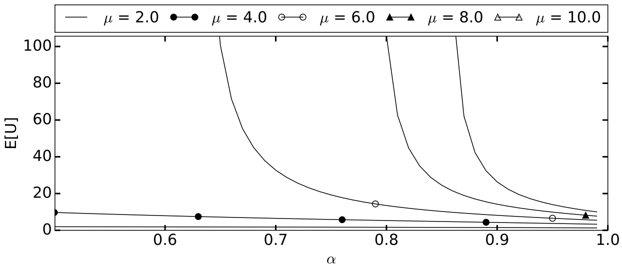


Fig. 5: Impact of Clustering on the System Occupancy

CONCLUSIONS AND FUTURE RESEARCH

In this paper we have analysed the performance of a discrete-time two-class single-server queueing system with variable capacity batch service. The capacity of the batch server is determined by the length of the sequence of same-class customers at the head of the queue at service initiation. The service times of a batch of either class are geometrically distributed, and we also considered correlation between the classes of consecutive customers. During the analysis, we have derived the steady-state pgf of the number of customers in the system, also called the system occupancy, at random slot boundaries. Using the generating function technique, we have demonstrated the impact of the mean service time, asymmetry and clustering in the arrival process on two performance characteristics, more specifically on the idle probability and on the mean system occupancy.

There are a number of possible extensions that could be considered for this model. A first extension would be to find the steady-state pgf for the number of customers that are being processed by the batch server, and the customer delay. In a second extension we could extend the model to use a class-dependent general service time distribution for class A and B batches. Another possible extension is introducing an upper bound for the service capacity. We can also look at systems capable of processing more than 2 classes of customers. We expect that this will introduce an extra level of complexity because the class of the next batch, if not all customers in the queue at service initiation were of the same class, is not deterministic.

REFERENCES

Baetens, J., Steyaert, B., Claeys, D., Bruneel, H., 2016. System occupancy of a two-class batch-service queue with class-dependent variable server capacity. In: International Conference on Analytical and Stochastic Modeling Techniques and Applications. Springer, pp. 32–44.

Banerjee, A., Gupta, U., 2012. Reducing congestion in bulk-service finite-buffer queueing system using batch-size-dependent service. Performance Evaluation 69(1), 53–70.

Boxma, O., van der Wal, J., Yechiali, U., 2008. Polling with batch service. Stochastic Models 24(4), 604–625.

Chaudhry, M., Chang, S., 2004. Analysis of the discrete-time bulk-service queue $Geo/G^Y/1/N + B$. Operations Research Letters 32 (4), 355–363.

Chen, Y., Qiao, C., Yu, X., 2004. Optical burst switching: a new area in optical networking research. IEEE network 18 (3), 16–23.

Claeys, D., Steyaert, B., Walraevens, J., Laevens, K., Bruneel, H., 2012. Tail distribution of the delay in a general batch-service queueing model. Computers and Operations Research 39, 2733–2741.

Claeys, D., Steyaert, B., Walraevens, J., Laevens, K., Bruneel, H., 2013. Analysis of a versatile batch-service queueing model with correlation in the arrival process. Performance Evaluation 70(4), 300–316.

Dorsman, J., der Mei, R. V., Winands, E., 2012. Polling with batch service. OR Spectrum 34, 743–761.

Germs, R., Foreest, N. V., 2010. Loss probabilities for the $M^X/G^Y/1/K+B$ queue. Probability in the Engineering and Informational Sciences 24(4), 457–471.

Germs, R., Foreest, N. V., 2013. Analysis of finite-buffer state-dependent bulk queues. OR Spectrum 35(3), 563–583.

Goswami, V., Mohanty, J., Samanta, S., 2006. Discrete-time bulk-service queues with accessible and non-accessible batches. Applied Mathematics and Computation 182, 898–906.

Lu, K., Wu, D., Fang, Y., Qiu, R. C., 2005. Performance analysis of a burst-frame-based mac protocol for ultrawideband ad hoc networks. In: Communications, 2005. ICC 2005. 2005 IEEE International Conference on. Vol. 5. IEEE, pp. 2937–2941.

Pradhan, S., Gupta, U., Samanta, S., 2015. Queue-length distribution of a batch service queue with random capacity and batch size dependent service: M/g r y/1. OPSEARCH, 1–15.

Reddy, G., Nadarajan, R., Kandasamy, P., 1993. A nonpreemptive priority multiserver queueing system with general bulk service and heterogeneous arrivals. Computers and operations research 20 (4), 447–453.

Sikdar, K., Samanta, S., 2016. Analysis of a finite buffer variable batch service queue with batch markovian arrival process and server's vacation. OPSEARCH 53 (3), 553–583.

Weng, W., Leachman, R., 1993. An improved methodology for real-time production decisions at batch-process work stations. IEEE Transactions on Semiconductor Manufacturing 6 (3), 219–225.

AUTHOR BIOGRAPHIES

JENS BAETENS obtained his master in Computer Science Engineering at Ghent University in 2014. After graduation, he started his doctorate studies at the Stochastic Modelling and Analysis of Communication Systems (SMACS) research group within the department of Telecommunications and Information Processing (TELIN) at Ghent University. His main research interests include discrete-time batch service models and its applications.

BART STEYAERT received his Phd in Engineering Sciences in 2008 from Ghent University, Belgium. Since January 1990, he has been working as a researcher at the SMACS Research Group within the TELIN-Department at the same university. His main research interests include discrete-time queueing models, traffic control, and stochastic modelling of high-speed communications networks.

DIETER CLAEYS obtained his Ph.D. degree in Engineering in 2011 and has since worked at the SMACS research group at Ghent University. Since October 2015, he is an assistant professor at the department of Industrial Systems Engineering and Product Design at Ghent University. His main research interests include methods and time engineering, and the analysis of discrete-time queueing models and its applications to operations research and telecommunications systems.

HERWIG BRUNEEL is full Professor in the Faculty of Engineering and head of the TELIN-Department at Ghent University. He also leads the SMACS Research Group within this department. His main personal research interests include stochastic modeling and analysis of communication systems, and (discrete-time) queueing theory. He has published more than 500 papers on these subjects and is coauthor of the book H. Bruneel and B. G. Kim, "Discrete-Time Models for Communication Systems Including ATM" (Kluwer Academic Publishers, Boston, 1993). From October 2001 to September 2003, he served as the Academic Director for Research Affairs at Ghent University. Since 2009, he holds a career-long Methusalem grant from the Flemish Government at Ghent University, specifically on Stochastic Modeling and Analysis of Communication Systems.

ACKNOWLEDGEMENT

Dieter Claeys is a Postdoctoral Fellow with the Research Foundation Flanders (FWO-Vlaanderen), Belgium. Part of the research has been funded by the Interuniversity Attraction Poles Programma initiated by the Belgian Science Policy Office.

MODELLING FOR ENSURING INFORMATION SECURITY OF THE DISTRIBUTED INFORMATION SYSTEMS

Alexander A. Grusho, Elena E. Timonina and Sergey Ya. Shorgin
Institute of Informatics Problems,
Federal Research Center "Computer Science and Control"
of the Russian Academy of Sciences
Vavilova 44-2, 119333, Moscow, Russia
Email: grusho@yandex.ru, eltimon@yandex.ru, sshorgin@ipiran.ru

KEYWORDS

Information security of the distributed information systems, models of the permitted interactions, statistical models of interactions

ABSTRACT

In the paper the concept of the permitted interactions is defined, i.e. such interactions are necessary for the tasks which are legally started at present. Any other interactions are considered as forbidden.

The main objective is definition, what interactions are permitted during this period of time. For the solution of this problem it is offered to model the needs for interactions depending on existence of legally started tasks. Such modeling is possible on the basis of meta data about the used tasks and their information requirements.

The main idea of control is that the started task appeals to meta data for the permission of interaction with other task, necessary for her decision. On the basis of meta data a need of such interaction is defined and permission which can't be forged or bypassed is given.

INTRODUCTION

The malicious code and harmful influences (Rieck et al.(Eds), 2013; Skorobogatov and Woods, 2012) can move through the distributed information system (DIS), using independently organized interactions of the DIS components. This assertion is true for the distributed DIS components. Therefore it is expedient to control all interactions of the DIS components. First of all it concerns to interactions of software applications. The set of software applications is part of a set of tasks which can be realized in DIS. In this paper software applications are also called tasks.

In the paper the concept of the permitted interactions is defined, i.e. such interactions are necessary for the tasks which are legally started at present. Any other interactions are considered as forbidden.

For such security policy it is necessary to develop special means of its realization. It is easy to construct mechanisms of control of network flows with use of cryptography. However the main objective is definition of component interactions which are allowed during this period of time. For the solution of this problem it is offered to model interactions depending on existence of legally started tasks. Such modelling is possible on the basis of meta data about the used tasks and their information requirements.

The main idea of control is that the started task appeals to meta data for the permission of interaction with other task, necessary for her decision. On the basis of meta data the need of such interaction is defined and permission which can't be forged or bypassed is given. Really interactions are implemented through network by means of sessions and information flows in network.

Usually security of information flows is supported by Firewalls, Proxy servers, Intrusion Detection Systems. These mechanisms work when there can be malicious flows. The paper presents new security mechanism which can be used instead of traditional measures. This is due to the strong limitation for existence of non needed information flow.

In the paper the ways of creation of the required meta data is considered. Historically the first method for creation of meta data are statistical (it was used in Secret Net). However with development of information technologies other ways for creation of meta data were found, tools are developed for their realization.

Emergence of such new methods is connected with constantly arising situations in which the initial statistical method doesn't give the adequate answer.

Control methods of information flows have a long story. They are correctly realized in security policies MLS and Biba (TCSEC, 1985). However these policies are formulated only in terms of information flows. Real security policies consider interactions of all DIS components and are formulated at the levels of tasks (software applications), and hosts (Grusho et al., 2014).

But all mechanisms of their realization are at the lower level of hierarchy (network, computer system). Therefore we define a mapping of interactions of components of the top level to the lower level.

Certificates of open keys (Menezes et al., 1997) which allow a wide arbitrariness in the organization of network interactions are used long ago in problems of the organization of secure communication sessions in networks. Sometimes it is useful. Tracking of such connections with the help of audit allows to indirectly observe leakages of valuable information, or interaction with the risk hosts including a malicious code.

On the opposite side there is the system of permitted connections with the help of a priori set of keys for symmetric cryptography.

In the paper we offer the intermediate way of the organization of communication when session keys are created by the cryptographic center as a result of the positive decision on a possibility of interaction of hosts. This way is more suitable for cases when security policy is defined by interactions of the current business tasks, and demands frequent changes.

The paper is structured as follows. Section 2 introduces two-level hierarchical model of DIS. Section 3 defines the problem of information security in DIS. In Section 4 we construct the special protocol which allows controlling information flows. In the Section 5 we consider the security assessment by means of the constructed protocol. In Section 6 we give examples of meta data formation. In Conclusion we shortly analyze the results.

TWO-LEVEL MODEL OF THE DISTRIBUTED INFORMATION SYSTEMS

The DIS two-level hierarchical model consists of the level of tasks and level of network. The network consists of hosts. The connection equipment allows to connect each host with everyone host without an interactions with other hosts. Communication is implemented only by means of sessions, for example, under the TCP protocol. In further considerations the network equipment will not be considered.

Hosts are computer systems and contain computing resources, information resources and software for the solution of various tasks.

The task is a facility for a transformation of information. It consists of source data which it can receive from other tasks, means of transformation of information, and the output data which will be used by other tasks. According to (Nilsson, 1971) a task can be divided into subtasks (operation of a reduction). Thus, a task A can generate a schedule of subtasks (Fig. 1).

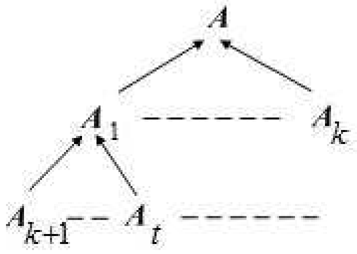


Figure 1: Graphic of subtasks

Let's note that subtasks carry out transformations of information, which results are used in the solution of an initial task A. Some tasks can be solved in other hosts.

Thus it is possible to consider that there is a mapping of a set of tasks and subtasks of a task A into a set of hosts. For any task A, a host on which it is solved, we will denote $H(A)$.

For the convenience tasks and hosts are named DIS components. Hosts on which tasks are solved form the level of network for tasks. All interactions of tasks (collection of information, distribution of information, starting) are carried out through network interactions.

As it was noted above, network interaction is implemented only through communication sessions. Each session is unambiguously defined by the identifiers (addresses, ports, etc.).

Thus, two-level decomposition of DIS (network, tasks) is constructed. Data transmission between hosts is implemented by means of information flows. Thus, each interaction of the DIS components is characterized by a set of information flows.

In traditional networks information flows can form transitive closure, i.e. information flows $V_1 \mapsto V_2$, $V_2 \mapsto V_3$ can generate information flow $V_1 \mapsto V_3$ (can be with a time lag). In certain cases such transitive closure is inadmissible from the point of view of information security. Therefore in DIS it is necessary to build the architecture preventing unauthorized closure of information flows.

INFORMATION SECURITY IN DIS

Let's consider questions of the information security in DIS. For generalizing the model (Grusho et al., 2014, 2015a), we will divide all information flows on admissible flows and non admissible flows.

The permitted interactions are defined by the scheme (Fig. 1), and the task A requests data for itself (ready, or demanding for calculations), i.e. the data flow is directed to A. Then for all levels of the reduction of the task A, and for network schedules the data flows are directed to A or to the following element of the network schedule. When the immediate task is solved, the user or the managing program starts a following task.

Admissible information flows are defined by the permitted interactions of tasks for hosts on which they are solved.

Experience in flows control in security policies (MLS and Biba (TCSEC, 1985)) shows that to construct an information security of DIS basing only on a concept of information flows is not enough because of a possibility of transitive closure of information flows. Therefore the information security models will be constructed on two bases:

- admissible/ non admissible information flows;
- isolation of tasks in computer systems.

Isolation of a task A in a computer system assumes that there is an isolated domain in computer system such that:

- in it there are all source data for solution of task A;
- there is no malicious code in it;
- if on a host two or more tasks are being solved, then their domains are guaranteed to be isolated from each other; necessary data exchange is possible only with the permission of some managing process $\mathcal{N}(H(A))$.

The concept of isolation of a task on a host allows to exclude non admissible transit of information flows. Basic element of providing information security is the model of tasks $\mathcal{N}$ on a host H_0. Model $\mathcal{N}$ is also a task containing a meta information about other tasks.

For its execution the task A on a host $H(A)$ has to address an immediate task A_1 as it follows from scheme of a reduction (Fig. 1). For this purpose the task A generates request for a possibility of the appeal of the task A to the task A_1, and addresses the managing program $\mathcal{N}(H(A))$. Each host H in DIS has in the managing task $\mathcal{N}(H)$ cryptographic facilities and a unique key $k(H)$ for communication with the managing task $\mathcal{N}$ on a host H_0.

The task $\mathcal{N}(H(A))$ forms the encoded message on key $k(H(A))$ with a request to allow interaction of the task A with the task A_1. The task $\mathcal{N}$ checks need of the appeal to A_1 with the help of available for it meta data of the solution of the task A. At the positive decision $\mathcal{N}$ forms the encoded message for the task $\mathcal{N}(H(A))$ in which there is an address of the host $H(A_1)$, number of a port for communication with the task $\mathcal{N}(H(A_1))$ and a session key $k(A, A_1)$. The similar message is also formed for a host $H(A_1)$. The address of the host $H(A)$, port of the task $\mathcal{N}(H(A))$, permission for starting of the task A_1 for the benefit of the task A, and the common key for their communication $k(A, A_1)$ are specified in this message. After obtaining this information the host $H(A)$ initiates a session of the encoded communication with the host $H(A_1)$.

Completion of a session happens standardly. If there is a failure, then it comes to light by means of identification codes MAC (Message Identification Code). In case of need there is a restart of the protocol.If we have agents $\mathcal{N}(H)$ in every host, then there is no need to use all other protocols to control parameters of the net. All necessary control information can be gathered by secure interaction of available hosts with H_0. It can be done by several sets of meta data. One of them may be net control meta data. It helps to forbid service flows.

Cryptographical part of the model resembles well known protocol Kerberos, but it supports different functionality.

SECURITY ASSESSMENT BY MEANS OF THE PROTOCOL

For verification of security of system by means of the offered method it is necessary to prove the following.

1. All admissible information flows are implemented in system.
2. Non admissible information flows, including transit flows, are absent.
3. All failures are identified.

Assertion. Admissible information flows in network are generated by legal interactions of tasks according to the scheme of meta data in $\mathcal{N}$, and non admissible information flows are impossible.

Proof. In the task $\mathcal{N}$ there is information about the required interactions of the task A with other tasks from which it has to obtain source data. If at least one of such tasks A_1 is on other host, i.e. $H(A) \neq H(A_1)$, then according to Protocol A requests at $\mathcal{N}$ an interaction with the task A_1. The task $\mathcal{N}$ finds the host $H(A_1)$ and allows opening of the protected session between $H(A)$ and $H(A_1)$. Information flows of this session are admissible according to the definition. Besides $\mathcal{N}$ can initiate all service information gathering. It can be done by initiating a legal session with any available host. Even protocol "keep alive" (control of perimeter of net) can be constructed in such a way. That proves the first part of the assertion.

Let's assume that the host H wants to organize a session with host H' beyond of an authorization system $\mathcal{N}$ (according to the assumption the UDP connection is forbidden). However the port and, therefore, the software application on the host H' aren't defined. Standard ports in secure system are closed. Communication with any legal task is carried out by means of enciphering. Other ways of information transfer from host H to host H' don't exist in the considered system. Thus, non admissible information flow between hosts H and H', ignoring an authorization system, is impossible in the system, where interactions are under $\mathcal{N}(H)$ control.

It is necessary to check impossibility of unauthorized transitive closure of information flows. If the host H has the permitted session with a host H', and the host H' has the permitted session with a host H'', then two situations are possible.

- Host H' organizes a session with a host H'' for the task A which has generated a session between host H and host H'. Then it is possible a transit information flow from host H to a host H'' and vice versa. However these flows are necessary for the solution of the task A and therefore they are legal information flows.

- If the host H' organizes the permitted session with a host H'' for the decision of a task A', unconnected with the task A, then in the assumption of domains isolation concerning the task A doesn't get into information flow from host H' to a host H''. Also information concerning to task A' doesn't get into information flow from the host H' to the host H. So it follows that there is no exist non admissible transit closure of information flows.

The assertion is proved.

FORMATION OF META DATA

The protocol forbids any connections which aren't reflected as legal in meta data. Therefore questions of completeness, consistency, a possibility of modification, and scalability are connected with the organization of meta data.

Questions of the organization of meta data are connected with possible examples of interrelation of tasks. Without applying for completeness, we will give several such examples.

Example 1. Addressing of a task A to database. In meta data it perhaps the access of the task A to the database on subject T. T is a parameter of the task A. The protocol will organize a session with the host containing the DBMS, but the access is possible only about subject T.

Let's show how it is supported by functionality of Oracle DBMS. The addressing to the task A_1 with parameter T (the appeal to the DBMS with a request from the task A) is equivalent to creation of the user process on a host $H(A_1)$. In Oracle DBMS the user process generates the server process isolated by means of TCB (Trusted Computing Base) (Oracle7 and Trusted Oracle7, 1994). For an isolation of server processes the mechanisms of implementation of discretionary and mandatory security policies are used (TCSEC, 1985). In this regard restriction of T is implemented by the standard policies of access control which are built in the DBMS. The trust to these functions is determined by certification documents (Oracle7 and Trusted Oracle7, 1994).

Example 2. Formation of meta data by means of models of business processes. Formation of meta data about interactions of tasks can be made on the basis of business process modelling methods. A set of advanced methods is developed for business process modeling: IDEF, ARIS, UML, BPMN, etc. (Samuylov et al., 2009).

The methodology of the functional simulation of IDEF0 considers system as a set of actions, each of which will transform some object or a set of objects. These actions correspond in the considered terminology to tasks and information transforms. Application of IDEF0 is presented in the form of hierarchy of the charts connected by cross-references. These models, in particular, allow to estimate distribution of resources for implementation of the target task.

For creation of models an automation software, for example, of BPwin and ERwin (Samuylov et al., 2009) are created. For creation of sequential diagrams it is possible to use methodology of IDEF3 (Samuylov et al., 2009). The methodology of IDEF3 is supported by software of the Computer Associates companies, etc.

The methodology of ARIS assumes several abstraction layers and serves for the complex description of activities of the enterprizes. In this methodology the set of models for the adequate description of system and its processes is created. Software tools are created to support of ARIS which are added by modeling languages of UML, BPMN and etc.

Example 3. Meta data for tasks with uncertainty. It is the most difficult to apply the offered approach of support of information security in DIS to tasks with uncertainties. An example of uncertainty is the question to the task $\mathcal{N}$: "Whether the task A can be solved by means of the task A_1?" For the consideration of such questions it is possible to use semantic methodology. Semantic methods are based on the description of ontologies. The ontology is understood as (Samuylov et al., 2009) hierarchical data structure containing meanings of information and their communications. The formal language of descriptions of ontologies is the standard of Web Ontology Language (Samuylov et al., 2009).

Example 4. The description of admissible communications of tasks by means of data mining. This method of formation of an authorization system is connected with search of such tasks which can be the useful in the analysis of the task A. This method has common features with the example 3. However complication of search in comparison with an example 3 consists that there is no accurate description of the required ontologies. Therefore for creation of an algorithm of the decision about admissible communications for this class of tasks more thin methods of data mining are necessary (Finn (Eds)., 2009).

Example 5. Statistical method of formation of meta data. Let some time the DIS, servicing technological processes of the organization, works in the free mode. Meta data are created by results of observation over interactions of tasks in DIS in the free mode. In some time point all interactions of solvable tasks in each of information technologies are fixed. The received mold of interactions defines meta data for monitoring of further interactions.

In case of such method of formation of meta data next errors are possible:

- some operation modes of information technologies can demand further additional requests for information resources or the software. I.e. in this method we get a bigger number of non admissible interactions, than it is necessary;

- in the free mode some interactions could be excessive, and can generate information flows, dangerous to functioning of information technologies.

Example 6. A method of bans for formation of meta data. Let's assume that all interactions are permitted, except some set of couples of the forbidden interactions (bans). Theory of bans in discrete probability spaces developed since 2011 year (Grusho A. and Timonina E., 2011). Bans form the graph in which vertices are tasks, and edges are bans. This method is rough, but it allows to enter dynamics into system of permissions of interactions. For example, the monitoring system registers events which can be signs of the attacks to assets of the organization. Then for support of information security it is necessary to block urgently any accesses to these assets. At the same time remaining assets need to be still available not to block operation of DIS (Grusho et al., 2015b).

CONCLUSION

Simulation in DIS is widely used for the organization and optimization of computation. An example of successful usage of models for the analysis of behavior of a network is MiniNet [(Lantz et al., 2010). The main idea of efficiency of MiniNet consists that experiments are made on the reduced, cut down information on a traffic.

In this paper it is offered to use the simulation based on the reduced information on solvable tasks and their interactions for implementation of the security policies connected to control of information flows on a network. Component interactions of DIS which generate information flows on a network are modelled. The security policy is built on the basis of division of all of information flows on a network on the admissible flows and non admissible flows. Permission is created due to the model of the allowed interactions in case of decision of legal tasks in DIS. This information by means of the special protocol allows controlling information flows.

The permitted interactions are defined by meta data of tasks and their communications. In the paper ways of creation of meta data about permitted interactions of tasks are considered.

Acknowledgements
This work was supported by the Russian Scientific Foundation (grant No. 16-11-10227).

REFERENCES

Grusho, A. and E. Timonina. 2011. "Prohibitions in discrete probabilistic statistical problems". *Discrete Math. and Appl.* 21, No.3, 275–281.

Grusho, A., N. Grusho, S. Shorgin and E. Timonina. 2014. "Secure architecture of the distributed systems". *Systems and means of informatics* 24, No.3, 18-31.

Grusho, A., N. Grusho, S. Shorgin and E. Timonina. 2015. "Possibilities of Secure Architecture Creation for Dynamically Changing Information Systems". *Systems and means of informatics* 25, No.3, 78-93.

Grusho, A., M. Levykin, E. Timonina, V. Piskovski and A. Timonina. 2015. "Architecture of consecutive identification of attack to information resources". *2015 7th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Brno, 265-268.

Lantz, Bob, Brandon Heller and Nick McKeown. "A Network in a Laptop: Rapid Prototyping for Software-Defined Networks". *9th ACM Workshop on Hot Topics in Networks*, October 20-21, 2010, Monterey, CA.

Menezes, Alfred J., Paul C. van. Oorschot and Scott A.. Vanstone. 1997. *Handbook of Applied Cpyptography*, CRC Press LLC, 780 p.

Nilsson, Nils J. 1971. *Problem-Solving Methods in Artificial Intelligence*, New York: McGraw-Hill Pub. Co., 255 p.

Final Evaluation Report Oracle Corporation's Oracle7 and Trusted Oracle7. Report No. CSC-EPL-94/004. C-Evaluation No. 07-95. 5 April 1994.

Rieck, K., P. Stewin and J.-P. Seifert (Eds). 2013. "Detection of Intrusions and Malware, and Vulnerability Assessment". *Proc. of 10th International Conference, DIMVA 2013*, LNCS 7967, Springer Berlin Heidelberg, 207 p.

Samuylov, K.E., A. V. Chukarin and N. V. Yarkina. 2009. *Business processes and information technologies in management of the telecommunication companies*, Moscow: Alpina Pablisherz, 2009. 442 p.

Skorobogatov, S. and Ch. Woods. 2012. "Breakthrough Silicon Scanning Discovers Backdoor in Military Chip". *Cryptographic Hardware and Embedded Systems - CHES 2012* LNCS 7428. Springer, Heidelberg, 23-40.

TCSEC. Department of Defense Trusted Computer System Evaluation Criteria. 1985, DoD.

Finn, V.K. (Eds), 2009. *Automatic Hypotheses Generation in Intelligent Systems*, Moscow: KD "LIBROKOM", 528 p.

AUTHOR BIOGRAPHIES

ALEXANDER A. GRUSHO, Professor (1993), Doctor of Science in physics and mathematics (1990). He is Head of laboratory in Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS) and Professor of Moscow State University.

Research interests: probability theory and mathematical statistics, information security, discrete mathematics, computer sciences.

His email is `grusho@yandex.ru`.

ELENA E. TIMONINA has graduated from the Moscow Institute of Electronics and Mathematics and obtained the Candidate degree (PhD) in physics and mathematics (1974). She is Doctor in Technical Science (2005), Professor (2007). Now she works as leading scientist in Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS).

Research interests: probability theory and mathematical statistics, information security, cryptography, computer sciences.

Her email is `eltimon@yandex.ru`.

SERGEY Ya. SHORGIN, Professor (2003), Doctor of Science in physics and mathematics (1997). He is Deputy Director of Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS) and principal scientist of Institute of Informatics Problems of FRC CSC RAS. Research interests: probability theory and mathematical statistics, information security, communication systems modeling, computer sciences.

His email is `sshorgin@ipiran.ru`.

ON ASYMPTOTIC APPROXIMATIONS TO THE DISTRIBUTIONS OF STATISTICS CONSTRUCTED FROM SAMPLES WITH RANDOM SIZES

Vladimir Bening
Moscow State University,
Moscow, Russia
RUDN University,
Moscow, Russia

Victor Korolev
Moscow State University,
Moscow, Russia
IPI FRC CSC RAS,
Moscow, Russia
Hangzhou Dianzi University,
Hangzhou, China

Alexander Zeifman
Vologda State University,
Vologda, Russia
IPI FRC CSC RAS;
ISEDT RAS
Email: a_zeifman@mail.ru

KEYWORDS

Sample of random size; asymptotic expansions; mixtures of probability laws; Laplace distribution; Student distribution; transfer theorem.

ABSTRACT

Due to the stochastic character of the intensities of information flows in high performance information systems, the size of data available for the statistical analysis can be often regarded as random. In the paper general theorem concerning the asymptotic expansions of the distribution function of the statistics based on the sample of random size was proved. Some examples are presented for the cases where the sample size has the negative binomial or discrete Pareto distributions.

INTRODUCTION

In classical problems of mathematical statistics, the size of the available sample, i. e., the number of available observations, is traditionally assumed to be deterministic. In the asymptotic settings it plays the role of infinitely increasing *known* parameter. At the same time, in practice very often the data to be analyzed is collected or registered during a certain period of time and the flow of informative events each of which brings a next observation forms a random point process. Therefore, the number of available observations is unknown till the end of the process of their registration and also must be treated as a (random) observation. For example, this is so in insurance statistics where during different accounting periods different numbers of insurance events (insurance claims and/or insurance contracts) occur and in high performance information systems where due to the stochastic character of the intensities of information flows, the size of data available for the statistical analysis can be often regarded as random. Say, the statistical algorithms applied in high-frequency financial applications must take into consideration that the number of events in a limit order book during a time unit essentially depends on the intensity of order flows. Moreover, contemporary statistical procedures of insurance and financial mathematics do take this circumstance into consideration as one of possible ways of dealing with heavy tails. However, in other fields such as medical statistics or quality control this approach has not become conventional yet although the number of patients with a certain disease varies from month to month due to seasonal factors or from year to year due to some epidemic reasons and the number of failed items varies from lot to lot. In these cases the number of available observations as well as the observations themselves are unknown beforehand and should be treated as random to avoid underestimation of risks or error probabilities.

In asymptotic settings, statistics constructed from samples with random sizes are special cases of random sequences with random indices. The randomness of indices usually leads to that the limit distributions for the corresponding random sequences are heavy-tailed even in the situations where the distributions of non-randomly indexed random sequences are asymptotically normal see, e. g., [1], [2], [3]. For example, if a statistic which is asymptotically normal in the traditional sense, is constructed on the basis of a sample with random size having negative binomial

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

distribution, then instead of the expected normal law, the Student distribution with power-type decreasing heavy tails appears as an asymptotic law for this statistic.

In the present paper, asymptotic expansions (a.e:s) are obtained for the distribution functions (d.f:s) of statistics constructed from samples with random sizes. These results continue the studies started in [1] – [6]. The obtained a.e:s directly depend on the a.e. for the d.f. of the random sample size and on the a.e. for the d.f. of the statistic based on the sample with a non-random size. Such statements are conventionally called transfer theorems. So, we may say that in this paper transfer theorems are presented for the a.e:s of the d.f:s of statistics constructed from samples with random size. Unlike previous works, here we concentrate our attention on the case of non-normalized statistics.

We use conventional notation: $\mathbb{R}$ is the set of real numbers, $\mathbb{N}$ is the set of natural numbers, $\Phi(x)$ and $\varphi(x)$ are the d.f. and the probability density of the standard normal law, respectively.

In Section 2 the main result is presented and proved, Section 3 contain some examples, Section 4 is devoted to the normalized statistics.

MAIN RESULTS FOR NON-NORMALIZED STATISTICS

Consider random variables (r.v:s) $N_1, N_2, \ldots$ and $X_1, X_2, \ldots$, defined on the same probability space $(\Omega, \mathcal{A}, \mathsf{P})$. The r.v:s $X_1, X_2, \ldots X_n$ will be treated as observations with n being a non-random sample size, whereas the r.v:s N_n will be treated as random sample sizes depending on the parameter $n \in \mathbb{N}$. For example, if the r.v. N_n has the geometric distribution $\mathsf{P}\big(N_n = k\big) = frac1n(1-\frac{1}{n})^{k-1}$, $k \in \mathbb{N}$, then $\mathsf{E}N_n = n$, that is, the r.v. N_n is parameterized by its expectation n.

Assume that for each $n \geq 1$ the r.v. N_n takes only natural values, that is, $N_n \in \mathbb{N}$ and are independent of the sequence $X_1, X_2, \ldots$. Everywhere in what follows consider the r.v:s $X_1, X_2, \ldots$ to be independent and identically distributed (i.i.d) with some common d.f. $F(x)$. By $T_n = T_n(X_1, \ldots, X_n)$ denote a statistic, that is, real measurable function of observations $X_1, \ldots, X_n$. We focus on the situation where the number of available observations is large, that is, $n \to \infty$. Assume that the d.f. of the non-normalized statistic T_n weakly converges to some d.f. $G(x)$, that is,

$$\mathsf{P}\big(T_n < x\big) \longrightarrow G(x), \quad n \to \infty, \qquad (2.1)$$

In every continuity point of $G(x)$. Assume that, as $n \to \infty$, the random sample size N_n tends to infinity in probability, that is, for any $M > 0$

$$\mathsf{P}(N_n \leq M) \longrightarrow 0, \quad n \to \infty. \qquad (2.2)$$

Consider the limit behavior of the d.f. of the statistic constructed from the sample of a random size, that is, of the statistic $T_{N_n}(\omega) \equiv T_{N_n(\omega)}(X_1(\omega), \ldots, X_{N_n(\omega)}(\omega))$, $\omega \in \Omega$. As is shown in the following lemma, under the conditions (2.1) and (2.2) the limit law for T_{N_n} is the same as for T_n.

Lemma 2.1. *Let conditions* (2.1) *and* (2.2) *hold. Then* $\mathsf{P}\big(T_{N_n} < x\big) \longrightarrow G(x)$, $n \to \infty$, *at every continuity point of* $G(x)$.

The **proof** is a simple exercise on the application of the formula of total probability.

Now assume that the d.f. of the *non-normalized* statistic T_n admits an a.e. described by the following condition.

Condition 1. *There exist constants* $l \in \mathbb{N}$, $\alpha > \frac{1}{2}$, $C_1 > 0$, *a differentiable d.f.* $G(x)$ *and differentiable bounded functions* $g_i(x)$, $i = 1, \ldots, l$ *such that*

$$\sup_x \Big|\mathsf{P}\big(T_n < x\big) - G(x) - \sum\nolimits_{i=1}^{l} \frac{g_i(x)}{n^{i/2}}\Big| \leq \frac{C_1}{n^{\alpha}}, \; n \in \mathbb{N}.$$

Also assume that the d.f. of the *normalized* random sample size T_n admits an a.e. described by the following condition.

Condition 2. *There exist constants* $m \in \mathbb{N}$, $\beta > m/2$, $C_2 > 0$, *functions* $0 < v(n) \uparrow \infty$, $(n \to \infty)$, $u(n) \in \mathbb{R}$, *a d.f.* $H(x)$ *with* $H(0+) = 0$ *and functions with bounded variations* $h_j(x)$, $j = 1, \ldots, m$, *such that*

$$\sup_{x \geq 0}\Big|\mathsf{P}\Big(\frac{N_n}{v(n)} - u(n) < x\Big) - H(x) -$$
$$- \sum\nolimits_{j=1}^{m} \frac{h_j(x)}{n^{j/2}}\Big| \leq \frac{C_2}{n^{\beta}}, \quad n \in \mathbb{N}.$$

Everywhere in what follows we denote $y_n = y - u(n)$.

Theorem 2.1. *Let conditions* 1 *and* 2 *hold. Then*

$$\sup_x \Big|\mathsf{P}\big(T_{N_n} < x\big) - G_n(x)\Big| \leq$$
$$\leq C_1 \mathsf{E}N_n^{-\alpha} + 2\frac{C_2}{n^{\beta}} \sup_x \sum\nolimits_{i=1}^{l} |g_i(x)|,$$

where

$$G_n(x) = G(x) + \sum\nolimits_{i=1}^{l} \frac{g_i(x)}{(v(n))^{i/2}} \int_{1/v(n)}^{\infty} y^{-i/2} \times$$
$$\times d\Big(H(y_n) + \sum\nolimits_{j=1}^{m} n^{-j/2} h_j(y_n)\Big) =$$

$$= G(x)+\sum_{i=1}^{l} g_i(x)\int_1^\infty z^{-i/2} d\Big(H(z/v(n)-u(n))+$$

$$+\sum_{j=1}^{m} n^{-j/2}\ h_j(z/v(n)-u(n))\Big).$$

Proof. Estimate the difference under consideration above in the following way:

$$\sup_x\Big|\mathsf{P}\big(T_{N_n}<x\big)-G_n(x)\Big|\le I_{1n}+I_{2n}, \qquad (2.4)$$

where

$$I_{1n}\equiv\sup_x\Big|\mathsf{P}\big(T_{N_n}<x\big)-\mathsf{E}\Big(G(x)+\sum_{i=1}^{l}\frac{g_i(x)}{N_n^{i/2}}\Big)\Big|,$$

$$I_{2n}\equiv\sup_x\Big|\mathsf{E}\Big(G(x)+\sum_{i=1}^{l}\frac{g_i(x)}{N_n^{i/2}}\Big)-G_n(x)\Big|.$$

Estimate I_{1n} using condition 1 and the formula of total probability. We have

$$I_{1n}=\sup_x\Big|\sum_{k=1}^{\infty}\mathsf{P}\big(N_n=k\big)\Big(\mathsf{P}\big(T_k<x\big)-$$

$$-G(x)-\sum_{i=1}^{l}k^{-i/2}g_i(x)\Big)\Big|\le$$

$$\le\sum_{k=1}^{\infty}\mathsf{P}\big(N_n=k\big)\sup_x\Big|\mathsf{P}\big(T_k<x\big)-$$

$$-G(x)-\sum_{i=1}^{l}k^{-i/2}g_i(x)\Big|\le$$

$$\le C_1\sum_{k=1}^{\infty}\frac{1}{k^\alpha}\mathsf{P}\big(N_n=k\big)=C_1\mathsf{E}N_n^{-\alpha}. \qquad (2.5)$$

To estimate I_{2n}, use condition 2 and integration by parts. We have

$$I_{2n}=\sup_x\Big|\mathsf{E}\Big(G(x)+\sum_{i=1}^{l}N_n^{-i/2}g_i(x)\Big)-G(x)-$$

$$-\sum_{i=1}^{l}\big(v(n)\big)^{-i/2}g_i(x)\int_{1/v(n)}^{\infty}y^{-i/2}d\Big(H(y_n)+$$

$$+\sum_{j=1}^{m}\frac{h_j(y_n)}{n^{j/2}}\Big)\Big|=\sup_x\Big|\sum_{i=1}^{l}\mathsf{E}N_n^{-i/2}g_i(x)-$$

$$-\sum_{i=1}^{l}g_i(x)(v(n))^{-i/2}\int_{1/v(n)}^{\infty}\frac{1}{y^{i/2}}\times$$

$$\times d\Big(H(y_n)+\sum_{j=1}^{m}\frac{h_j(y_n)}{n^{j/2}}\Big)\Big|=$$

$$=\sup_x\Big|\sum_{i=1}^{l}g_i(x)\int_1^\infty y^{-i/2}d\mathsf{P}\big(N_n<y\big)-$$

$$-\sum_{i=1}^{l}g_i(x)\big(v(n)\big)^{-i/2}\int_{1/v(n)}^{\infty}y^{-i/2}\times$$

$$\times d\Big(H(y_n)-\sum_{j=1}^{m}\frac{h_j(y_n)}{n^{j/2}}\Big)\Big|.$$

Changing the variables in the first integral we obtain $I_{2n}=$

$$=\sup_x\Big|\sum_{i=1}^{l}\frac{g_i(x)}{(v(n))^{i/2}}\int_{\frac{1}{v(n)}}^{\infty}\frac{1}{y^{i/2}}d\Big(\mathsf{P}\Big(\frac{N_n}{v(n)}<y\Big)-$$

$$-H(y_n)-\sum_{j=1}^{m}\frac{h_j(y_n)}{n^{j/2}}\Big)\Big|.$$

Using integration by parts, the boundedness of the functions $g_i(x)$, $i=1,...,l$ and condition 2, we obtain the inequalities

$$I_{2n}\le\frac{C_2}{n^\beta}\sup_x\sum_{i=1}^{l}\big|g_i(x)\big|+$$

$$+\sup_x\Big|\sum_{i=1}^{l}\frac{g_i(x)}{(v(n))^{i/2}}\int_{\frac{1}{v(n)}}^{\infty}\Big(\mathsf{P}\Big(\frac{N_n}{v(n)}-u(n)<$$

$$<y_n\Big)-H(y_n)-\sum_{j=1}^{m}n^{-j/2}h_j(y_n)\Big)dy^{-i/2}\Big)\Big|\le$$

$$\le2\frac{C_2}{n^\beta}\sup_x\sum_{i=1}^{l}\big|g_i(x)\big|. \qquad (2.6)$$

Now the desired assertion follows from (2.4), (2.5) and (2.6). The theorem is proved.

EXAMPLES

Here we present two examples of the application of theorem 2.1 for statistics constructed from samples with special random sample sizes. We will consider the a.e:s for the d.f. of the sample mean constructed from samples with random sizes. Similar results can be obtained for statistics admitting the Edgeworth-type a.e:s for the d.f. under a non-random sample size.

Let $X_1,X_2,\dots$ be i.i.d. r.v:s with $\mathsf{E}X_1=\mu$, $0<\mathsf{D}X_1=\sigma^2$, $\mathsf{E}|X_1|^{3+2\delta}<\infty$, $\delta\in(0,\frac{1}{2})$ and $\mathsf{E}\big(X_1-\mu\big)^3=\mu_3$. For $n\in\mathbb{N}$ denote

$$T_n=\frac{X_1+\ldots+X_n-n\mu}{\sigma\sqrt{n}}. \qquad (3.1)$$

Also assume that the r.v. X_1 satisfies the CramEr condition (C)

$$\limsup_{|t|\to\infty}|\mathsf{E}\exp\{itX_1\}|<1.$$

Then with the account of theorem 6.3.2 from [7] we obtain

$$\sup_x\Big|\mathsf{P}\big(T_n<x\big)-\Phi(x)-$$

$$-\frac{\mu_3}{6\sqrt{n}\sigma^3}(1-x^2)\varphi(x)\Big|\le\frac{C_1}{n^{1/2+\delta}} \qquad (3.2)$$

with $C_1>0$, $\delta\in(0,\frac{1}{2})$, $n\in\mathbb{N}$. Thus, statistic (3.1) satisfies condition 1 of theorem 2.1 with $\alpha=\frac{1}{2}+\delta$, $l=1$, $G(x)=\Phi(x)$, $g_1(x)=\frac{\mu_3}{6\sigma^3}(1-x^2)\varphi(x)$. It is easy to see that $\sup_x|g_1(x)|<\infty$.

A. Sample size with the negative binomial distribution

Let the random sample size N_n have the negative binomial distribution with parameters $p = \frac{1}{n}$ and $r > 0$, that is

$$\mathsf{P}\big(N_n = k\big) = \frac{(k+r-2)\cdots r}{(k-1)!}\frac{1}{n^r}\Big(1-\frac{1}{n}\Big)^{k-1}, \quad k \in \mathbb{N}. \tag{3.4}$$

With $r = 1$ we obtain the geometric distribution. In [8] (relation (6.112) on p. 233) the following bound was presented for the rate of convergence of the normalized sample size N_n to the gamma-distribution:

$$\sup_{x\geq 0}\Big|\mathsf{P}\Big(\frac{N_n}{\mathsf{E}N_n} < x\Big) - H_r(x)\Big| \leq \begin{cases} \dfrac{C_r}{n}, & r \geq 1, \\ \dfrac{C_r}{n^r}, & r \in (0,1), \end{cases} \tag{3.5}$$

where $C_r > 0$, $n \in \mathbb{N}$ and

$$H_r(x) = \frac{r^r}{\Gamma(r)}\int_0^x e^{-ry}y^{r-1}dy, \quad x \geq 0, \tag{3.6}$$

is the gamma-d.f. with shape parameter $r > 0$ coinciding with scale parameter. In this case

$$\mathsf{E}N_n = r(n-1)+1. \tag{3.7}$$

So, from (3.5) – (3.7) it follows that the random sample size N_n satisfies condition 2 with $v(n) = r(n-1)+1$, $H(x) = H_r(x)$, $m = 1$, $h_1(x) \equiv 0$, $C_2 = C_r > 0$, $u(n) = 0$, $n \in \mathbb{N}$,

$$\beta = \begin{cases} 1, & r \geq 1, \\ r, & r \in (1/2, 1). \end{cases}$$

Further, using the equality $\big(1+x\big)^\gamma = \sum_{k=0}^\infty \gamma(\gamma-1)\cdots(\gamma-k+1)x^k/k!$, $|x| < 1$, $\gamma \in \mathbb{R}$, it is easy to obtain that for $r > 0$, $r \neq 1$, $n \in \mathbb{N}$

$$\mathsf{E}N_n^{-1} = [(n-1)(1-r)]^{-1}(n^{1-r}-1) = O(n^{-r}). \tag{3.8}$$

For $r = 1$ we have $\mathsf{E}N_n^{-1} = (n-1)^{-1}\log n$, $n > 1$. Now using the Hölder inequality, we obtain

$$\mathsf{E}N_n^{-\alpha} \leq \big(\mathsf{E}N_n^{-1}\big)^\alpha (\alpha \leq 1),$$

$$\mathsf{E}N_n^{-\alpha} = O(n^{-r(1/2+\delta)}), \ r > 0, \ r \neq 1,$$

$$\mathsf{E}N_n^{-\alpha} = O\Big(\Big(\frac{\log n}{n}\Big)^{1/2+\delta}\Big), \quad r = 1, \ n \in \mathbb{N}. \tag{3.9}$$

So, using theorem 2.1, relations (3.2), (3.3), (3.5)–(3.9) and the equality

$$\int_{(r(n-1)+1)^{-1}}^\infty \sqrt{y}dH_r(y) = \int_0^\infty \sqrt{y}dH_r(y) + O\big(1/n\big) =$$

$$= \int_0^\infty \sqrt{y}\frac{r^r}{\Gamma(r)}e^{-ry}y^{r-1}dy + O\big(1/n\big) =$$

$$= \frac{r^r}{\Gamma(r)}\int_0^\infty e^{-y}\frac{y^{r-1/2}}{r^{r+1/2}}dy + O\big(1/n\big) =$$

$$= \frac{\Gamma(r+1/2)}{\Gamma(r)\sqrt{r}} + O\big(1/n\big),$$

we obtain the following assertion.

Theorem 3.1. *Let a statistic T_n have the form* (3.1), *where $X_1, X_2, \ldots$ are i.i.d. r.v:s with $\mathsf{E}X_1 = \mu$, $0 < \mathsf{D}X_1 = \sigma^2$, $\mathsf{E}|X_1|^{3+2\delta} < \infty$, $\delta \in (0, \frac{1}{2})$ and $\mathsf{E}(X_1 - \mu)^3 = \mu_3$. Also assume that the r.v. X_1 satisfies the Cramér condition (C). Assume that the r.v. N_n has the negative binomial distribution* (3.4) *with some $r > 0$. Then for $r > (1+2\delta)^{-1}$, as $n \to \infty$, the d.f. of T_{N_n} admits the a.e.*

$$\sup_x \Big|\mathsf{P}\big(T_{N_n} < x\big) - \Phi(x) -$$

$$- \frac{\mu_3\Gamma(r+1/2)}{6\sigma^3\Gamma(r)\sqrt{r^2(n-1)+r}}(1-x^2)\varphi(x)\Big| =$$

$$= \begin{cases} O\big((n^{-1}\log n)\big)^{1/2+\delta}, & r = 1, \\ O\big(n^{-\min\{1, r(1/2+\delta)\}}\big), & r > 1, \\ O\big(n^{-r(1/2+\delta)}\big), & (1+2\delta)^{-1} < r < 1. \end{cases}$$

B. Sample size with the discrete Pareto distribution

In [4], an example related to the theory of records was considered of a sequence of r.v:s $N_n(s)$ depending on a natural parameter $s \in \mathbb{N}$ such that

$$\mathsf{P}\big(N(s) \geq k\big) = \frac{s}{s+k-1}, \quad k \geq 1 \tag{3.10}$$

(also see [9], [10]). Let now $N^{(1)}(s), N^{(2)}(s), \ldots$ be i.i.d. r.v:s with distribution (3.10). Define the r.v:s $N_n(s) = \max_{1\leq j\leq n} N^{(j)}(s)$. Then, as was shown in [4],

$$\lim_{n\to\infty}\mathsf{P}\Big(\frac{N_n(s)}{n} < x\Big) = e^{-s/x}, \quad x > 0. \tag{3.11}$$

In [5] the following bound of the rate of convergence in (3.10) was obtained: there exists a constant $C_s \in (0, \infty)$ such that

$$\sup_{x\geq 0}\Big|\mathsf{P}\Big(\frac{N_n(s)}{n} < x\Big) - e^{-s/x}\Big| \leq \frac{C_s}{n}, \qquad n \in \mathbb{N}. \tag{3.12}$$

So, from (3.12) it follows that the r.v. $N_n(s)$ satisfies condition 2 of theorem 2.1 with

$$v(n) = n, \quad H(x) = e^{-s/x}, \quad m = 1, \quad h_1(x) \equiv 0,$$

$$C_2 = C_s > 0, \quad u(n) = 0, \quad \beta \ = \ 1. \tag{3.13}$$

Consider $\mathsf{E}N_n^{-1}(s)$ in more detail. From the definition of $N_n(s)$ and (3.10) we have

$$\mathsf{P}\big(N_n(s)=k\big)=\Big(\frac{k}{s+k}\Big)^n-\Big(\frac{k-1}{s+k-1}\Big)^n=$$
$$=sn\int_{k-1}^{k}\frac{x^{n-1}}{(s+x)^{n+1}}dx.$$

Therefore,

$$\mathsf{E}N_n^{-1}(s)=\sum\nolimits_{k=1}^{\infty}\frac{1}{k}\mathsf{P}\big(N_n(s)=k\big)=$$
$$=sn\sum\nolimits_{k=1}^{\infty}\frac{1}{k}\int_{k-1}^{k}\frac{x^{n-1}}{(s+x)^{n+1}}dx\le \qquad (3.14)$$
$$\le sn\sum\nolimits_{k=1}^{\infty}\int_{k-1}^{k}\frac{x^{n-2}}{(s+x)^{n+1}}dx=$$
$$=sn\int_{0}^{\infty}\frac{x^{n-2}}{(s+x)^{n+1}}dx=\frac{1}{s(n-1)}=O(n^{-1}),$$

see [11], formula 856.12. Now using the Hölder inequality we obtain

$$\mathsf{E}N_n^{-\alpha}\le\big(\mathsf{E}N_n^{-1}\big)^{\alpha},\ \alpha\le 1,$$
$$\mathsf{E}N_n^{-\alpha}=O(n^{-(1/2+\delta)}),\ \alpha=1/2+\delta. \qquad (3.15)$$

Now from theorem 2.1, relations (3.12)–(3.15) and the equality

$$\int_{n^{-1}}^{\infty}\sqrt{y}de^{-s/y}=\sqrt{s}\int_{0}^{\infty}y^{-1/2}e^{-y}dy+O(n^{-1})=$$
$$=\sqrt{s}\Gamma(1/2)+O(n^{-1})=\sqrt{\pi s}+O(n^{-1})$$

we directly obtain the following theorem.

Theorem 3.2. *Let a statistic T_n have the form* (3.1), *where $X_1, X_2, ...$ are i.i.d. r.v:s with $\mathsf{E}X_1=\mu$, $0<\mathsf{D}X_1=\sigma^2$, $\mathsf{E}|X_1|^{3+2\delta}<\infty$, $\delta\in(0,\frac{1}{2})$ and $\mathsf{E}(X_1-\mu)^3=\mu_3$. Also assume that the r.v. X_1 satisfies the Cramér condition (C). Assume that the r.v. N_n has the discrete Pareto distribution* (3.10). *Then, as $n\to\infty$, the d.f. of $T_{N_n(s)}$ admits the a.e.*

$$\sup_x\Big|\mathsf{P}\big(T_{N_n(s)}<x\big)-\Phi(x)-$$
$$-\frac{\mu_3\sqrt{\pi s}}{6\sigma^3\sqrt{n}}(1-x^2)\varphi(x)\Big|=O\Big(\frac{1}{n^{1/2+\delta}}\Big).$$

CONCLUSION

Due to the stochastic character of the intensities of information flows in high performance information systems, the size of data available for the statistical analysis can be often regarded as random. In the paper general theorem concerning the asymptotic expansions of the distribution function of the statistics based on the sample of random size was proved. Some examples are presented for the cases where the sample size has the negative binomial or discrete Pareto distributions.

ACKNOWLEDGEMENT

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008) and by Russian Foundation for Basic Research, project 15-07-02652.

REFERENCES

[1] *Gnedenko B. V., Korolev V. Yu.* Random Summation. Limit Theorems and Applications. – Boca Raton: CRC Press, 1996.

[2] *Bening V. E., Korolev V. Yu.* Generalized Poisson Models and Their Applications in Insurance and Finance. – Utrecht: VSP, 2002.

[3] *Bening V. E., Korolev V. Yu.* On an application of the Student distribution in the theory of probability and mathematical statistics // Theory of Probability and Its Applications, 2005. Vol. 49. No. 3. P. 377–391.

[4] *Bening V. E., Korolev V. Yu.* Some statistical problems related to the Laplace distribution // Informatics and Its Applications, 2008. Vol. 2. No. 2. P. 19 – 34.

[5] *Lyamin O. O.* On the rate of convergence of the distributions of some statistics to the Laplace and Student distributions // Bulletin of Moscow University. Ser. 15 Computational Mathematics and Cybernetics, 2011. No. 1. P. 39 – 47.

[6] *Bening V. E., Galieva N. K, Korolev V. Yu.* Asymptotic expansions for the distributiond functions of statistics constructed from samples with random sizes // Informatics and Its Applications, 2013. Vol. 7. No. 2. P. 75 – 83.

[7] *Petrov V. V.* Sums of Independent Random Variables. – Berlin–New York: Springer, 1975.

[8] *Bening V. E., Korolev V. Yu., Sokolov I. A., Shorgin S. Ya.* Randomized Models and Methods Of the Theory of Reliability of Information and Technical Systems. – Moscow: Torus Press, 2007 (in Russian).

[9] *Wilks S. S.* Recurrence of extreme observations // Journal of American Mathematical Society, 1959. V. 1, No. 1, P. 106 – 112.

[10] *Nevzorov V. B.* Records: Mathematical Theory. – Providence, RI: American Mathematical Society, 2000. [15]

[11] *Dwight H. B.* Tables of Integrals and Other Mathematical Data. – New York: The MacMillan Company, 1961.

AUTHOR BIOGRAPHIES

VLADIMIR BENING is Doctor of Science in physics and mathematics, professor, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; senior researcher, Peoples' Friendship University of Russia (RUDN University), Moscow. His email is `bening@yandex.ru`.

VICTOR KOROLEV is Doctor of Science in physics and mathematics, professor, Head of Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University; leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; professor, Hangzhou Dianzi University, Hangzhou, China. His email is victoryukorolev@yandex.ru.

ALEXANDER ZEIFMAN is Doctor of Science in physics and mathematics; professor, Heard of Department of Applied Mathematics, Vologda State University; senior scientist, Institute of Informatics Problems, Federal Research Center ”Computer Science and Control” of the Russian Academy of Sciences; principal scientist, Institute of Socio-Economic Development of Territories, Russian Academy of Sciences. His email is a_zeifman@mail.ru.

USING INTER-ARRIVAL TIMES FOR SCHEDULING IN NON-OBSERVABLE QUEUES

Mikhail Konovalov
Institute of Informatics Problems
of the FRC CSC RAS, Moscow, Russia,
Email: mkonovalov@ipiran.ru

Rostislav Razumchik
Institute of Informatics Problems
of the FRC CSC RAS, Moscow, Russia,
Peoples' Friendship University of Russia
(RUDN University), Moscow, Russia
Email: rrazumchik@ipiran.ru,
razumchik_rv@pfur.ru

KEYWORDS

Queueing system, customer assignment, partial information, deterministic policies, dispatching, simulation

ABSTRACT

In online dispatching systems, when there is no opportunity to observe the state of the systems' components, one may implement "blind" scheduling policies i.e. those which use incomplete/indirect observations of the system state or do not use any information at all. Here we deal with the well-known problem of scheduling in several non-observable parallel single server queues with single Poisson incoming flow, when the broker (scheduler) does not observe neither the current states of the queues and servers, nor the size of the incoming jobs. The only available information is the job size distribution, server's speeds and job's inter-arrival time distribution. For this problem setting it is known that if the scheduler can memorize the sequence of its previous decisions, then a deterministic policy is much better than the probabilistic policy (with respect to the job's mean waiting and mean sojourn time). But if the broker can memorize its decision, it is also very natural to assume that it can also memorize the time instants at which these decisions are made. In this paper we address the following question: can the deterministic policy be improved if the broker, in addition to decisions made, utilizes also the information about the lengths of time between the decisions? We give numerical evidence that it is indeed possible; the cases presented include three, five and nine parallel $\cdot|M|1$ queues. We describe the simple new policy according to which, the broker's decisions are based on the estimates of the queue sizes. In most of the numerical experiments, the new policy outperformed the deterministic policy The relative gain may reach 10% in the case of the job's mean sojourn time and 50% in the case of the job's mean waiting time.

INTRODUCTION

In this paper we consider the following well-known problem of scheduling in parallel non-observable queues. The system consists of N single server infinite capacity queues, operating in parallel. Each server is labelled with a number from 1 to N without repetitions. The service discipline in each server is FCFS. A single flow of jobs arrives at the system. There is a broker (or scheduler), which has to route (immediately[1]) an incoming job to one of the servers. We adopt the following assumptions:

- the scheduler has only static information about the system: CDF of job's inter-arrival times, CDF of job's size and servers' speeds;
- the scheduler can memorize its routing decisions and time instants at which the decisions were made.

Under assumptions made we seek "blind" scheduling policies, which minimize the following cost functions: job's mean waiting and sojourn times in the system. The peculiar feature of the problem is that the online information about the system's state (like queue sizes, remaining work in the servers etc.) is not available to the broker.

There is an extensive literature on this topic. For general and some recent results one can refer to Hordijk, Anneke and Jeroen (1998); Hordijk, Koole and Loeve (1994); Anselmi and Gaujal (2011, 2010); Anselmi, Gaujal and Nesti (2015); Brun (2016); Altman, Gaujal and Hordijk (2000); Gaujal, Hyon and Jean-Marie (2006); Hordijk and van der Laan (2004); Combe and Boxma (1994). Within the considered framework usually two policies attract the most attention: probabilistic and deterministic policies.

According to the probabilistic policy the next job is routed to server m, $1 \leq m \leq N$ with the probability p_m (independently of the previous decisions). For finding p_m there exist efficient numerical procedures (see, for example, Combe and Boxma (1994); Bell and Stidham (1983); Neely and Modiano (2005); Sethuraman and Squillante (1999)).

Deterministic policies, which can achieve significantly better performance than probabilistic policies[2], can be seen as an infinite sequences of numbers $\{a_1, a_2, \ldots, a_{n-1}, a_n, \ldots\}$, where a_m means that m-th job is routed to server a_m. The interest in such policies is justified by the intuitive idea that under such policies the incoming flow to each server is becoming more regular

[1] I.e. the broker does not have a queue to store the jobs.

[2] As, for example, demonstrated in Anselmi and Gaujal (2010).

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

(less random). And according to the folk theorem from queueing theory (see Humblet (1982); Hajek (1983)) determinism in the inter-arrival times minimizes the waiting time in the single server queue. For N queues in parallel, finding the optimal deterministic sequence is a difficult problem. Sometimes a so called billiard sequence can be used instead, which can be constructed using several known greedy algorithms (see Hordijk and van der Laan (2004)). Nevertheless, one still has to estimate the parameters of the algorithms, for which no general solutions to the best of our knowledge exist.

In this paper we propose the algorithm to find policies, which are better than deterministic policies (with respect to the two cost functions considered above) and are still size-oblivious. The key observation for the new algorithm is the following. In order to implement any deterministic policy the broker has to be able to memorize its previous decisions. It is very natural to assume that in addition to memorizing the decisions made, the broker can also memorize the time instants at which the decisions were made. From the technical point of view this cannot be seen as a problem because the broker may be granted access to the internal clock of the system it is implemented in. Thus each decision, say y_n, where to route the n-th customer is associated with the time instant t_n at which it is made (clearly t_n is the customer's arrival epoch). Now when the next job, say $(n+1)$, arrives to the broker it can base its routing decision on the sequence $(y_1, t_1), \dots, (y_n, t_n)$. Yet it does not observe the system's state and even the quantity, which has to be optimized (job's mean waiting or mean sojourn time).

The paper is organized as follows. In the next section we give the description of the exact model considered in this paper. The section is followed by the description of the proposed algorithm for construction the dispatching policy, based on a sequence $(y_1, t_1), \dots, (y_n, t_n)$. Then we present the numerical results provided by the algorithm, and compare them with the results, achieved by known deterministic policies.

MODEL DESCRIPTION

The system consists of N single server infinite capacity queues, operating in parallel. The speed of server m is denoted by v^m, $1 \leq m \leq N$. The service discipline of each server is FCFS, pre-emption is not allowed. Jobs arrive at the system one-by-one according to a Poisson flow with rate λ. The sizes of the jobs are i.i.d., having exponential distribution with mean 1. Upon arrival each job must be instantly routed to one of the servers. No jockeying between servers and queues is allowed.

DESCRIPTION OF THE ALGORITHM

The key to obtaining the feasible algorithm is the discrete-time setting as in De Vuyst, Bruneel and Fiems (2014). All time-related quantities in the model are discretized into fixed-length intervals of length Δ. Some comments on the choice of Δ will be given in the numerical section and for now we assume that the value of Δ is fixed.

As the job size S has exponential distribution with mean equal to 1, then the service time $S^{(m)}$ in server m has also exponential distribution with rate $v^{(m)}$. We used the following relation to discretize the continuous distribution of $S^{(m)}$:

$$F_0^{(m)} = \mathbf{P}\left\{S^{(m)} = 0\right\} = \mathbf{P}\left\{S < 0,5\Delta v^{(m)}\right\} = 1 - e^{0,5\Delta v^{(m)}},$$

$$F_j^{(m)} = \mathbf{P}\left\{S^{(m)} = j\right\} =$$

$$= \mathbf{P}\left\{S < (m+0,5)\Delta v^{(m)}\right\} - \mathbf{P}\left\{S < (m-0,5)\Delta v^{(m)}\right\} =$$

$$= e^{(m-0,5)\Delta v^{(m)}} - e^{(m+0,5)\Delta v^{(m)}}, \ 1 \leq m \leq J_m,$$

where $J_m = \min\{J : \sum_{j=0}^{J} F_j^{(m)} > 0,95\}$. The rest of the mass is concentrated in the point $J_m + 1$ and eventually the service times of jobs in server m are i.i.d. with the probability mass function $\{F_j^{(m)}\}$.

Let $0 \leq t_1 < t_2 < \dots$ denote the arrival instants of consecutive jobs. Let $y_1, y_2, \dots$ denote the server numbers whereto the jobs are routed i.e. y_n is the server whereto the job arrived at instant t_n was routed. Clearly each y_n takes the value in the set $\{1, 2, \dots, N\}$.

Denote by $u_n^{(m)}$ the workload in the server m at instant $(t_n - 0)$ i.e. just before the job, arrived at instant t_n, was routed to server y_n. By $\tilde{u}_n^{(m)}$ denote the workload in the server m but after the job arrived at instant t_n was routed to server y_n. According to the Lindley equation in queueing theory we have for $n \geq 1$:

$$\tilde{u}_n^{(m)} = u_n^{(m)} + \frac{S}{v^{(m)}}\delta_{m,y_n},$$

$$u_{n+1}^{(m)} = \left(\tilde{u}_n^{(m)} - (t_{n+1} - t_n)\right)^+,$$

where $(\cdot)^+$ is the shorthand notation for $\max(0, \cdot)$ and δ_{ij} is the Kronecker symbol. In this paper we do not pursue any generalizations and assume that initially the system is empty and thus upon arrival of the first job all the servers are empty i.e.

$$u_1^{(1)} = u_1^{(2)} = \dots = u_1^{(N)} = 0. \quad (1)$$

Due to the discrete-time setting the continuous distributions of the workloads $u_n^{(m)}$ and $\tilde{u}_n^{(m)}$ have to be replaced by their discretized versions, which are denoted further by $\{G_{n,k}^{(m)}\}$ and $\{\tilde{G}_{n,k}^{(m)}\}$ i.e.

$$G_{n,k}^{(m)} = \mathbf{P}\left\{u_n^{(m)} = k\right\}, \ k = 0, 1, \dots.$$

$$\tilde{G}_{n,k}^{(m)} = \mathbf{P}\left\{\tilde{u}_n^{(m)} = k\right\}, \ k = 0, 1, \dots.$$

From (1) it follows that the distribution $\{G_{1,k}^{(m)}\}$ is concentrated in one point i.e.

$$G_{1,0}^{(m)} = 1, 1 \leq m \leq N,$$

$$G_{1,k}^{(m)} = 0, \ k \geq 1, 1 \leq m \leq N.$$

For any given sequence $y_1, y_2, \ldots$, one is able to calculate recursively the probability mass functions $\{G^{(m)}_{n,k}\}$ and $\{\tilde{G}^{(m)}_{n,k}\}$ and the expected workloads $U^{(m)}_n = \mathbf{E}u^{(m)}_n$ using, for example, the approach[3] from De Vuyst, Bruneel and Fiems (2014) (see expressions (7) and (11)). In order to reduce the computational complexity, we had to introduce several refinements into the algorithm. Firstly we have truncated the distributions at a given level (in those numerical results, which are presented here, the distributions were truncated at level below which 95% of the mass is concentrated). Secondly the summations of zero terms were avoided. Below we present relations for the recursive computation of the distributions $\{G^{(m)}_{n,k}\}$ and $\{\tilde{G}^{(m)}_{n,k}\}$. The details are omitted and will appear elsewhere.

Define the following sequence of constants

$$\begin{aligned} K^{(m)}_n &= \sup\left\{k : G^{(m)}_{n,k} > 0\right\}, \\ \tilde{K}^{(m)}_n &= \sup\left\{k : \tilde{G}^{(m)}_{n,k} > 0\right\}, \\ \tilde{K}^{(m)*}_n &= \sup\left\{K : \sum_{k=0}^{K} \tilde{G}^{(m)}_{n,k} \geq 0,95\right\}, \\ L^{(m)} &= \sup\left\{j : F^{(m)}_j > 0\right\}. \end{aligned}$$

For $\{\tilde{G}^{(m)}_{n,k}\}$ and $\{G^{(m)}_{n+1,k}\}$ we have for $n \geq 1$:

$$\tilde{G}^{(m)}_{n,k} = \begin{cases} G^{(m)}_{n,k}, & m \neq y_n, \\ \sum\limits_{i=\max(0,k-L^{(m)})}^{\min(k,K^{(m)}_n)} G^{(m)}_{n,i} F^{(m)}_{k-i}, & 0 \leq k \leq \tilde{K}^{(m)*}_n, \\ & m = y_n, \end{cases} \quad (2)$$

$$G^{(m)}_{n+1,0} = \sum_{i=0}^{\min(T_n,\tilde{K}^{(m)}_n)} \tilde{G}^{(m)}_{n,i}, \quad (3)$$

$$G^{(m)}_{n+1,k} = \tilde{G}^{(m)}_{n,k+T_n}, \ 1 \leq k \leq \tilde{K}^{(m)}_n - T_n, \quad (4)$$

where T_n is integer number nearest to $(t_{n+1} - t_n)/\Delta$.

In order to define the rule for producing the sequence $y_1, y_2, \ldots$ we have to come back to the system description. Remind that there is no opportunity to observe neither the queues' sizes nor the remaining work in any server. Let the broker decide where to route the incoming job based on the estimates of the expected workloads $U^{(1)}_n, \ldots, U^{(N)}_n$ in each server (including its queue) upon n-th job arrival. The decision rule is the following: send the first job to the fastest server; send the n-th job to the server y_n, where

$$y_n = \text{argmin}_m \left\{U^{(m)}_n + h\frac{1}{v^{(m)}},\ 1 \leq m \leq N\right\},\ n \geq 2. \quad (5)$$

Here $0 \leq h \leq 1$ is a constant. Due to (1), the expected workload $U^{(m)}_1 = 0$ for each m. As the values of the expected workloads $U^{(m)}_n$, $n \geq 2$ can be calculated from (2), (3) and (4) as

$$U^{(m)}_n = \sum_{k=0}^{\tilde{K}^{(m)*}_n} k\, G^{(m)}_{n,k}, \quad (6)$$

we can write out the algorithm for choosing the next server for an arriving job (see Algorithm 1).

Algorithm 1 Algorithm for choosing the server for an arriving job based on inter-arrival times and estimations of the expected workloads

```
Route the job arrived at instant $t_1$ to the fastest server;
for $i \geq 2$ do
    for job arrived at instant $t_i$ do
        calculate (2) for $n = i - 1$;
        calculate (3), (4) for $n = i - 1$;
        calculate $y_i$ using (5), (6);
        route the job to server $y_i$;
    end for
end for
```

The constant $0 \leq h \leq 1$ which appears in (5), in general, depends on the system's parameters and the cost function. As our experiments show, the value of h allows optimization in the interval [0, 1]. Although in case of minimization of the job's mean waiting time (mean sojourn time) the value of h in (5) has to be equal to zero (one), using the intermediate values sometimes led to better results (lower values of the cost function).

NUMERICAL EXPERIMENT

The purpose of this numerical section is to give the raw comparison of the two scheduling policies – the near optimal deterministic policy and the new policy, proposed in this paper – with respect to job's mean waiting and mean sojourn times.

Some results of such comparison in the system with two servers have been reported in Konovalov and Razumchik (2016). Here we will present the results for the three special cases. The first case is the system with 3 servers with speeds 1, 3 and 7 (see Tables 1 and 2), the second case is the system with 5 servers with speeds 1, 2, 3, 4 and (see Tables 3 and 4) and the third case is the system with 9 servers with speeds 0,9, 1, 1,1, 2,9, 3, 3,1, 6,9, 7 and 7,1 (see Table 5).

In all cases it is assumed that the incoming flow of jobs is Poisson and the job size distribution is exponential with mean equal to 1. Thus the service time distributions are in each case exponential with the rates equal to servers' speeds.

As the near optimal deterministic policy we used the billiard sequence, generated by the SG (Special Greedy) algorithm from (Hordijk and van der Laan, 2004, p.184). In order to construct the dispatching sequence this algorithm takes as input the probabilities p_m of sending an arriving job to server m, $1 \leq m \leq N$ and non-negative rational numbers x_m, $1 \leq m \leq N$. We assumed (as in Anselmi

[3] Another approach and the expression for the expected workload can be found in (Abate and Whitt, 1994, Eq.(18)). Yet in the experiments reported we did not utilize it. Other methods for the calculation of the time-dependent workload can be also found in Stadje (1997); Ackroyd (1986).

and Gaujal (2011)) that $x_m = 1$ if m is the fastest server i.e. if $v_m = \max_j v_j$ and $x_m = 0$ otherwise. In the experiments, results of which we present here, we did not have serves with equal speeds and thus the fastest server had always been unique. The values of p_m were taken as the solution of PA1 problem in Combe and Boxma (1994), when the objective was to minimize the mean waiting time[4]. When the mean sojourn time was to be minimized p_m were taken as the solution of the modified version of the PA1 problem[5].

Sometimes we also report the results for two size-based policies: join-the-shortest-queue (JSQ) and myopic. According to the JSQ policy the arriving job is routed to the server with minimum number of jobs; with the myopic policy the broker chooses such server, which minimizes the sojourn time for the arriving job. We used our own implementation of the JSQ and myopic policies, using the simulation framework Konovalov and Razumchik (2014); Konovalov (2014). The ties in the JSQ policy were broken according to the rule: if two or more servers have the same number of jobs (in the queue and in the server), choose the fastest server. It is worth mentioning that the performance of the JSQ policy heavily depends on how one breaks the ties. If the ties are broken, for example, randomly, the results for the JSQ policy will be much poorer than those presented.

From the tables presented below one can see that the new policy usually outperforms the deterministic policy. The advantage of the new policy is mostly noticeable with respect to the mean waiting time: the relative gain from the new policy may reach 50 %.

Let us consider the results for the 3-server case in more detail (see Tables 1 and 2). In the the 4-th and 5-th rows of the Table 1 one can see the values (rounded to the third decimal place) of the mean waiting times obtained by methods from Combe and Boxma (1994) and Hordijk, Koole and Loeve (1994) and reported in Hordijk, Koole and Loeve (1994). The sign x means that values for such values of the system's load ρ have not been reported in Hordijk, Koole and Loeve (1994). These two methods produce periodic policies. As one can notice these policies are better than the policy, produced by the SG algorithm and worse than the new policy. As the new policy is not periodic, we cannot compare it with respect to the length of the period as done in Hordijk, Koole and Loeve (1994).

From the Table 2 it is worth noticing that policies, which do not use any online information, may outperform size-based policies. Under low system's load, for example $\rho = 0,375$, the relative gain of the deterministic policy (constructed by the SG algorithm) with respect to JSQ policy is $\frac{0{,}267-0{,}247}{0{,}258} \times 100\% \approx 3{,}3$ % and relative gain of the new policy with respect to JSQ is even higher and equal to $\frac{0{,}267-0{,}247}{0{,}267} \times 100$ % $\approx 7{,}5$ %. According to our numerical experiments this relative gain monotonically increases with the decrease of the system's load.

The other observation about the new policy is that it is hard to obtain any improvement over the SG algorithm, when the system's load is high (see the last rows of the Tables 1, 2, 3 and 4). Our guess here is that under high load all queues are non-empty most of the time and additional information about the inter-arrival times does not pay any role.

The size of the discretization step size requires more careful treatment and justification. We have obtained the results for the new policy using the discretization step size equal to 0,1. In those experiments with the Algorithm 1, which we report here, making step size smaller did not lead to any improvements.

In the 5-server case (see Tables 3 and 4) the tendency is similar: the new policy is always better than the deterministic one. With respect to the mean waiting time the relative gain is more pronounced.

As the number of servers increases, construction of the dispatching sequence using the new algorithm becomes more challenging: its running time increases and thus it requires high performance computing nodes. It also becomes too sensitive to the value of Δ. Finding good values of p_i for the deterministic policy is another challenge. For the demonstration purposes we have considered the case with 9 servers. The results are presented in the Table 5. Here as well the new policy performs better than the deterministic one. Although for this case we have used the same algorithm as for the two previous cases, our experiments show that the approaches to solve the problems with all distinct servers and with several group of similar servers should not be the same.

SUMMARY

The main result of this paper is the evidence that there is a general opportunity to increase the performance of the non-observable system (of parallel single server queues) by constructing policies, that utilize the decision history more efficiently.

In most of the presented cases the policy proposed in the paper, which utilizes the history of the inter-arrival times, outperforms the near-optimal deterministic policy. The relative gain may reach ≈ 10% with respect to job's mean sojourn time and ≈ 50% with respect to job's mean waiting time.

The improvements that we have demonstrated here of course come at price. The deterministic policy (using SG algorithm) can be implemented in the broker at very limited costs, whereas the new policy requires quite some computational efforts. But whenever there is an opportunity to implement the deterministic policy in the broker, there is also an opportunity to implement the new policy and thus the decision, which one to prefer, is a matter of compromise between costs and benefits.

[4] These values of p_m may not be the optimal ones for the deterministic policy.

[5] The expression for the mean waiting time in queue m was replaced by the expression for the mean sojourn time in queue m.

Table 1: Mean waiting times in the three server system ($N = 3$) under different dispatching policies. Service time is exponential. The service rates are 1, 4 and 7. The arrival flow is Poisson with rate λ.

$\rho = \lambda/12$	SG	MEM			JSQ	Myopic	(SG-MEM)/SG
0,25	0,039	0,024	0,032	0,031	0,007	0,017	38,5 %
0,375	0,073	0,057	x	x	0,026	0,034	21,9 %
0,5	0,128	0,112	0,120	0,119	0,066	0,061	12,5 %
0,625	0,222	0,207	x	x	0,141	0,108	6,7 %
0,750	0,420	0,405	0,412	0,414	0,286	0,208	3,6 %
0,917	1,650	1,650	x	x	1,041	0,852	0%

Table 2: Mean sojourn times in the three server system ($N = 3$) under different dispatching policies. Service time is exponential. The service rates are 1, 4 and 7. The arrival flow is Poisson with rate λ.

$\rho = \lambda/12$	SG	MEM	JSQ	Myopic	(SG-MEM)/SG
0,25	0,220	0,204	0,222	0,175	7,3 %
0,375	0,258	0,247	0,267	0,200	4,3 %
0,5	0,323	0,314	0,321	0,235	2,0 %
0,625	0,440	0,422	0,401	0,294	4,0 %
0,750	0,649	0,632	0,546	0,410	2,7 %
0,917	1,894	1,894	1,296	1,082	0 %

Table 3: Mean waiting times in the system with 5 servers ($N = 5$) under SG and MEM dispatching policies. Service time is exponential. The service rates are 1, 2, 3, 4, 5. The arrival flow is Poisson with rate λ.

$\rho = \lambda/15$	SG	MEM	(SG-MEM)/SG
0,20	0,023	0,011	52,2 %
0,30	0,046	0,032	30,4 %
0,40	0,081	0,067	17,3 %
0,50	0,135	0,122	9,6 %
0,60	0,221	0,210	5,0 %
0,73	0,448	0,437	2,4 %
0,87	1,154	1,145	0,9 %

Table 4: Mean sojourn times in the system with 5 servers ($N = 5$) under SG and MEM dispatching policies. Service time is exponential. The service rates are 1, 2, 3, 4, 5. The arrival flow is Poisson with rate λ.

$\rho = \lambda/15$	SG	MEM	(SG-MEM)/SG
0,20	0,279	0,263	5,7 %
0,30	0,317	0,301	5,0 %
0,40	0,364	0,351	3,6 %
0,50	0,432	0,417	3,5 %
0,60	0,530	0,515	2,8 %
0,73	0,766	0,754	1,6 %
0,87	1,480	1,469	0,7 %

Table 5: Mean waiting and sojourn times in the system with 9 servers ($N = 9$) under SG and MEM dispatching policies. Service time is exponential. The service rates are 0,9, 1, 1,1, 2,9, 3, 3,1, 6,9, 7 and 7,1. The arrival flow is Poisson with rate λ.

	$\rho = \lambda/33$	SG	MEM	(SG-MEM)/SG
mean waiting times	0,25	0,018	0,010	44,4 %
	0,375	0,041	0,032	21,9 %
mean sojourn times	0,25	0,187	0,180	3,7 %
	0,375	0,231	0,222	3,9 %

Acknowledgements This work was supported by the Russian Foundation for Basic Research (grant 15-07-03406).

REFERENCES

Ackroyd, M. 1986. Approximate characterisation of nonstationary discrete time G/G/1 systems. Performance Evaluation. Vol. 6. No. 2. Pp. 117–123.

Abate, J., Whitt W. 1994. Transient Behavior of the M/G/1 Workload Process. Operations Research. Vol. 42. No. 4. Pp. 750–764.

Anselmi, J., Gaujal B. 2010. The price of anarchy in parallel queues revisited. ACM Sigmetrics. Pp. 353–354.

Anselmi, J., Gaujal B. 2011. The price of forgetting in parallel and non-observable queues. Performance Evaluation. Vol. 68. Issue 12. Pp. 1291–1311.

Altman, E., Gaujal B., Hordijk A. 2000. Balanced Sequences and Optimal Routing. Journal of American Computing Machinery. Vol. 47. Issue 4. Pp. 752–775.

Anselmi, J., Gaujal B., Nesti T. 2015. Control of parallel non-observable queues: asymptotic equivalence and optimality of periodic policies. Stochastic Systems. Vol. 5. Issue 1. Pp. 120–145.

Bell, C. H., Stidham S. 1983. Individual versus social optimization in the allocation of customers to alternative servers. Management Science. Vol. 29. No. 7. Pp. 831–839.

Brun, O. 2016. Performance of non-cooperative routing over parallel non-observable queues. Probability in the Engineering and Informational Sciences. Vol. 30. Issue 3. Pp. 455–469.

Combe, M. B., Boxma O. J. 1994. Optimization of static traffic allocation policies. Theor. Comput. Sci. Vol. 125. No. 1. Pp. 17–43.

Gaujal, B., Hyon E., Jean-Marie A. 2006. Optimal routing in two parallel queues with exponential service times. Discrete Event Dynamic Systems. Vol. 16. Issue 1. Pp. 71–107.

Hajek, B. 1983. The proof of a folk theorem on queuing delay with applications to routing in networks. Journal of the ACM. Vol. 30. No. 4. Pp. 834–851.

Humblet, P. 1982. Determinism minimizes waiting time in queues. The Laboratory for Information and Decision Systems Technical Report ser. LIDS-P/1207.

Hordijk, A., Anneke L., Jeroen T. 1998. Analysis of a finite-source customer assignment model with no state information. Mathematical Methods of Operations Research. Vol. 47. Issue 2. Pp. 317–336.

Hordijk, A., Koole G. M. and Loeve J. A. 1994. Analysis of a customer assignment model with no state information. Probability in the Engineering and Informational Sciences. Vol. 8. Pp. 419–429..

Hordijk, A., van der Laan D. A. 2004. Periodic routing to parallel queues and billiard sequences. Math. Method. Oper. Res., 2004. Vol. 59. No. 2. Pp. 173–192.

Konovalov, M., Razumchik R. 2014. Simulation Of Task Distribution In Parallel Processing Systems. Proceedings of the 6th International Congress on Ultra Modern Telecommunications and Control Systems. Pp. 657–663.

Konovalov, M. G. 2014. Building a simulation model for solving scheduling problems of computing resources. Systems and Means of Informatics. Vol. 24. No. 4. Pp. 45–62. (in Russian)

Konovalov, M., Razumchik R. 2016. Dispatching to two parallel nonobservable queues using only static information. Informatics and its applications. Vol. 10. No. 4. Pp. 57–67.

Neely, M. J., Modiano E. 2005. Convexity in queues with general inputs. IEEE Transactions on Information Theory. Vol. 51. No. 2. Pp. 706–714.

Sethuraman, J., Squillante M. S. 1999. Optimal stochastic scheduling in multi-class parallel queues. SIGMETRICS. Pp. 93–102.

Stadje, W. 1997. A new approach to the Lindley recursion. Statistics & Probability Letters. Vol. 31. No. 3. Pp. 169-175.

De Vuyst, S., Bruneel H., Fiems D. 2014. Computationally efficient evaluation of appointment schedules in health care. European Journal of Operational Research. Vol. 237. No. 3. Pp. 1142–1154.

AUTHOR BIOGRAPHIES

MIKHAIL KONOVALOV is a Doctor of Sciences in Technics and holds position of the principal scientist at Information Technologies Department at Institute of Informatics Problems of the Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences. His research activities are focused on adaptive control of random sequences, modelling and simulation of complex systems. His email address is `mkonovalov@ipiran.ru`.

ROSTISLAV RAZUMCHIK received his Ph.D. degree in Physics and Mathematics in 2011. Since then, he has worked as a leading research fellow at Institute of Informatics Problems of the Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS). Currently he also holds the associate professor position at Peoples' Friendship University of Russia (RUDN University). His current research activities are focused on queueing theory and its applications for performance evaluation of stochastic systems. His email address is `rrazumchik@ipiran.ru`

INFINITE-SERVER QUEUEING TANDEM WITH MMPP ARRIVALS AND RANDOM CAPACITY OF CUSTOMERS

Alexander Moiseev[1,2], Svetlana Moiseeva[1], Ekaterina Lisovskaya[1]
[1]Tomsk State University, 36 Lenina ave., Tomsk, 634050, Russia
[2]Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya st., Moscow, 117198, Russia
E-mail: moiseev.tsu@gmail.com, smoiseeva@mail.ru, ekaterina_lisovs@mail.ru

KEYWORDS
Infinite-server queueing system, random capacity of customer, Markovian modulated Poisson process.

ABSTRACT

Tandem of two queueing systems with infinite number of servers is considered. Customers arrive at the first stage of the tandem according to Markovian modulated Poisson process, and after a completion of their services, they go to the second stage. Each customer carries some data package with a random value (capacity of the customer). Service time does not depend on the customers' capacities in this study, capacities are used just to fix some additional characteristic of the system evolution. It is shown that two-dimensional probability distribution of total capacities at the stages of the system is two-dimensional Gaussian under the asymptotic condition of a high rate of arrivals. Presented numerical experiments and simulations allow to determine an applicability area of the asymptotic result.

INTRODUCTION

Infinite-server queuing systems are used as relevant models in some fields such as finance, insurance, etc. Furthermore, they may be applicable in case of models with a limited number of server devices as described in (Moiseev and Nazarov 2016a).

Queues with random capacities of customers are also useful for analysis and design problems in information and economic systems (Tikhonenko 1991; Tikhonenko and Klimovich 2001). In the case of information systems, the object under study is data received in the form of random-sized messages. In the case of economic systems, capacity of a customer refers to money that the customer pays to bank account. Also, such models are important in modeling of engineering devices where it is necessary to calculate a sufficient volume of buffer for data storing (Tikhonenko 2005 and 2015).

Results for single-server queues with limited buffer and LIFO service discipline were presented in the papers (Pechinkin 1998; Tikhonenko 2010). Algorithms for calculation of stationary characteristics were obtained for the models.

In the work (Cascone et al. 2011), the system Geo$_k$/G/1/∞ was studied under a condition of limited total capacity of customers. In that paper, capacities were modeled as discrete random variables that allowed the authors to obtain simple and efficient algorithms for the calculation of basic stationary characteristics of the system evolution.

In the paper (Naumov et al. 2016), a multi-server queue with losses is considered. Losses in the model are caused by the lack of resources required for customers' service. A customer accepted for service takes a random amount of resources of several types according to given distribution functions. Under the assumption of Poisson arrivals and exponential service time, authors derived an asymptotical joint probability distribution of the number of customers in the system and a distribution of vector of occupied resources' volumes. There is an example which illustrates applying of the model for analysis of characteristics of video conference service in a wireless network LTE.

A new trend in the study of queuing systems is analysis of the systems with non-Poisson arrivals and non-exponential service time. So, in the works (Pankratova and Moiseeva 2014 and 2015; Moiseev and Nazarov 2016a), queues and networks with MAP and renewal arrivals are studied under various asymptotic conditions.

Tandem queues (Reich 1957) are the models with sequential processing of customers at the stages of the system. When a customer arrives at the system, it goes to the first stage of the tandem. There, it is serviced during a random period, and when the service is complete, it goes to the next stage, and so on, until its service will be completed at the last stage of the system. The analytical results for the number of customers at the stages of the system were obtained in the papers (Moiseev and Nazarov 2014 and 2016b) for tandem queues with renewal and MAP arrivals and non-exponential service time. Analytical results about distributions of total capacities of customers at the stages of tandem queues for such models are not obtained yet.

The goal of the paper is to study total capacities' volume at the stages of the tandem system with incoming Markovian modulated Poisson process, two stages with infinite number of servers and non-exponential service time distribution.

MATEMATICAL MODEL

Consider a queue tandem with two stages and infinite number of servers at each stage. Customers arrive according to Markovian modulated Poisson process (MMPP). The process is given by generator matrix

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Q=||q_{ij}|| of size $K \times K$ and conditional rates $\lambda_1,...,\lambda_K$ which we compose into a diagonal matrix $\mathbf{\Lambda} = \text{diag}\{\lambda_1,...,\lambda_K\}$. Denote the underlying Markov chain of the MMPP as $k(t) \in 1,2,...,K$. Let each customer has some random capacity $\upsilon > 0$ with distribution function $G(y)$. Arriving customer instantly occupies a server at the first stage of the system. Service time at this stage has distribution function $B_1(x)$. When the service is complete, the customer moves to the second stage for the further service. Service time at the second stage has distribution function $B_2(x)$. When the service is complete at the second stage, the customer leaves the system. Customers' capacities, service times are not dependent on each other and are not dependent on epochs of customers' arrivals.

Denote the number of customers at the first and at the second stages of the system at a moment t by $i_1(t)$ and $i_2(t)$, and denote the total capacities of all customers at the first and at the second stages by $V_1(t)$ and $V_2(t)$ respectively. Let us obtain probabilistic characteristics of multi-dimensional process $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$. This process is not Markovian, therefore, we use the dynamic screening method (Moiseev and Nazarov 2016a) for its investigation.

Consider three time axes that are numbered from 0 to 2 (Fig. 1). Let axis 0 show the epochs of customers' arrivals. Axes 1 and 2 correspond to the stages of the system.

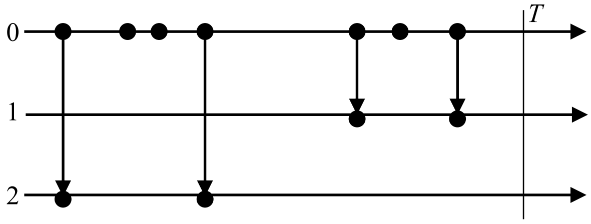


Figure 1: Screening of the Customers' Arrivals

We introduce a set of two functions $S_1(t)$, $S_2(t)$ (dynamic probabilities) that satisfy the conditions

$$0 \le S_1(t) \le 1, \quad 0 \le S_2(t) \le 1, \quad S_1(t) + S_2(t) \le 1 .$$

We assume that an epoch t of customer's arrival may be screened to axis 1 with the probability $S_1(t)$, or to axis 2 with the probability $S_2(t)$, or may be not screened anywhere with the probability $S_0(t) = 1 - S_1(t) - S_2(t)$.

Let the system be empty at moment t_0, and let us fix some arbitrary moment T in the future. We will use $S_1(t)$ as the probability that a customer arrived at the moment t will be serviced at the first stage at the moment T, and $S_2(t)$ as the probability that a customer arrived at the moment t will be serviced at the second stage of the system at the moment T. It is shown (Moiseev and Nazarov 2014) that

$$S_1(t) = 1 - B_1(T - t), \quad S_2(t) = B_1(T - t) - B_2^*(T - t)$$

for $t_0 \le t \le T$, where

$$B_2^*(\tau) = (B_1 * B_2)(\tau) = \int_0^\tau B_2(\tau - y) dB_1(y)$$

is a convolution of functions $B_1(x)$ and $B_2(x)$.

Denote the number of arrivals screened before the moment t on axes 1 and 2 by $n_1(t)$ and $n_2(t)$, and denote the total capacities of customers screened on axis 1 and 2 by $W_1(t)$ and $W_2(t)$ respectively.

As it is shown in (Moiseev and Nazarov 2016b), the multi-dimensional joint probability distribution of the number of customers at the stages of the tandem system at the moment T coincides with multi-dimensional joint probability distribution of the number of screened arrivals on respective axes:

$$P\{i_1(T) = m_1, i_2(T) = m_2\} = P\{n_1(T) = m_1, n_2(T) = m_2\}$$

for all $m_1, m_2 = 0, 1, 2, \dots$ It is easy to prove the same property for extended process $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$:

$$P\{i_1(T) = m_1, V_1(T) < z_1, i_2(T) = m_2, V_2(T) < z_2\} = \\ P\{n_1(T) = m_1, W_1(T) < z_1, n_2(T) = m_2, W_2(T) < z_2\} \quad (1)$$

for all $m_1, m_2 = 0, 1, 2, \dots$ and $z_1, z_2 \ge 0$. We use Equalities (1) for investigation of the process $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$ via analysis of the process $\{n_1(t), W_1(t), n_2(t), W_2(t)\}$.

KOLMOGOROV DIFFERENTIAL EQUATIONS

Let us consider the five-dimensional Markovian process $\{k(t), n_1(t), W_1(t), n_2(t), W_2(t)\}$. Denoting the probability distribution of this process by $P(k, n_1, w_1, n_2, w_2, t) = P\{k(t) = k, n_1(t) = n_1, W_1(t) < w_1, n_2(t) = n_2, W_2(t) < w_2\}$ and taking into account the formula of total probability, we can write the following system of Kolmogorov differential equations:

$$\frac{\partial P(k, n_1, w_1, n_2, w_2, t)}{\partial t} = \\ -\lambda_k (S_1(t) + S_2(t)) P(k, n_1, w_1, n_2, w_2, t) + \\ \lambda_k S_1(t) \int_0^{w_1} P(k, n_1 - 1, w_1 - y, n_2, w_2, t) dG(y) + \\ \lambda_k S_2(t) \int_0^{w_2} P(k, n_1, w_1, n_2 - 1, w_2 - y, t) dG(y) + \\ \sum_{\nu} q_{\nu k} P(\nu, n_1, w_1, n_2, w_2, t)$$

for $k = 1 \dots K$, n_1, $n_2 = 0,1,2,\dots$, w_1, $w_2 > 0$.

We introduce the partial characteristic function:

$$h(k, u_1, v_1, u_2, v_2, t) = \\ \sum_{n_1=0}^{\infty} e^{ju_1 n_1} \int_0^{\infty} e^{jv_1 w_1} \sum_{n_2=0}^{\infty} e^{ju_2 n_2} \int_0^{\infty} e^{jv_2 w_2} P(k, n_1, dw_1, n_2, dw_2, t),$$

where $j = \sqrt{-1}$ is an imaginary unit. Then we can write the following equations:

$$\frac{\partial h(k, u_1, v_1, u_2, v_2, t)}{\partial t} = \lambda_k h(k, u_1, v_1, u_2, v_2, t) \cdot$$

$$\left[S_1(t)\left(e^{ju_1}G^*(v_1)-1\right)+S_2(t)\left(e^{ju_2}G^*(v_2)-1\right)\right]+ \sum_{\nu} q_{\nu k} h(\nu,n_1,w_1,n_2,w_2,t)$$

for $k = 1\ldots K$, where $G^*(v)=\int\limits_0^\infty e^{jvy}dG(y)$.

Let us rewrite this system in the matrix form:

$$\frac{\partial \mathbf{H}(u_1,v_1,u_2,v_2,t)}{\partial t}=\mathbf{H}(u_1,v_1,u_2,v_2,t)\cdot \left[\mathbf{\Lambda}\left(S_1(t)\left(e^{ju_1}G^*(v_1)-1\right)+S_2(t)\left(e^{ju_2}G^*(v_2)-1\right)\right)+\mathbf{Q}\right] \quad (2)$$

with the initial condition $\mathbf{H}(u_1,v_1,u_2,v_2,t_0)=\mathbf{r}$, where

$$\mathbf{H}(u_1, v_1, u_2, v_2, t) = [h(1, u_1, v_1, u_2, v_2, t), \ldots, h(K, u_1, v_1, u_2, v_2, t)],$$

and $\mathbf{r} = [r(1), \ldots, r(K)]$ is a vector of the stationary distribution of the underlying Markov chain. Vector $\mathbf{r}$ satisfies the following linear system:

$$\begin{cases}\mathbf{rQ}=\mathbf{0},\\ \mathbf{re}=1,\end{cases} \quad (3)$$

where $\mathbf{e}$ is a column vector with all entries equal to 1.

ASYMPTOTIC ANALYSIS

The exact solution of Equation (2) is not possible in general case, but it may be solved under an asymptotic condition. In the paper, we consider the asymptotic condition of an infinitely growing arrivals' rate. Let us substitute $\mathbf{\Lambda} = N\mathbf{\Lambda}_1$ and $\mathbf{Q} = N\mathbf{Q}_1$ into Equation (2), where N is some parameter which is used for the asymptotic analysis ($N \to \infty$ in theoretical studies). Then Equation (2) may be rewritten as follows:

$$\frac{1}{N}\frac{\partial \mathbf{H}(u_1,v_1,u_2,v_2,t)}{\partial t}=\mathbf{H}(u_1,v_1,u_2,v_2,t)[\mathbf{Q}_1+ \mathbf{\Lambda}_1\left(S_1(t)\left(e^{ju_1}G^*(v_1)-1\right)+S_2(t)\left(e^{ju_2}G^*(v_2)-1\right)\right)] \quad (4)$$

with the initial condition

$$\mathbf{H}(u_1,v_1,u_2,v_2,t_0)=\mathbf{r}\,. \quad (5)$$

We solve Problem (4)–(5) under the asymptotic condition and we obtain a solution in the form of approximations which are named as "first-order asymptotic" $\mathbf{H}(u_1,v_1,u_2,v_2,t)\approx \mathbf{H}^{(1)}(u_1,v_1,u_2,v_2,t)$ and "second-order asymptotic" $\mathbf{H}(u_1,v_1,u_2,v_2,t)\approx \mathbf{H}^{(2)}(u_1,v_1,u_2,v_2,t)$. These approximations have different order of accuracy.

First-order Asymptotic Analysis

We formulate and prove the following statement.

Lemma. The first-order asymptotic characteristic function of the probability distribution of the process $\{k(t), n_1(t), W_1(t), n_2(t), W_2(t)\}$ has the form

$$\mathbf{H}^{(1)}(u_1,v_1,u_2,v_2,t)=\mathbf{r}\exp\left\{N\kappa_1\left[(ju_1+jv_1a_1)\int\limits_{t_0}^{t}S_1(\tau)d\tau+ (ju_2+jv_2a_1)\int\limits_{t_0}^{t}S_2(\tau)d\tau\right]\right\},$$

where $\kappa_1=\mathbf{r\Lambda}_1\mathbf{e}$, and $a_1=\int\limits_0^\infty ydG(y)$ is the mean of a customer capacity.

Proof.

Let us perform the substitutions

$$\varepsilon=\frac{1}{N},\ u_1=\varepsilon x_1,\ v_1=\varepsilon y_1\ \ u_2=\varepsilon x_2,\ v_2=\varepsilon y_2,$$
$$\mathbf{H}(u_1,v_1,u_2,v_2,t)=\mathbf{F}_1(x_1,y_1,x_2,y_2,t,\varepsilon)$$

in Expressions (4) and (5). Using such substitutions allows us to exclude a direct influence of an asymptotic parameter from the variables x_1, x_2, y_1, y_2. Then we obtain the following equation

$$\varepsilon\frac{\partial \mathbf{F}_1(x_1,y_1,x_2,y_2,t,\varepsilon)}{\partial t}=\mathbf{F}_1(x_1,y_1,x_2,y_2,t,\varepsilon)\cdot \quad (6)$$
$$\left\{\mathbf{Q}_1+\mathbf{\Lambda}_1\left[S_1(t)\left(e^{j\varepsilon x_1}G^*(\varepsilon y_1)-1\right)+S_2(t)\left(e^{j\varepsilon x_2}G^*(\varepsilon y_2)-1\right)\right]\right\}$$

with initial condition

$$\mathbf{F}_1(x_1,y_1,x_2,y_2,t_0,\varepsilon)=\mathbf{r}\,. \quad (7)$$

Let us find the asymptotic solution of Problem (6)–(7) $\mathbf{F}_1(x_1,y_1,x_2,y_2,t)=\lim\limits_{\varepsilon\to 0}\mathbf{F}_1(x_1,y_1,x_2,y_2,t,\varepsilon)$ in two steps.

Step 1. Substituting $\varepsilon = 0$ in (6), we obtain

$$\mathbf{F}_1(x_1,y_1,x_2,y_2,t)\mathbf{Q}_1=\mathbf{0}\,.$$

Comparing this equation with the first one in (3), we can conclude that $\mathbf{F}_1(x_1,y_1,x_2,y_2,t)$ can be expressed as

$$\mathbf{F}_1(x_1,y_1,x_2,y_2,t)=\mathbf{r}\Phi_1(x_1,y_1,x_2,y_2,t)\,, \quad (8)$$

where $\Phi_1(x_1,y_1,x_2,y_2,t)$ is some scalar function which satisfies the condition

$$\Phi_1(x_1,y_1,x_2,y_2,t_0)=1\,. \quad (9)$$

Step 2. Let us multiply (6) by vector $\mathbf{e}$, substitute (8), divide the results by ε and perform the asymptotic transition $\varepsilon \to 0$. Then taking into account $\mathbf{Q}_1\mathbf{e} = \mathbf{0}$ and $\mathbf{re} = 1$, we obtain the following differential equation for the function $\Phi_1(x_1,y_1,x_2,y_2,t)$

$$\frac{\partial \Phi_1(x_1,y_1,x_2,y_2,t)}{\partial t}=\Phi_1(x_1,y_1,x_2,y_2,t)\cdot \kappa_1\left[S_1(t)(jx_1+jy_1a_1)+S_2(t)(jx_2+jy_2a_1)\right]. \quad (10)$$

The solution of Problem (9)–(10) is as follows:

$$\Phi_1(x_1,y_1,x_2,y_2,t)=\exp\left\{\kappa_1\left[(jx_1+jy_1a_1)\int\limits_{t_0}^{t}S_1(\tau)d\tau+\right.\right.$$

$$\left.\left.\left(jx_2 + jy_2 a_1\right)\int_{t_0}^{t} S_2(\tau)d\tau\right]\right\}.$$

Substituting this expression into (8) and making inverse substitutions, we obtain

$$\mathbf{H}(u_1,v_1,u_2,v_2,t) \approx \mathbf{H}^{(1)}(u_1,v_1,u_2,v_2,t) =$$

$$\mathbf{r}\exp\left\{\kappa_1\left[\left(jx_1 + jy_1 a_1\right)\int_{t_0}^{t} S_1(\tau)d\tau + \right.\right.$$

$$\left.\left.\left(jx_2 + jy_2 a_1\right)\int_{t_0}^{t} S_2(\tau)d\tau\right]\right\}.$$

Thus, the proof is complete.

Second-order Asymptotic Analysis

The main result is the following theorem.

Theorem. The second-order asymptotic characteristic function of the probability distribution of the process $\{k(t), n_1(t), W_1(t), n_2(t), W_2(t)\}$ has the form

$$\mathbf{H}^{(2)}(u_1,v_1,u_2,v_2,t) = \mathbf{r}\exp\left\{N\kappa_1\left(ju_1 + jv_1 a_1\right)\int_{t_0}^{t} S_1(\tau)d\tau + \right.$$

$$N\kappa_1\left(ju_2 + jv_2 a_1\right)\int_{t_0}^{t} S_2(\tau)d\tau +$$

$$\frac{(ju_1)^2}{2}\left(N\kappa_1\int_{t_0}^{t} S_1(\tau)d\tau + N\kappa_2\int_{t_0}^{t} S_1{}^2(\tau)d\tau\right) +$$

$$\frac{(jv_1)^2}{2}\left(N\kappa_1 a_2\int_{t_0}^{t} S_1(\tau)d\tau + N\kappa_2 a_1{}^2\int_{t_0}^{t} S_1{}^2(\tau)d\tau\right) +$$

$$ju_1\, jv_1\left(N\kappa_1 a_1\int_{t_0}^{t} S_1(\tau)d\tau + N\kappa_2 a_1\int_{t_0}^{t} S_1{}^2(\tau)d\tau\right) + \quad (11)$$

$$\frac{(ju_2)^2}{2}\left(N\kappa_1\int_{t_0}^{t} S_2(\tau)d\tau + N\kappa_2\int_{t_0}^{t} S_2{}^2(\tau)d\tau\right) +$$

$$\frac{(jv_2)^2}{2}\left(N\kappa_1 a_2\int_{t_0}^{t} S_2(\tau)d\tau + N\kappa_2 a_1{}^2\int_{t_0}^{t} S_2{}^2(\tau)d\tau\right) +$$

$$ju_2\, jv_2\left(N\kappa_1 a_1\int_{t_0}^{t} S_2(\tau)d\tau + N\kappa_2 a_1\int_{t_0}^{t} S_2{}^2(\tau)d\tau\right) +$$

$$ju_1\, ju_2\, N\kappa_1\int_{t_0}^{t} S_1(\tau)S_2(\tau)d\tau +$$

$$jv_1\, jv_2\, N\kappa_2 a_1{}^2\int_{t_0}^{t} S_1(\tau)S_2(\tau)d\tau +$$

$$\left.\left(ju_1\, jv_2 + ju_2\, jv_1\right)N\kappa_2 a_1\int_{t_0}^{t} S_1(\tau)S_2(\tau)d\tau\right\},$$

where $\kappa_2 = 2\mathbf{g}(\mathbf{\Lambda}_1 - \kappa_1\mathbf{I})\mathbf{e}$, $a_2 = \int_0^\infty y^2 dG(y)$, and a row vector $\mathbf{g}$ satisfies the linear matrix system

$$\begin{cases}\mathbf{gQ}_1 = \mathbf{r}(\kappa_1\mathbf{I} - \mathbf{\Lambda}_1),\\ \mathbf{ge} = 1.\end{cases}$$

Proof.

Let $\mathbf{H}_2(x_1,y_1,x_2,y_2,t)$ be a vector function that satisfies the equation

$$\mathbf{H}(u_1,v_1,u_2,v_2,t) = \mathbf{H}_2(u_1,v_1,u_2,v_2,t)\cdot \quad (12)$$

$$\exp\left\{\kappa_1\left[\left(jx_1 + jy_1 a_1\right)\int_{t_0}^{t} S_1(\tau)d\tau + \left(jx_2 + jy_2 a_1\right)\int_{t_0}^{t} S_2(\tau)d\tau\right]\right\}.$$

Substituting this expression into (4) and (5), we obtain the following problem:

$$\frac{1}{N}\frac{\partial\mathbf{H}_2(u_1,v_1,u_2,v_2,t)}{\partial t} + \mathbf{H}_2(u_1,v_1,u_2,v_2,t)\kappa_1\cdot$$

$$\left[\left(ju_1 + jv_1 a_1\right)S_1(t) + \left(ju_2 + jv_2 a_1\right)S_2(t)\right] =$$

$$\mathbf{H}_2(u_1,v_1,u_2,v_2,t)\cdot \quad (13)$$

$$\left[\mathbf{\Lambda}_1\left(S_1(t)\left(e^{ju_1}G^*(v_1) - 1\right) + S_2(t)\left(e^{ju_2}G^*(v_2) - 1\right)\right) + \mathbf{Q}_1\right]$$

with the initial condition

$$\mathbf{H}_2(u_1,v_1,u_2,v_2,t_0) = \mathbf{r}\,. \quad (14)$$

Let us make the substitutions

$$\varepsilon^2 = \frac{1}{N}\,,\ u_1 = \varepsilon x_1\,,\ w_1 = \varepsilon y_1\,,\ u_2 = \varepsilon x_2\,,\ w_2 = \varepsilon y_2\,,$$

$$\mathbf{H}_2(u_1,v_1,u_2,v_2,t) = \mathbf{F}_2(u_1,v_1,u_2,v_2,t,\varepsilon). \quad (15)$$

Using these notations, Problem (13)–(14) can be rewritten in the form

$$\varepsilon^2\frac{\partial\mathbf{F}_2(x_1,y_1,x_2,y_2,t,\varepsilon)}{\partial t} + \mathbf{F}_2(x_1,y_1,x_2,y_2,t,\varepsilon)\kappa_1\cdot$$

$$\left[S_1(t)\left(j\varepsilon x_1 + j\varepsilon y_1 a_1\right) + S_2(t)\left(j\varepsilon x_1 + j\varepsilon y_1 a_1\right)\right] =$$

$$\mathbf{F}_2(x_1,y_1,x_2,y_2,t,\varepsilon)\cdot \quad (16)$$

$$\left[\mathbf{\Lambda}_1\left(S_1(t)\left(e^{j\varepsilon x_1}G^*(\varepsilon y_1) - 1\right) + S_2(t)\left(e^{j\varepsilon x_2}G^*(\varepsilon y_2) - 1\right)\right) + \mathbf{Q}_1\right]$$

with the initial condition

$$\mathbf{F}_2(x_1,y_1,x_2,y_2,t_0,\varepsilon) = \mathbf{r}\,. \quad (17)$$

Let us find the asymptotic solution of this problem $\mathbf{F}_2(x_1,y_1,x_2,y_2,t) = \lim_{\varepsilon\to 0}\mathbf{F}_2(x_1,y_1,x_2,y_2,t,\varepsilon)$ in three steps.

Step 1. Substituting $\varepsilon = 0$ in (16)–(17), we obtain the following system of equations:

$$\begin{cases}\mathbf{F}_2(x_1,y_1,x_2,y_2,t)\mathbf{Q}_1 = \mathbf{0},\\ \mathbf{F}_2(x_1,y_1,x_2,y_2,t_0) = \mathbf{r}.\end{cases}$$

Therefore, taking into account (3), we can write

$$\mathbf{F}_2(x_1,y_1,x_2,y_2,t) = \mathbf{r}\Phi_2(x_1,y_1,x_2,y_2,t), \quad (18)$$

where $\Phi_2(x_1,y_1,x_2,y_2,t)$ is some scalar function which satisfies the condition

$$\Phi_2(x_1,y_1,x_2,y_2,t_0) = 1\,. \quad (19)$$

Step 2. Using (18), the function $\mathbf{F}_2(x_1, y_1, x_2, y_2, t)$ can be represented in the expansion form

$$\mathbf{F_2}(x_1, y_1, x_2, y_2, t, \varepsilon) = \Phi_2(x_1, y_1, x_2, y_2, t)[\mathbf{r} + \\ \mathbf{g}(S_1(t)(j\varepsilon x_1 + j\varepsilon y_1 a_1) + S_2(t)(j\varepsilon x_2 + j\varepsilon y_2 a_1))] + \mathbf{O}(\varepsilon^2), \quad (20)$$

where $\mathbf{g}$ is the row vector that satisfies the condition $\mathbf{ge} = 1$, and $\mathbf{O}(\varepsilon^2)$ is a row vector of infinitesimals of the order ε^2. Let us use substitution (20) and expansion $e^{j\varepsilon x} = 1 + j\varepsilon x + O(\varepsilon^2)$ in Equation (16). Taking into account (3) and making a transition $\varepsilon \to 0$, we obtain matrix equation for the vector $\mathbf{g}$

$$\mathbf{gQ}_1 = \mathbf{r}(\kappa_1 \mathbf{I} - \mathbf{\Lambda}_1),$$

where $\mathbf{I}$ is an identity matrix.

Step 3. We multiply Equation (16) by vector $\mathbf{e}$ and use Expression (20) and the second-order expansion

$$e^{j\varepsilon x} = 1 + j\varepsilon x + \frac{(j\varepsilon x)^2}{2} + O(\varepsilon^3).$$

After some transformations, using the notation

$$\kappa_2 = 2\mathbf{g}(\mathbf{\Lambda}_1 - \kappa_1 \mathbf{I})\mathbf{e},$$

we obtain the following differential equation for the function $\Phi_2(x_1, y_1, x_2, y_2, t)$

$$\frac{\partial \Phi_2(x_1, y_1, x_2, y_2, t)}{\partial t} = \Phi_2(x_1, y_1, x_2, y_2, t) \cdot \\ \left[\frac{(jx_1)^2}{2}\left(\kappa_1 S_1(t) + \kappa_2 S_1^2(t)\right) + \frac{(jy_1)^2}{2}\left(\kappa_1 a_2 S_1(t) + \kappa_2 a_1^2 S_1^2(t)\right)\right. \\ + jx_1\, jy_1\left(\kappa_1 a_1 S_1(t) + \kappa_2 a_1 S_1^2(t)\right) + \\ jx_2\, jy_2\left(\kappa_1 a_1 S_2(t) + \kappa_2 a_1 S_2^2(t)\right) + \\ jx_1\, jx_2\, \kappa_1 S_1(t) S_2(t) + jy_1\, jy_2\, \kappa_2 a_1^2 S_1(t) S_2(t) + \\ (jx_1\, jy_2 + jx_2\, jy_1)\kappa_2 a_1 S_1(t) S_2(t) + \frac{(jx_2)^2}{2}\left(\kappa_1 S_2(t) + \kappa_2 S_2^2(t)\right) \\ \left. + \frac{(jy_2)^2}{2}\left(\kappa_1 a_2 S_2(t) + \kappa_2 a_1^2 S_2^2(t)\right)\right].$$

The solution of this equation with initial condition (19) is as follows:

$$\Phi_2(x_1, y_1, x_2, y_2, t) = \\ \exp\left\{\frac{(jx_1)^2}{2}\left(\kappa_1 \int_{t_0}^{t} S_1(\tau)d\tau + \kappa_2 \int_{t_0}^{t} S_1^2(\tau)d\tau\right) + \right. \\ \frac{(jy_1)^2}{2}\left(\kappa_1 a_2 \int_{t_0}^{t} S_1(\tau)d\tau + \kappa_2 a_1^2 \int_{t_0}^{t} S_1^2(\tau)d\tau\right) + \\ jx_1\, jy_1\left(\kappa_1 a_1 \int_{t_0}^{t} S_1(\tau)d\tau + \kappa_2 a_1 \int_{t_0}^{t} S_1^2(\tau)d\tau\right) + \\ \frac{(jx_2)^2}{2}\left(\kappa_1 \int_{t_0}^{t} S_2(\tau)d\tau + \kappa_2 \int_{t_0}^{t} S_2^2(\tau)d\tau\right) + \\ \frac{(jy_2)^2}{2}\left(\kappa_1 a_2 \int_{t_0}^{t} S_2(\tau)d\tau + \kappa_2 a_1^2 \int_{t_0}^{t} S_2^2(\tau)d\tau\right) + \\ jx_2\, jy_2\left(\kappa_1 a_1 \int_{t_0}^{t} S_2(\tau)d\tau + \kappa_2 a_1 \int_{t_0}^{t} S_2^2(\tau)d\tau\right) + \\ jx_1\, jx_2\, \kappa_1 \int_{t_0}^{t} S_1(\tau) S_2(\tau)d\tau + \\ jy_1\, jy_2\, \kappa_2 a_1^2 \int_{t_0}^{t} S_1(\tau) S_2(\tau)d\tau + \\ \left.(jx_1\, jy_2 + jy_2\, jx_1)\kappa_2 a_1 \int_{t_0}^{t} S_1(\tau) S_2(\tau)d\tau\right\}.$$

Substituting this expression in Formula (18) and performing the substitutions that are inverse to (15) and (12), we obtain Expression (11) for the asymptotic characteristic function of the process $\{k(t), n_1(t), W_1(t), n_2(t), W_2(t)\}$. The proof is complete.

Corollary 1. Assuming $t = T$ and $t_0 \to -\infty$ and using Equalities (1), we obtain the steady-state characteristic function of the process under study $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$:

$$h(u_1, v_1, u_2, v_2) = \exp\{N\kappa_1(ju_1 + jv_1 a_1)b_{11} + \\ N\kappa_1(ju_2 + jv_2 a_1)b_{21} + \frac{(ju_1)^2}{2}(N\kappa_1 b_{11} + N\kappa_2 b_{12}) + \\ \frac{(jv_1)^2}{2}\left(N\kappa_1 a_2 b_{11} + N\kappa_2 a_1^2 b_{12}\right) + \\ ju_1\, jv_1(N\kappa_1 a_1 b_{11} + N\kappa_2 a_1 b_{12}) + \frac{(ju_2)^2}{2}(N\kappa_1 b_{21} + N\kappa_2 b_{22}) + \\ \frac{(jv_2)^2}{2}\left(N\kappa_1 a_2 b_{21} + N\kappa_2 a_1^2 b_{22}\right) + \\ ju_2\ \ jv_2(N\kappa_1 a_1 b_{21} + N\kappa_2 a_1 b_{22}) + \\ ju_1\, ju_2\, N\kappa_1 b + jv_1\, jv_2\, N\kappa_2 a_1^2 b + \\ (ju_1\, jv_2 + ju_2\, jv_1)N\kappa_2 a_1 b\}, \quad (21)$$

where

$$b_{11} = \int_0^\infty (1 - B_1(\tau))d\tau,\ b_{12} = \int_0^\infty (1 - B_1(\tau))^2 d\tau,$$

$$b_{21} = \int_0^\infty \left(B_1(\tau) - B_2^*(\tau)\right)d\tau,\ b_{22} = \int_0^\infty \left(B_1(\tau) - B_2^*(\tau)\right)^2 d\tau,$$

$$b = \int_0^\infty (1 - B_1(\tau))\left(B_1(\tau) - B_2^*(\tau)\right)d\tau.$$

From the form of characteristic function (21), it is clear that the probability distribution of four-dimensional process $\{i_1(t), V_1(t), i_2(t), V_2(t)\}$ is asymptotically Gaussian.

Corollary 2. The steady-state joint probability distribution of two-dimensional process of the total capacity at the stages of the system is asymptotically Gaussian with a vector of means

$$\mathbf{a} = N \cdot \left[\kappa_1 a_1 b_{11} \quad \kappa_1 a_1 b_{21} \right]$$

and a covariance matrix

$$\mathbf{K} = N \cdot \begin{bmatrix} \kappa_1 a_2 b_{11} + \kappa_2 a_1{}^2 b_{12} & \kappa_2 a_1{}^2 b \\ \kappa_2 a_1{}^2 b & \kappa_1 a_2 b_{21} + \kappa_2 a_1{}^2 b_{22} \end{bmatrix}.$$

NUMERICAL EXAMPLE

Result (21) is obtained under the asymptotic condition $N \to \infty$. Therefore, the result may be used just as an approximation and it is applicable when N is great enough. So, we need in determining of a low boundary of parameter N which cause the approximation (21) be applicable. To do this we make series of simulation experiments and compare asymptotic distributions with empiric ones by using the Kolmogorov distance

$$\Delta = \max_x \left| F(x) - A(x) \right| \qquad (22)$$

as an accuracy. Here $F(x)$ is a cumulative distribution function of total capacity of customers at a stage of the tandem which is constructed on the basis of simulation results, and $A(x)$ is a Gaussian cumulative distribution function with respective mean and variance from Expression (21). Increasing value of parameter N step by step from one experiment to another, we can find the value of N at which the accuracy (22) is small enough.

Consider the following numerical example. The MMPP is given by parameters $\mathbf{Q} = N\mathbf{Q}_1$ and $\mathbf{\Lambda} = N\mathbf{\Lambda}_1$ where

$$\mathbf{Q}_1 = \begin{bmatrix} -0{,}8 & 0{,}4 & 0{,}4 \\ 0{,}3 & -0{,}6 & 0{,}3 \\ 0{,}4 & 0{,}4 & -0{,}8 \end{bmatrix}, \quad \mathbf{\Lambda}_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}.$$

Fundamental rate of arrivals is $N\kappa_1 = N\mathbf{r}\mathbf{\Lambda}_1\mathbf{e} = 3 \cdot N$. Capacities of customers have uniform distribution in the range [0; 1]. Service time has gamma distribution with shape and inverse scale parameters $\alpha_1 = 1{,}5$ and $\beta_1 = 2$ at the first stage of the system, and $\alpha_2 = 0{,}5$ and $\beta_2 = 1{,}5$ at the second stage. So, the fundamental rate of arrivals exceeds exactly by N times the service rate at the second stage, therefore, we consider marginal distributions of the total capacity at this stage.

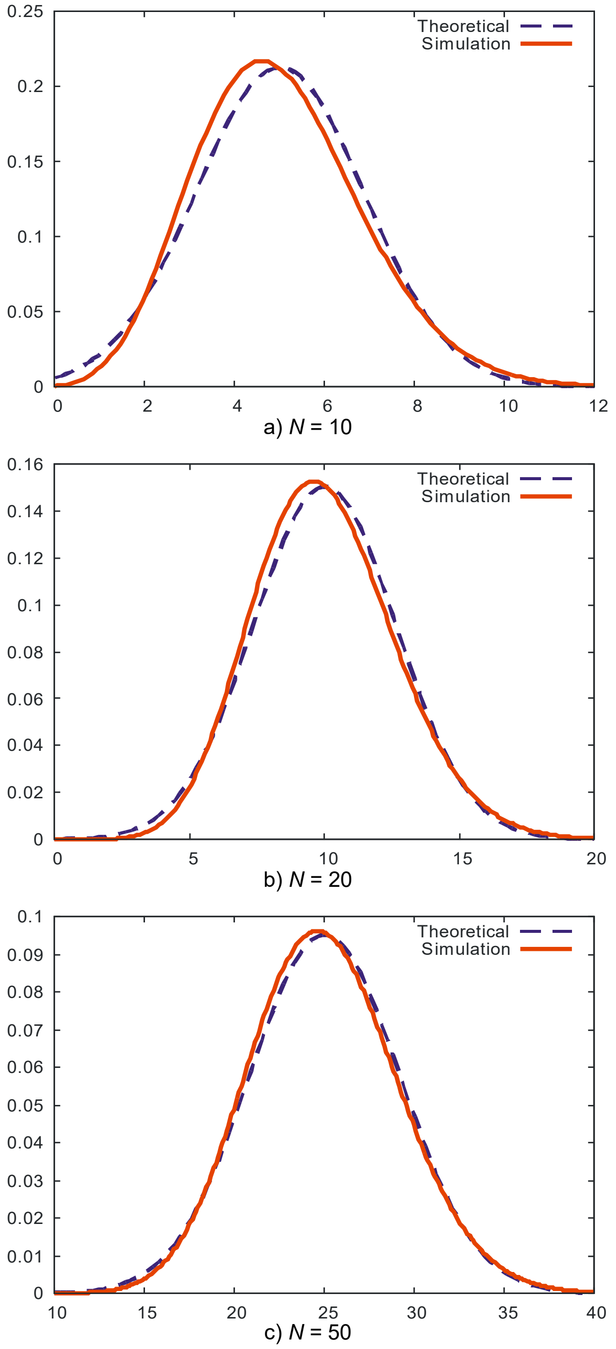
A vector of means and a covariance matrix of the Gaussian approximation for this example are as follows:

$$\mathbf{a} = N \cdot \left[0{,}125 \quad 0{,}5 \right], \; \mathbf{K} = N \cdot \begin{bmatrix} 0{,}921 & 0{,}039 \\ 0{,}039 & 0{,}352 \end{bmatrix}.$$

So, in Formula (22) $F(x)$ will be a cumulative distribution function of total capacity of customers at the second stage of the system constructed on simulation results, and $A(x)$ will be a Gaussian cumulative distribution function with mean and variance equal to 0,5N and 0,352N respectively. Values of the Kolmogorov distance for increasing values of parameter N are presented in Table 1. We can notice that the asymptotic results become more accurate while a value of the parameter N (fundamental rate of arrivals) is increasing. Figures 2 show probability densities of asymptotic and empiric distributions at the second stage of the system and they confirm the effect.

Table 1: Kolmogorov Distances between Simulation and Asymptotic Results for the Total Capacity

N	5	10	20	50	100
Δ	0,056	0,035	**0,027**	**0,015**	**0,009**



a) N = 10

b) N = 20

c) N = 50

Figures 2: Probability Densities of Asymptotic (Marked as “Theoretical”) and Empiric (Marked as “Simulation”) Distributions of the Total Capacity

We suppose that an approximation is applicable if its Kolmogorov distance less than 0,03. Then we can draw a conclusion that the asymptotic results are applicable for values of the parameter N equal to 20 or more (marked by boldface in Table 1).

CONCLUSION

In the paper, the queue tandem with MMPP arrivals, infinite number of servers and non-exponential service time is considered. The problem under study is a capacity which each customer brings to the system. The analysis is performed under the asymptotic condition of high rate of arrivals (high values of a fundamental rate of the MMPP). It is shown that two-dimensional probability distribution of total capacities at the stages of the system is two-dimensional Gaussian under this asymptotic condition. Numerical results show that asymptotic results have enough accuracy for marginal distribution of total capacity at a stage of the system when a fundamental rate of arrivals exceeds service rate at the stage by 20 times or more.

Future studies may be devoted to analysis of customers' capacities in queueing tandems with MAP arrivals and systems in random environment.

ACKNOWLEGEMENTS

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008). Authors also thank Dr. Michele Pagano (University of Pisa) for his help in numerical experiments.

REFERENCES

Cascone, A.; R. Manzo; A.V. Pechinkin; and S.Ya. Shorgin. 2011. "Geom/G/1/n System with LIFO Discipline without Interrupts and Constrained Total Amount of Customers". *Automation and Remote Control* 72, No.1 (Jan), 99-11.

Moiseev, A.N. and A.A. Nazarov. 2014. "Asymptotic Analysis of a Multistage Queuing System with a High-rate Renewal Arrival Process". *Optoelectronics, Instrumentation and Data Processing* 50, No.2, 163-171.

Moiseev, A. and A. Nazarov. 2016. "Queueing Network MAP–(GI–∞)^K with High-rate Arrivals". *European Journal of Operational Research* 254, 161-168.

Moiseev, A. and A. Nazarov. 2016. "Tandem of Infinite-server Queues with Markovian Arrival Process". *Communications in Computer and Information Science* 601, 323-333.

Naumov, V.A.; K.E. Samuilov; and A.K. Samuilov. 2016. "On the Total Amount of Resources Occupied by Serviced Customers". *Automation and Remote Control* 77, No.8 (Aug), 1419-1427.

Pankratova, E. and S. Moiseeva. 2014. "Queueing System MAP|M|∞ with n Types of Customers". *Communications in Computer and Information Science* 487, 356-366.

Pankratova, E. and S. Moiseeva. 2015. "Queueing System with Renewal Arrival Process and Two Types of Customers". In *Proceedings of the 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshop* (St. Petersburg, 2014, Oct. 06-08). IEEE, St. Petersburg, 514-517.

Pechinkin, A. 1998. "The M/G/1/n System with LIFO Service Discipline with Interruptions and Limitations on the Total Amount of Requests". *Automation and* Remote *Control* 59, No.4 (Apr), 545-553.

Reich, E. 1957. "Waiting Times When Queues are in Tandem". *Annals of Mathematical Statistics* 28, No.3 (Mar), 768-773.

Tikhonenko, O.M. 1991. "Queuing System for Random Length with Restrictions." *Automation and Remote Control* 52, No.10 (Oct), 1431-1437.

Tikhonenko, O.M. 2005. "Generalized Erlang Problem for Service Systems with Finite Total Capacity". *Problems of Information Transmission* 41, No.3 (Mar), 243-253.

Tikhonenko, O.M. 2010. "Queuing Systems with Processor Sharing and Limited Resources". *Automation and Remote Control* 71, No.5 (May), 803-815.

Tikhonenko, O.M. 2015. "Queuing Systems with Processor Sharing and Limited Memory under Control of the AQM Mechanism". *Automation and Remote Control* 76, No.10 (Oct), 1784-1796.

Tikhonenko, O.M and K.G. Klimovich. 2001. "Analysis of Queuing Systems for Random-length Arrivals with Limited Cumulative Volume". *Problems of Information Transmission* 37, No.1 (Jan), 70-79.

AUTHOR BIOGRAPHIES

ALEXANDER MOISEEV graduated from Tomsk State University (Russia) in 1993. He obtained Ph.D. degree in engineering in 1999 and became a Doctor in Physics and Mathematics in 2016. He works as a lecturer and as a scientist in Tomsk State University and RUDN University now. The fields of his scientific interest are queueing theory, mathematical modeling and simulation. He is a regular organizer of the International conference "Information Technologies and Mathematical Modelling". His e-mail address is: moiseev.tsu@gmail.com.

SVETLANA MOISEEVA is a Doctor of Science in Physics and Mathematics. She is a professor of Department of probability theory and mathematical statistics in Tomsk State University, Russia. Her scientific interest is theory of stochastic processes, queueing theory and their applications to telecommunications, call centers, economic systems, etc. She is one of main organizers of the International conference "Information Technologies and Mathematical Modelling". Her e-mail address is: smoiseeva@mail.ru.

EKATERINA LISOVSKAYA is an assistant of Department of probability theory and mathematical statistics in Tomsk State University, Russia. Her scientific interest is theory of stochastic processes, queueing theory and their applications to telecommunications, call centers, etc. Her e-mail address is: ekaterina_lisovs@mail.ru.

ANALYSIS OF UNRELIABLE MULTI-SERVER QUEUEING SYSTEM WITH BREAKDOWNS SPREAD AND QUARANTINE

Alexander Dudin[a]

Sergei Dudin[a,b]
Olga Dudina[a,b]

Konstantin Samouylov[b]

[a]Department of Applied Mathematics
and Computer Science
Belarusian State University,
4 Nezavisimosti Ave.,220030,
Minsk, Belarus

[b]Department of Applied Probability
and Informatics
RUDN University,
6 Miklukho-Maklaya st., 117198,
Moscow, Russia

Email:dudin@bsu.by

Email: dudins@bsu.by
dudina@bsu.by

Email:ksam@sci.pfu.edu.ru

KEYWORDS

Multi-server queueing system, breakdowns, quarantine, Markovian arrival flow

ABSTRACT

We consider an unreliable multi-server queue in which the rate of servers' breakdowns increases when the number of broken servers grows. To prevent quick degradation of the system, it is proposed to switch to a quarantine regime when the number of broken servers exceeds some threshold and to maintain this regime until the number of broken servers becomes less than another threshold. During the quarantine, service of customers is stopped, new breakdowns do not arrive while the broken servers continue recovering. Under the fixed values of the thresholds, behavior of the system is described by the multi-dimensional continuous time Markov chain. The steady state distribution of the chain and the key performance measures of the system are computed as the functions of the thresholds. Possibility of the optimal choice of the thresholds providing the minimal value of an economical criterion is numerically illustrated.

INTRODUCTION

Queueing theory provides powerful mathematical and algorithmic tool for analysis of a variety systems where a certain scarce resource is shared between the competitive users who generate requests at random moments. As an example of such a resource we can mention the channels and servers of a telecommunication system or operators and equipment of a contact center. Important feature of many queueing systems is the unreliability of the servers, i.e., possibility of their failure. Such systems are called unreliable queueing systems. As the first papers devoted to analysis of the multi-server unreliable queueing systems, the papers (Mitrani and Avi-Itzhak 1968) and (Neuts and Lucantoni 1979) deserve to be mentioned. In these papers, the multi-server unreliable queueing systems with identical servers, the stationary Poisson arrival processes of customers and breakdowns and the exponential distribution of customers service time and servers recovering time are considered. Behavior of the systems is described by the two-dimensional Markov chains where one component defines the number of customers in the system and the second component is the number of operable (non-broken) servers. Such Markov chains were analysed using the partial generating functions in (Mitrani and Avi-Itzhak 1968) and the matrix analytic approach in (Neuts and Lucantoni 1979).

Assumption about the stationary Poisson arrival processes (in such a process, the inter-arrival times are independent identical exponentially distributed random variables) is not realistic for queueing systems describing operation of modern telecommunication networks where information flows exhibit significant burstiness and correlation. As more realistic model of information flows, the model of the Markovian Arrival Process (MAP) was developed, see, e.g., (Chakravarthy 2001, Lucantoni 1991). Unreliable multi-server queueing systems with the MAP were considered, e.g., in (Klimenok et al. 2008, Dudin et al. 2015, Al-Begain et al. 2012).

Distinguishing feature of the model analysed in this paper is consideration of the breakdowns arrival process which, to the best of our knowledge, is not considered in the queueing literature. The most popular in the queueing literature assumption is that each operable server breaks down independently of other servers and the intensity of its breakdowns is constant, i.e., the total intensity of breakdowns is proportional to the number of *operable* servers and, therefore, it decreases

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

when the number of the broken servers grows. In this paper, we make the opposite assumption, not considered in the queueing literature previously. We assume that the total intensity of breakdowns increases when the number of *broken* servers grows.

This assumption may be true in the situations when the breakdown consists of infecting the server in some database by a computer virus or the human operator in contact center by a virus. Because the servers may use common hardware, tables and indices of databases and software while the operators may be compactly located and infections spreads by airborne transmission, infection of one server may provoke the quick spread of infection to another servers. Another situation where the total intensity of breakdowns increases when the number of broken servers grows is as follows. Smooth operation of the servers in some manufacturing system is provided via implementation of certain preventive works. If these works are done by a limited team of repairmen along with recovering the broken servers, this team may have a lack of time for preventive works when the number of broken servers becomes large. In such a situation, each operable server experiences the increased load and this, along with the absence of enough preventive work, may cause more quick failure of the server.

It is clear that the increase of the total intensity of breakdowns when the number of broken servers grows may eventually lead to the full degradation of the system if a certain additional (to the routine repair works) mechanism for reducing the number of broken servers will be not applied. To propose such a mechanism, it is worth to note that an effective way to struggle against the spread of infection in medicine is to impose the quarantine (ward closure) when the number of infected persons becomes large. When the regime of quarantine is established in some entity (school class, children group in kindergarten, hospital, etc), this entity stops its routine operation and members of the entity are isolated in the maximally possible extent until the level of infection drops to admissible level. By analogy, to prevent quick spread of breakdowns (up to the complete collapse of the system), we assume that the considered queueing system may switch to quarantine regime. This regime suggests that all servers stop operation, new customers may arrive but they are not allowed to enter service and are stored to the buffer. New breakdowns do not arrive. The broken servers are repaired. Thus, in this paper, we consider the unreliable multi-server queueing system with possibility to use the quarantine regime. The strategy of using this regime is as follows. The system switches from the operable mode to quarantine regime when the number of broken servers exceeds some predefined threshold, say, M_1. The system switches back to the operable mode when the number of broken servers drops below other threshold, say M_2, where $M_2 < M_1$. The final goal of the study is to provide the way for choosing the optimal values of the thresholds M_1 and M_2 providing the best quality of the system operation.

MATHEMATICAL MODEL

We consider an N-server queueing system with an infinite buffer and Markovian arrival process of customers. The structure of the system under study is presented in Figure 1.

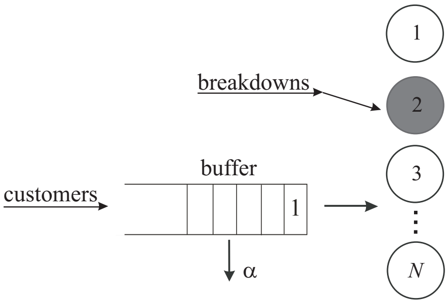


Figure 1: Queueing system under study

Customers arrive at the system according to the *MAP*. The advantage of the *MAP* pattern of the arrival process comparing to the popular in the queueing literature model of the stationary Poisson process (which is a very particular case of the *MAP*) is that the *MAP* allows to take into account correlation of the successive inter-arrival times and burstyness typical for modern telecommunication networks. Arrivals in the *MAP* are directed by an irreducible continuous-time Markov chain $w_t,\ t \geq 0$, with the finite state space $\{0, 1, ..., W\}$. The intensities of transitions that are accompanied by the arrival of k customers, are combined to the square matrices $D_k,\ k = 0, 1$, of size $W + 1$. Formulas for computation of the average intensity λ (fundamental rate) of the *MAP*, the squared coefficient of the variation and the coefficient of correlation of intervals between successive arrivals can be found, e.g., in (Chakravarthy 2001).

The service time of a customer by a server has an exponential distribution with the parameter μ, $\mu > 0$. The servers can break down. The intensity of breakdowns arrival depends on the current number of broken servers. When the number of broken servers is k, the intensity of breakdowns arrival is equal to $\gamma_k,\ k = \overline{0, N}$. By default, we assume that the intensity γ_k does not decrease when k grows, i.e. $\gamma_k \leq \gamma_{k+1},\ k = \overline{0, N-1}$. Arrival of a breakdown implies the failure of an arbitrary non-broken server. If the broken server provided service to a customer, this customer moves to a free server if it is available, or joins the queue from the head, otherwise.

Each broken server is recovering, independently of another broken servers, during the exponentially distributed time with the intensity β. After recovering, the server immediately resumes service.

Because we assume that the intensity of breakdowns γ_k, generally speaking, increases when the number of broken servers k grows, to prevent the quick spread of servers failures (and possible complete crash of the system), we suggest that when the number of broken servers becomes large the system can pass to the regime of quarantine. This means that all servers stop service, customers, which are getting service, move to the head of the queue in the random order. During the quarantine, new customers can arrive and move to the buffer while the arrivals of the new breakdowns are ignored. The strategy of the quarantine beginning and ending is defined by the thresholds, M_1 and M_2, such as $1 \leq M_1 \leq N,\ \ 0 \leq M_2 < M_1$. If the number of broken servers reaches the level M_1, the quarantine starts. It finishes when the number of broken servers falls to the level M_2. When the quarantine ends, all non-broken servers immediately resume service.

The customers staying in the buffer are impatient, i.e., the customer leaves the buffer and the system after an exponentially distributed waiting time defined by the parameter α, $0 < \alpha < \infty$.

Our goal is to analyse the stationary behavior of the described system under the fixed parameters of the system and the thresholds M_1 and M_2 and to illustrate the effect of these thresholds. It is clear that the problem of the optimal choosing the thresholds is not the trivial one. If M_1 is large, the system benefits from late interruption of service but then it suffers from many servers being broken and not available for service provisioning. If M_1 is small, the situation when many servers are broken is avoided, however, the systems wastes time due to the frequent use of quarantines. If M_2 is small, the quarantine is long and performance of the servers is lost. If M_2 is large, the system weakly uses an opportunity to temporarily ignore new breakdowns arrival and repair the broken servers. Therefore, likely the next quarantine will be required very soon. Thus, to provide the best conditions for customers service, the careful quantitative analysis of the model is required.

PROCESS OF SYSTEM STATES

It is easy to see that the behavior of the system under study is described by the following regular irreducible continuous-time Markov chain

$$\xi_t = \{i_t, r_t, k_t, w_t\},\ t \geq 0,$$

where, during the epoch t,

- i_t is the number of customers in the system, $i_t \geq 0$;
- r_t is an indicator that specifies the state of the system: $r_t = 0$ corresponds to the operable state and $r_t = 1$ corresponds to the quarantine regime;
- k_t is the number of broken servers, $k_t = \overline{0, M_1}$;
- w_t is the state of the underlying process of the MAP, $w_t = \overline{0, W}$.

The Markov chain ξ_t, $t \geq 0$, has the following state space:

$$\left(\{i, 0, k, w\},\ k = \overline{0, M_1 - 1}\right) \bigcup$$

$$\left(\{i, 1, k, w\},\ k = \overline{M_2 + 1, M_1}\right), i \geq 0,\ w = \overline{0, W}.$$

To simplify the analysis of the Markov chain ξ_t, let us enumerate the states of this chain in the direct lexicographic order of the components r, k, w and refer to the set of the states of the Markov chain having values (i, r) of the first two components as the macro-state (i, r). Let Q be the generator of the Markov chain ξ_t. It consists of the blocks $Q_{i,j},\ i, j \geq 0$, each of which contains four blocks $Q_{i,j}^{(r,r')},\ r, r' = 0, 1$, defining the intensities of the transition from the macro-state (i, r) to the macro-state (j, r').

Analysing all possible transitions of the Markov chain ξ_t during an interval of an infinitesimal length and rewriting the intensities of these transitions in the block matrix form we obtain the following result.

Theorem 1. The infinitesimal generator $Q = (Q_{i,j})_{i,j \geq 0}$ of the Markov chain $\xi_t,\ t \geq 0$, has a block-tridiagonal structure:

$$Q = \begin{pmatrix} Q_{0,0} & Q^+ & O & O & \dots \\ Q_{1,0} & Q_{1,1} & Q^+ & O & \dots \\ O & Q_{2,1} & Q_{2,2} & Q^+ & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The matrix $Q_{i,i}$, $i \geq 0$, has the following form:

$$Q_{i,i} = \begin{pmatrix} Q_{i,i}^{(0,0)} & Q_{i,i}^{(0,1)} \\ Q_{i,i}^{(1,0)} & Q_{i,i}^{(1,1)} \end{pmatrix}$$

where

$$Q_{i,i}^{(0,0)} = I_{M_1} \otimes D_0 - (\alpha A_i + \beta C_1 (I_{M_1} - E_1^-) +$$

$$G(I_{M_1} - E^+) + \mu N_i) \otimes I_{\bar{W}},$$

$$Q_{i,i}^{(0,1)} = G\hat{I} \otimes I_{\bar{W}},\ \ Q_{i,i}^{(1,0)} = \beta \tilde{I} \otimes I_{\bar{W}},$$

$$Q_{i,i}^{(1,1)} = I_{M_1 - M_2} \otimes D_0 - i\alpha I_{(M_1 - M_2)\bar{W}} -$$

$$\beta C_2 (I_{M_1 - M_2} - E_2^-) \otimes I_{\bar{W}},\ i \geq 0,$$

$$Q_{i,i-1} = \text{diag}\{(\mu N_i + \alpha A_i) \otimes I_{\bar{W}}, i\alpha I_{(M_1 - M_2)\bar{W}}\},\ i \geq 1,$$

$$Q^+ = I_{2M_1 - M_2} \otimes D_1.$$

Here,

- I is the identity matrix, and O is a zero matrix of appropriate dimension;
- $\otimes$ indicates the symbol of Kronecker product of matrices, see (Graham 1981);
- $\bar{W} = W + 1$;
- diag$\{\dots\}$ denotes the diagonal matrix with the diagonal blocks listed in the brackets;
- $A_i = \text{diag}\{\max\{i - N, 0\}, \max\{i - (N-1), 0\}, \dots, \max\{i - (N - M_1 - 1), 0\}\}$;
- $N_i = \text{diag}\{\min\{i, N\}, \max\{\min\{i, N\} - 1, 0\}, \dots, \max\{\min\{i, N\} - M_1 - 1, 0\}\}$;
- $C_1 = \text{diag}\{0, 1, \dots, M_1 - 1\}$;
- $C_2 = \text{diag}\{M_2 + 1, M_2 + 2, \dots, M_1\}$;
- $G = \text{diag}\{\gamma_0, \gamma_1, \dots, \gamma_{M_1 - 1}\}$;
- E^+ is the square matrix of size M_1 with all zero entries except the entries $(E^+)_{l,l+1}$, $l = \overline{0, M_1 - 2}$, which are equal to 1;
- E_1^- is the square matrix of size M_1 with all zero entries except the entries $(E_1^-)_{l,l-1}$, $l = \overline{1, M_1 - 1}$, which are equal to 1;
- E_2^- is the square matrix of size $M_1 - M_2$ with all zero entries except the entries $(E_2^-)_{l,l-1}$, $l = \overline{1, M_1 - M_2}$, which are equal to 1;

- $\hat{I}$ is the matrix of size $M_1 \times (M_1 - M_2)$ with all zero entries except the entry $(\hat{I})_{M_1-1, M_1-M_2-1}$ which is equal to 1;
- $\tilde{I}$ is the matrix of size $(M_1 - M_2) \times M_1$ with all zero entries except the entry $(\hat{I})_{0,M_2}$ which is equal to $M_2 + 1$.

Proof. The entries of the blocks Q^+ define the intensities of the transition of the Markov chain $\xi_t = \{i_t, r_t, k_t, w_t\}$, $t \geq 0$, that lead to increase of the component i_t (the number of customers in the system) by one. This can happen only if a customer arrives to the system, i.e., the underlying process w_t of the MAP arrival flow makes a transition with generation of a new customer. The intensities of such transition are defined by the entries of the matrix D_1. Arrival of a customer does not change the regime of the system operation (component r_t) and the number of broken servers (component k_t). Therefore, the matrix Q^+ has the form $Q^+ = \mathrm{diag}\{I_{M_1} \otimes D_1, I_{M_1-M_2} \otimes D_1\} = I_{2M_1-M_2} \otimes D_1$. Note, that operation of Kronecker product of matrices ($\otimes$) is very useful for describing the intensity or the probability of simultaneous transition of several independent Markovian components.

The entries of the blocks $Q_{i,i-1}$, $i \geq 1$, define the intensities of the transition of the Markov chain ξ_t, $t \geq 0$, that lead to the decrease in the number of customers in the system by one. This can happen if 1) a customer completes service or 2) a customer leaves the system due to impatience. If the system is in the operable state, both the options 1) and 2) are possible. The intensities of service completions depend on the number of customers in the system and the number of broken servers and are defined by the corresponding entries of the matrix μN_i. The matrix N_i defines the number of busy servers for each possible number of broken servers. The intensities of customers abandonment also depend on the number of customers in the system and the number of broken servers and are defined by the corresponding entries of the matrix αA_i.

If the system in the quarantine regime, only option 2) is possible. In this case, all i customers stay in the buffer, therefore, the intensity of a customer abandonment is $i\alpha$. A customer departure from the system does not change the regime of the system operation (component r_t), the number of broken servers (component k_t), and the state of the MAP underlying process (component w_t). Therefore, the matrix $Q_{i,i-1}$ has the form $Q_{i,i-1} = \mathrm{diag}\{(\mu N_i + \alpha A_i) \otimes I_{\bar{W}}, i\alpha I_{(M_1-M_2)\bar{W}}\}$, $i \geq 1$.

The non-diagonal entries of the blocks $Q_{i,i}$, $i \geq 0$, define the intensities of the transition of the chain $\xi_t, t \geq 0$, that do not lead to the change of the number of customers in the system. The entries of the matrix $Q_{i,i}^{(0,0)}$ define such intensities when the system is in the operable regime. The events that lead to such transitions are the following:

1) A breakdown arrives when the number of broken server is less then $M_1 - 1$. In this case, the number of broken servers increases by one and the component w_t does not change. The intensities of these transitions are given by the entries of the matrix $GE^+ \otimes I_{\bar{W}}$.

2) The underlying process of the MAP transits to another state without generation of a customer. In this case, the number of broken servers does not change. Thus, the intensities of these transitions are defined as the non-diagonal entries of the matrix $I_{M_1} \otimes D_0$.

3) The repairing of one broken server is finished. In this case, the number of broken servers decreases by one and the component w_t does not change. The intensities of these transitions are given by the entries of the matrix $\beta C_1 E_1^- \otimes I_{\bar{W}}$.

The entries of the matrix $Q_{i,i}^{(0,1)}$ define the intensities of the transitions that do not lead to the change the number of customers in the system, but lead to the transition to the quarantine regime. These transitions are possible only when the system is in the operable regime, the number of broken servers is $M_1 - 1$ and a new breakdown arrives. The intensities of these transitions are defined by the entries of the matrix $Q_{i,i}^{(0,1)} = G\hat{I} \otimes I_{\bar{W}}$.

The entries of the matrix $Q_{i,i}^{(1,1)}$ define the intensities of the transitions that also do not lead to the termination of the use of the quarantine regime and to the change the number of customers in the system. The events that lead to such transitions are the following: the repair completion of a broken server when the number of broken servers is greater than $M_2 + 1$ (the intensities are defined by the entries of the matrix $\beta C_2 E_2^- \otimes I_{\bar{W}}$) and the transition of the underlying MAP process without generation of a customer (the intensities are defined by the non-diagonal entries of the matrix $I_{M_1-M_2} \otimes D_0$).

The entries of the matrix $Q_{i,i}^{(1,0)}$ define the intensities of the transitions that do not lead to the change the number of customers in the system, but lead to the transition from the quarantine regime to the operable regime. These transitions are possible only if the number of broken servers is $M_2 + 1$ and one server is repaired. The intensities of these transitions are defined by the entries of the matrix $Q_{i,i}^{(1,0)} = \beta\tilde{I} \otimes I_{\bar{W}}$.

The diagonal entries of the matrix $Q_{i,i}$ are negative and the modulus of each entry defines the total intensity of leaving the corresponding state of the chain $\xi_t, t \geq 0$. The modules of the diagonal entries of the matrix $I_{M_1} \otimes D_0 - (\alpha A_i + \beta C_1 + G + \mu N_i) \otimes I_{\bar{W}}$ define the intensities of leaving the corresponding state of the Markov chain when the system is in operable mode. Thus, the matrix $Q_{i,i}^{(0,0)}$ is defined by the formula $Q_{i,i}^{(0,0)} = I_{M_1} \otimes D_0 - (\alpha A_i + \beta C_1(I_{M_1} - E_1^-) + G(I_{M_1} - E^+) + \mu N_i) \otimes I_{\bar{W}}$. If the system is in the quarantine regime, the intensities of leaving the corresponding state of the Markov chain ξ_t are defined by the modulus of the diagonal entries of the matrix $I_{M_1-M_2} \otimes D_0 - i\alpha I_{(M_1-M_2)\bar{W}} - \beta C_2 \otimes I_{\bar{W}}$. Thus, the matrix $Q_{i,i}^{(1,1)}$ has the form $Q_{i,i}^{(1,1)} = I_{M_1-M_2} \otimes D_0 - i\alpha I_{(M_1-M_2)\bar{W}} - \beta C_2(I_{M_1-M_2} - E_2^-) \otimes I_{\bar{W}}$.

Because the probability that the number of customers decreases or increases by more than one during the interval of an infinitesimal length is negligible, the blocks $Q_{i,j}$ are zero matrices for all i, j when $|i-j| > 1$. Therefore, the generator Q has the block-tridiagonal structure. Theorem 1 is proved.

Corollary 1. The Markov chain $\xi_t,\ t \geq 0$, belongs to the class of continuous-time asymptotically quasi-Toeplitz Markov chains ($AQTMC$), see (Klimenok and Dudin 2006).

Proof directly follows from comparison of the properties of the blocks $Q_{i,j}$ of the block-tridiagonal generator Q for large values of i with the required properties of the generator listed in the definition of the $AQTMC$.

As the customers staying in the buffer are impatient ($\alpha > 0$), based on the results for $AQTMC$ it can be shown that the stationary probabilities of the system states $p(i, r, k, w),\ i \geq 0,\ r = \overline{0, 1},\ k = \overline{0, M_1 - 1},\ w = \overline{0, W}$, exist for all possible values of the system parameters. Let us form the row vectors $\mathbf{p}_i$ of these probabilities enumerated in the lexicographic order of the components r, k, w. To this end, we sequentially form the row vectors

$$\mathbf{p}(i, 0, k) = (p(i, 0, k, 0), p(i, 0, k, 1), \ldots, p(i, 0, k, W)),$$

$$k = \overline{0, M_1 - 1},$$

$$\mathbf{p}(i, 0) = (\mathbf{p}(i, 0, 0), \mathbf{p}(i, 0, 1), \ldots, \mathbf{p}(i, 0, M_1 - 1)),$$

$$\mathbf{p}(i, 1, k) = (p(i, 1, k, 0), p(i, 1, k, 1), \ldots, p(i, 1, k, W)),$$

$$k = \overline{M_2 + 1, M_1},$$

$$\mathbf{p}(i, 1) = (\mathbf{p}(i, 1, M_2{+}1), \mathbf{p}(i, 1, M_2{+}2), \ldots, \mathbf{p}(i, 1, M_1)),$$

$$\mathbf{p}_i = (\mathbf{p}(i, 0), \mathbf{p}(i, 1)),\ i \geq 0.$$

It is well known that the probability vectors $\mathbf{p}_i,\ i \geq 0$, satisfy the following system of linear algebraic equations:

$$(\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_i, \ldots)Q = \mathbf{0},\ \ (\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_i, \ldots)\mathbf{e} = 1$$

where Q is the infinitesimal generator of the Markov chain $\xi_t,\ t \geq 0$, $\mathbf{0}$ is a zero row vector, and $\mathbf{e}$ denotes a unit column vector. Due to the existing dependence of the blocks $Q_{i,i-1}$ and $Q_{i,i}$ on i, the problem of solving this infinite system of linear algebraic equations for the components of the vectors $\mathbf{p}_i,\ i \geq 0$, is far of easy. We omit the details of derivations and just mention that, to solve this system, we used the numerically stable algorithm that takes into account that the matrix Q has a block-tridiagonal structure, see (Dudina et al. 2013).

PERFORMANCE MEASURES

As soon as the vectors $\mathbf{p}_i,\ i \geq 0$, have been calculated, we are able to find various performance measures of the system.

The average number L of customers in the system is computed by $L = \sum\limits_{i=1}^{\infty} i\mathbf{p}_i\mathbf{e}$.

Note that some formulas for performance measures contain the infinite sums. However, computation of these sums do not create difficulties. It is well known that if the Markov chain is ergodic the stationary probability vectors $\mathbf{p}_i$ converge in norm to a zero vector when i approaches infinity. Therefore, the computation of an sum may be terminated if the norm of the summand becomes less than a preassigned value ϵ.

The average number N_{serv} of busy servers is computed by $N_{serv} = \sum\limits_{i=1}^{\infty}\sum\limits_{k=0}^{M_1-1} \max\{\min\{i, N\} - k, 0\}\mathbf{p}(i, 0, k)\mathbf{e}$.

The average number N_{buffer} of customers in the buffer is computed by $N_{buffer} = \sum\limits_{i=1}^{\infty}(\sum\limits_{k=0}^{M_1-1} \max\{i-(N-k), 0\}\mathbf{p}(i, 0, k)\mathbf{e} + \sum\limits_{k=M_2+1}^{M_1} i\mathbf{p}(i, 1, k)\mathbf{e})$.

The average number N_{broken} of broken servers is computed by $N_{broken} = \sum\limits_{i=0}^{\infty}(\sum\limits_{k=1}^{M_1-1} k\mathbf{p}(i, 0, k)\mathbf{e} + \sum\limits_{k=M_2+1}^{M_1} k\mathbf{p}(i, 1, k)\mathbf{e})$.

The average intensity λ_{out} of flow of customers who receive service is computed by $\lambda_{out} = \mu N_{serv}$.

The probability P_{loss} that an arbitrary customer will be lost is computed by

$$P_{loss} = 1 - \frac{\lambda_{out}}{\lambda} = \frac{\alpha N_{buffer}}{\lambda}.$$

The average intensity γ_{quar} of the quarantine beginning can be found as

$$\gamma_{quar} = \gamma_{M_1-1}\sum_{i=0}^{\infty}\mathbf{p}(i, 0, M_1 - 1)\mathbf{e}.$$

The average duration T_{quar} of the quarantine is computed by $T_{quar} = \beta^{-1}\sum\limits_{k=M_2+1}^{M_1}\frac{1}{k}$.

The probability P_{quar} that at an arbitrary moment system is in the quarantine regime is calculated as

$$P_{quar} = \sum_{i=0}^{\infty}\mathbf{p}(i, 1)\mathbf{e} = \gamma_{quar}T_{quar}.$$

NUMERICAL EXAMPLE

The goal of the numerical example is to demonstrate the feasibility of the proposed results and illustrate possible way for optimizing the system operation by means of the optimal choosing the thresholds M_1 and M_2 for starting and finishing the regime of quarantine, correspondingly.

Let consider the system with $N = 20$ servers. We assume that the MAP arrival flow of customers is defined by the matrices

$$D_0 = \begin{pmatrix} -12.168 & 0 \\ 0 & -0.395 \end{pmatrix}, D_1 = \begin{pmatrix} 12.087 & 0.081 \\ 0.22 & 0.175 \end{pmatrix}.$$

This arrival flow has the average intensity of customers arrival $\lambda = 9$, the coefficient of correlation of successive inter-arrival intervals is $c_{cor} = 0.2$ and the coefficient of variation of such intervals is $c_{var} = 12.34$.

The rest of the system parameters are chosen as follows: the intensities $\gamma_k,\ k = \overline{0, M_1 - 1}$, of breakdowns arrival are defined as $\gamma_0 = 0.00001$, $\gamma_1 = 0.0007$, $\gamma_2 =$

0.0015, $\gamma_l = \gamma_{l-1} + 0.0015l$, $l = \overline{3, M_1 - 1}$; the service intensity $\mu = 0.8$; the intensity of customers impatience $\alpha = 0.2$; the recovering intensity $\beta = 0.0005$.

Let us vary the threshold M_1 over the interval $[1, N]$ and the threshold M_2 over the interval $[0, M_1 - 1]$.

Figure 2 illustrates the dependence of the average number N_{broken} of broken servers on the thresholds M_1 and M_2.

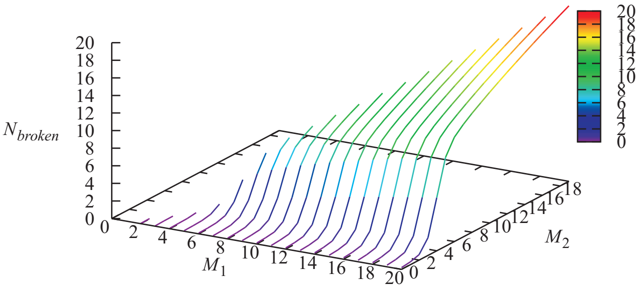


Figure 2: Dependence of the average number N_{broken} of broken servers on the thresholds M_1 and M_2

It is evidently seen in Figure 2 that the grows of the threshold M_1 (more late switching to the quarantine regime) causes the significant increase of the average number N_{broken} of broken servers. This increase becomes more essential when the threshold M_2 increases (the work in the quarantine regime becomes shorter).

Figure 3 illustrates the dependence of the average intensity γ_{quar} of the quarantine beginning on the thresholds M_1 and M_2.

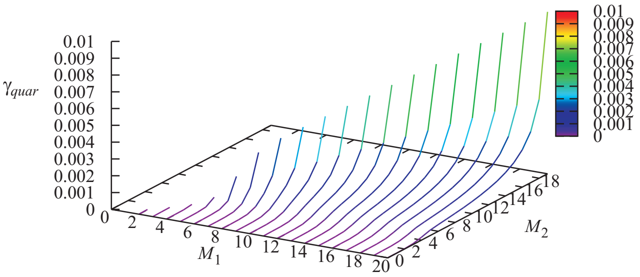


Figure 3: Dependence of the average intensity γ_{quar} on the thresholds M_1 and M_2

The intensity γ_{quar} of the quarantine beginning essentially increases when M_2 grows (what causes short duration of each quarantine while a frequent switching the system to the quarantine regime).

Figure 4 illustrates the dependence of the probability P_{quar} that an arbitrary moment system is in the quarantine regime on the thresholds M_1 and M_2.

The essential increase of P_{quar} when M_2 grows has the same explanation as the one given to Figure 3. More flat shape of the dependence of P_{quar} on the thresholds M_1 and M_2 in comparison to the shape of the dependence of γ_{quar} for large values of M_1 and M_2 is easily explained by the fact that the frequent switching to the quarantine regime (illustrated by Figure 3) is accompanied by the shorter duration of each period of using the quarantine regime.

Figure 5 illustrates the dependence of the loss probability P_{loss} on the thresholds M_1 and M_2.

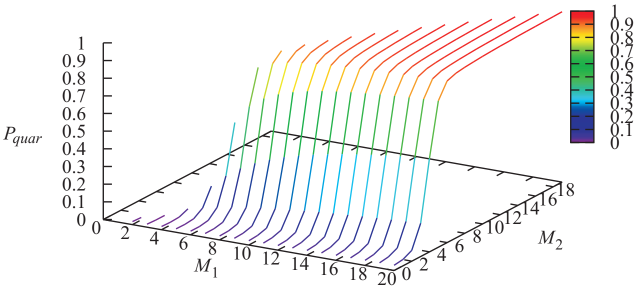


Figure 4: Dependence of the probability P^{quar} on the thresholds M_1 and M_2

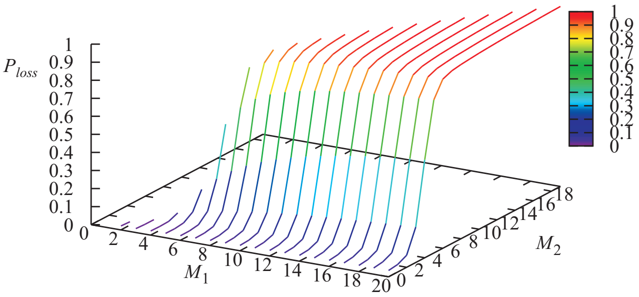


Figure 5: Dependence of the loss probability P_{loss} on the thresholds M_1 and M_2

The shape of the surface at Figure 5 well agrees with the shape of the surface at Figure 4. The large probability P_{quar} implies that during an essential share of time service is not provided. This causes large average queue length N_{buffer} and large value of the loss probability P_{loss} which is proportional to N_{buffer}. The minimal value of the loss probability $P_{loss} = 0.0182787$ is reached for $M_1 = 2$ and $M_2 = 1$.

Depending on application area of the considered queueing model, the quality of the system operation can be characterised in terms of different cost criteria. E.g., let us consider the following cost criterion:

$$E(M_1, M_2) = a\lambda_{out} - b\gamma_{quar} - cN_{broken}$$

where a is the profit obtained by the system from service of one customer, b is the charge paid by the system for the transition to the quarantine regime and c is the charge paid by the system for repair of one broken server.

Figure 6 illustrates the dependence of this cost criterion on the thresholds M_1 and M_2 when the cost coefficients are fixed as follows: $a = 1$, $b = 100$ and $c = 2$.

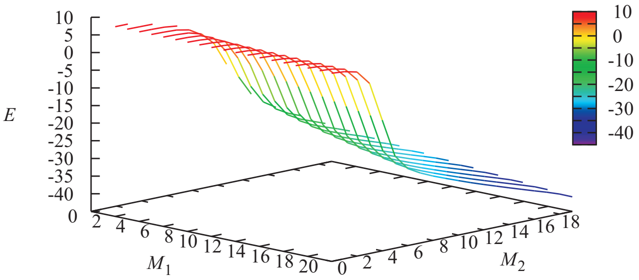


Figure 6: Dependence of the value $E(M_1, M_2)$ of the cost criterion on the thresholds M_1 and M_2

The optimal (maximal) value of the cost criterion is equal to $E^*(M_1, M_2) = 8.7419$ and is reached when $M_1 = 2$ and $M_2 = 0$, i.e., the quarantine starts when the number of broken servers reaches the value $M_2 = 2$ and the system resumes the work when all broken servers will be repaired.

CONCLUSIONS

The novel unreliable queueing model, in which the intensity of servers breakdowns increases when the number of broken servers grows, is analysed. To smooth the negative effect of the increase in this intensity, it is proposed to switch the system to quarantine regime during which new breakdowns cannot occur. Switching to quarantine regime is performed when the number of broken servers reaches the threshold value M_1. Switch back to operable regime is performed when the number of broken servers drops below the threshold value M_2. The problem of computation of performance measures of the model as function of the thresholds (M_1, M_2) is solved. This allows us to make managerial decisions providing the best quality of the system operation. Results may be extended to the systems with retrials and more complicated arrival flows.

ACKNOWLEDGMENT

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008).

REFERENCES

Al-Begain, K., Dudin, A., Klimenok, V., Dudin, S. 2012. "Generalised survivability analysis of systems with propagated failures", *Computers and Mathematics with Applications*, 64, 3777-3791.

Chakravarthy, S. 2001. "The batch Markovian arrival process: a review and future work", *Advances in Probability Theory and Stochastic Processes*, Notable Publications Inc., New Jersey, 21-29.

Dudin, A., Jacob, V., Krishnamoorthy, A. 2015. "A multi-server queueing system with service interruption, partial protection and repetition of service", *Annals of Operations Research*, 233, 101-121.

Dudina, O., Kim, Ch., Dudin, S. 2013. "Retrial queueing system with Markovian arrival flow and phase type service time distribution", *Computers and Industrial Engineering*, 66, 360-373.

Graham, A. 1981. "Kronecker products and matrix calculus with applications", Ellis Horwood, Cichester.

Klimenok, V., Orlovsky, D., Kim, Ch. 2008. "The $BMAP/PH/N$ retrial queue with Markovian flow of breakdowns", *European Journal of Operational Research*, 189, 1057-1072.

Klimenok, V.I. and Dudin, A.N. 2006. "Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory", *Queueing Systems*, 54, 245-259.

Lucantoni, D. 1991. "New results on the single server queue with a batch Markovian arrival process", *Communication in Statistics-Stochastic Models*, 7, 1-46.

Mitrani, I. and Avi-Itzhak, B. 1968. "A many-server queue with server interruptions", *Operations Research*, 163, 28-638.

Neuts, M. and Lucantoni, D. 1979. "A Markovian queue with N servers subject to breakdowns and repair", *Management Science*, 25, 49-861.

AUTHOR BIOGRAPHIES

ALEXANDER DUDIN has got PhD degree in Probability Theory and Mathematical Statistics in 1982 from Vilnius University and Doctor of Science degree in 1992 from Tomsk University. He is Head of Laboratory of Applied Probabilistic Analysis in Belarusian State University, Professor of the Probability Theory and Mathematical Statistics Department. He works also part time at the Peoples' Friendship University of Russia. He is author of 350 publications including more than 80 papers in top level Journals. Field of scientific interests are: Random Processes in Queueing Systems and Applications of Queueing Theory to Telecommunication. His email address is dudin@bsu.by.

SERGEI DUDIN was graduated from Belarusian State University in 2007. In 2010, he got PhD degree in Belarusian State University in System Analysis, Control and Information Processing and works currently as leading scientific researcher of Research Laboratory of Applied Probabilistic Analysis in Belarusian State University. He works also part time at the Peoples' Friendship University of Russia. His main fields of interests are queueing systems with session arrivals and controlled tandem models. His email address is dudins@bsu.by.

OLGA DUDINA was graduated from Belarusian State University in 2007. In 2010, she got PhD degree in Belarusian State University in Probability Theory and Mathematical Statistics and works currently as leading scientific researcher of Research Laboratory of Applied Probabilistic Analysis in Belarusian State University. She works also part time at the Peoples' Friendship University of Russia. Her main fields of interests are tandem queueing models with correlated arrival flows, non-markovian queueing systems. Her email address is dudina@bsu.by.

KONSTANTIN SAMOUYLOV received his Ph.D. from the Moscow State University and a Doctor of Sciences degree from the Moscow Technical University of Communications and Informatics. During 1985-1996 he held several positions at the Faculty of Science of the Peoples' Friendship University of Russia where he became a head of Telecommunication System Department in 1996. Since 2014 he is a head of the Department of Applied Informatics and Probability Theory. His current research interests are probability theory and theory of queuing systems, performance analysis of 4G/5G networks, teletraffic of triple play networks, and signaling networks planning. He is the author of more than 100 scientific and technical papers and three books. His email address is ksam@sci.pfu.edu.ru.

ASYMPTOTIC ANALYSIS OF MARKOVIAN RETRIAL QUEUE WITH TWO-WAY COMMUNICATION UNDER LOW RATE OF RETRIALS CONDITION [1]

Nazarov A.
Peoples' Friendship University of
Russia (RUDN University)
6 Miklukho-Maklaya St, Moscow,
117198, Russian Federation
National Research
Tomsk State University
36 Lenina ave., 634050, Russian
Federation
nazarov.tsu@gmail.com

Paul S.
National Research
Tomsk State University
36 Lenina ave., 634050, Russian
Federation
paulsv82@mail.ru

Gudkova I.
Peoples' Friendship University of
Russia (RUDN University)
6 Miklukho-Maklaya St, Moscow,
117198, Russian Federation
gudkova_ia@pfur.ru

KEYWORDS

Retrial queueing system with two-way communication, incoming and outgoing calls, asymptotic analysis method, Gaussian approximation.

ABSTRACT

In this paper we are reviewing the retrial queue with two-way communication and Poisson arrival process. If the server free, incoming call occupies it. The call that finds the server being busy joins an orbit and retries to enter the server after some exponentially distributed time. If the server is idle, it causes the outgoing call from the outside. The outgoing call can find server free, then it starts making an outgoing call in an exponentially distributed time. If the outgoing call finds the server occupied, then it is lost. To research the system in question we have derived first and second order asymptotics of a number of calls in the orbit in an asymptotic condition of a low rate of retrials. Based on found asymptotics we have built the Gaussian approximation of a number of calls in the orbit.

INTRODUCTION

Recently a lot of attention is being paid to the research of the retrial queues such as mathematical models of real call center systems, telecommunication networks, computer networks, economical systems (Artalejo and Gomez-Corral 2008). These systems are characterized by the fact that if the clients (calls, phone calls, messages etc.) couldn't be served immediately they have to enter the virtual orbit where they wait out some delay before they could access the server for service again (Flajolet and Sedgewick 2009).

As a rule, the ones that are considered are the retrial queues in which arriving calls are either served immediately or join the orbit where they are wait out a random delay before accessing the server again. Recently, however, server is more likely to have the ability to make an outgoing call. The example of that could be the common cellphone that has function of both incoming and outgoing calls. In different call centers operators could receive arriving calls but as soon as they have free time and are in standby mode they could make outgoing calls to advertise, promote and sell packages and services of the centre.

Falin (Falin 1979) derives integral formulas for partial generating functions and some explicit expressions for characteristics of the M|G|1|1 retrial queues with outgoing calls. Choi et al. (Choi et al. 1995) extends Falin's model for the M/G/1/K retrial queues. Artalejo and Resing (Artalejo and Resing 2010) have derived first moments for characteristics of the M/G/1/1 retrial queues, in which the times of serving arriving and outgoing calls are different.

Martin and Artalejo (Martin and Artalejo 1995) are considering M|G|1|1 retrial queues with outgoing calls in which calls from an orbit access the server after an exponentially distributed delay in the order of arrival.

Artalejo and Phung-Duc (Artalejo and Tuan 2012) are considering M|M|1|1 retrial queues with outgoing calls and a different service time for incoming and outgoing calls. In their paper the authors have found an explicit solution for two-dimensional probability distribution of a server state and a number of calls in an orbit. Likewise, the factorial moments are found, based on which the proposed numerical and recurrent algorithms may be applied.

In this paper the main method of research is the asymptotic analysis method which allows to find in M|M|1|1 retrial queue with two-way communication type of limit distribution of a number of calls in the orbit in an asymptotic condition of a low rate of retrials and to show that limit distribution is Gaussian.

[1] The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008)

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

This result is achieved by using the original asymptotic analysis method without needing to find the nonlimiting distribution. Furthermore, the discrete distribution is constructed which approximates discrete distribution of a number of calls in an orbit. This distribution will be addressed as Gaussian approximation. Research of retrial queueing system under the asymptotic condition that the retrial rate is extremely low is stated in the following papers (Nazarov and Chernikova 2014) (Nazarov and Izmailova 2016).

Furthermore, we have defined conditions of applicability of obtained approximation according to system defining parameters.

The remainder of the paper is presented as follows. In Section "Mathematical Model", we describe the model in detail and preliminaries for later asymptotic analysis. In Sections "Ferst order asymptotic" and "Second order acymptotic", we present our main contribution for the model with Poisson input. In Section "Approximation accuracy $P^{(2)}(i)$ and its application area" we have defined the conditions of applicability of the obtained approximation depending on values of system-defining parameters. Section "Conclusions" is devoted to concluding remark and future work.

MATHEMATICAL MODEL

Let's consider retrial queue (Figure 1) with Poisson arrival process of incoming calls with rate λ.

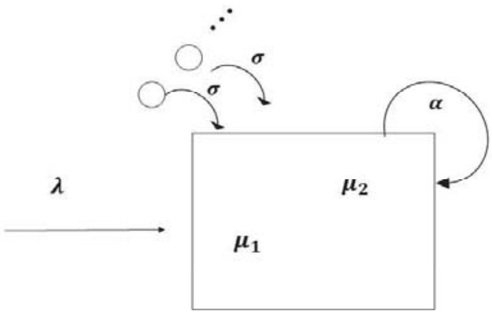


Figure 1: Retrial queue with two-way communication

The incoming call finds the server and goes into service for an exponentially distributed time with rate $\mu_{1.}$ If upon entering the system the call finds the server being busy the call immediately joins the orbit, where it stays during a random time distributed exponentially with rate σ.

If the server is idle (empty) it starts making outgoing calls from the outside with rate α. If the outgoing call finds the server free the call goes into service for an exponentially distributed time with rate $\mu_{2.}$ If upon entering the system the outgoing call finds the server being busy the call is lost and is not considered in the future. Let's denote:

$i(t)$ – number of calls in the orbit at the time t,

$n(t)$ – server state: 0 – server is free, 1 – server is busy serving an incoming call, 2 – server is busy serving an outgoing call.

Let's consider two-dimensional Markovian process $\{i(t), n(t)\}$ for probability distribution

$$P\{i(t) = i, n(t) = n\}= P_n(i, t)$$

setting up system of Kolmogorov equations

$$-(\lambda + i\sigma + \alpha)P_0(i) + \mu_1 P_1(i) + \mu_2 P_2(i) = 0,$$
$$-(\lambda+\mu_1)P_1(i) + \lambda[P_1(i-1) + P_0(i)] +$$
$$+(i+1)\sigma P_0(i+1) = 0\,,$$
$$-(\lambda+\mu_2)P_2(i) + P_0(i)\alpha + P_2(i-1)\lambda = 0\,. \qquad (1)$$

Introducing partial characteristic functions (Nazarov and Paul 2016), denoting $j = \sqrt{-1}$,

$$H_n(u) = \sum_{i=0}^{\infty} e^{jui} P_n(i)\,.$$

Rewriting system (1) in the following form

$$-(\lambda+\alpha)H_0(u) + j\sigma\frac{dH_0(u)}{du} +$$
$$+\mu_1 H_1(u) + \mu_2 H_2(u) = 0,$$
$$\left[\lambda\left(e^{ju}-1\right)-\mu_1\right]H_1(u) + \lambda H_0(u) - j\sigma e^{-ju}\frac{dH_0(u)}{du} = 0,$$
$$\left[\lambda\left(e^{ju}-1\right)-\mu_2\right]H_2(u) + \alpha H_0(u) = 0\,. \qquad (2)$$

Characteristic function $H(u)$ of a number of incoming calls in an orbit and server states probability distribution r_n are relatively easy expressed through partial characteristic functions $H_n(u)$ by the following equations

$$H(u) = Me^{jui(t)} = H_0(u) + H_1(u) + H_2(u)\,,$$
$$r_n = H_n(0),\; n = 0, 1, 2.$$

The task is put to find these characteristics of retrial queue with two-way communication. The main content of this paper is the solution of system (2) by using asymptotic analysis method in limit condition of a low rate of retrials, when σ → 0.

This is due to the fact that for the more complicated queues with an incoming MMPP, the equation system similar to (2) is analytically unsolvable, but a solution by using asymptotic analysis method is allowed.

Application of asymptotic results in prelimit situation is causing the necessity of specifying the area of its applicability, which is obtainable only through comparison of asymptotic and prelimit characteristics and that is relatively easy implemented for the retrial queue in question. For more complex systems prelimit characteristics are usually defined by results of imitational modeling or by using pretty complicated numerical algorithms. The asymptotic analysis method suggested below is implemented by sequential determination of first and second order asymptotics.

FIRST ORDER ASYMPTOTIC

We introduce the following notations

$$\sigma = \varepsilon,\quad u = \varepsilon w,\quad H_n(u) = F_n(w, \varepsilon),$$

then we will get this system

$$-(\lambda+\alpha)F_0(w,\varepsilon)+j\varepsilon\frac{\partial F_0(w,\varepsilon)}{\partial w}+$$
$$+\mu_1 F_1(w,\varepsilon)+\mu_2 F_2(w,\varepsilon)=0,$$
$$\left\{\lambda\left(e^{jw\varepsilon}-1\right)-\mu_1\right\}F_1(w,\varepsilon)+$$
$$+\lambda F_0(w,\varepsilon)-j\varepsilon e^{-jw\varepsilon}\frac{\partial F_0(w,\varepsilon)}{\partial w}=0,$$
$$\left\{\lambda\left(e^{jw\varepsilon}-1\right)-\mu_1\right\}F_2(w,\varepsilon)+\alpha F_0(w,\varepsilon)=0. \qquad (3)$$

Theorem 1. (First order asymptotic) Suppose *i*(*t*) is a number of calls in an orbit of stationary M|M|1 retrial queue with two-way communication, then the following equation is true

$$\lim_{\sigma\to 0} Me^{jw\sigma i(t)}=e^{jw\kappa_1},$$

where the parameter κ_1 is defined by the following

$$\kappa_1=\frac{\lambda}{\mu_2}\cdot\frac{\lambda\mu_2+\alpha\mu_1}{(\mu_1-\lambda)}.$$

Proof. Consider ε → 0, then we will get

$$-(\lambda+\alpha)F_0(w)+\mu_1F_1(w)+\mu_2F_2(w)=0,$$
$$-\mu_1F_1(w)+\lambda F_0(w)=0,$$
$$-\mu_2F_2(w)+\alpha F_0(w)=0, \qquad (4)$$

by denoting

$$\lim_{\varepsilon\to 0}F_n(w,\varepsilon)=F_n(w).$$

We will look for solution of the last system in form of

$$F_n(w)=\Phi(w)r_n, \qquad (5)$$

where r_n is the scalar server state probability distribution, and the function Φ(*w*) is defined in the following form

$$\Phi(w)=\exp\{jw\kappa_1\}.$$

Then, if we look at the system, we could see that the first equation is a sum of the second and the third in the system (4), and whilst keeping that in mind, let's review the normalization condition for stationary server state probability distribution

$$r_0+r_1+r_2=1,$$

we have the system

$$\begin{cases}\mu_1r_1=(\lambda+\kappa_1)r_0,\\ \mu_2r_2=\alpha r_0,\\ r_0+r_1+r_2=1.\end{cases} \qquad (6)$$

Value of the parameter κ_1 will be defined below. By summing equations of the system (3) we will get the following equation

$$F_1(w,\varepsilon)\left(e^{j\varepsilon w}-1\right)\lambda+F_2(w,\varepsilon)\left(e^{j\varepsilon w}-1\right)\lambda+$$
$$+je^{-j\varepsilon w}\left(e^{j\varepsilon w}-1\right)F_0'(w,\varepsilon)=o(\varepsilon), \qquad (7)$$

in which we will execute the limit transition while ε→0. Then we could write down this equation

$$F_1(w)\lambda+F_2(w)\lambda-jF_0'(w,\varepsilon)=0.$$

Let's substitute the product (5) in the obtained equation

$$\Phi_0'(w)=j\Phi(w)\frac{[r_1+r_2]\lambda}{r_0},$$

Then

$$\Phi(w)=\exp j\frac{[r_1+r_2]\lambda}{r_0}w.$$

Let's denote

$$\kappa_1=\frac{[r_1+r_2]\lambda}{r_0}. \qquad (8)$$

By taking into consideration the normalization condition for server state probability distribution and solving the system (6) alongside with the system (8) we will get the values of probabilities r_n and the value of parameter κ_1.

$$\begin{cases}r_1=\dfrac{\lambda}{\mu_1},\\ r_2=\dfrac{\alpha}{\mu_2}r_0,\\ r_0+\dfrac{\lambda}{\mu_1}+\dfrac{\alpha}{\mu_2}r_0=1.\end{cases}$$

will have this system in the end

$$\begin{cases}r_1=\dfrac{\lambda}{\mu_1},\\ r_2=\dfrac{\alpha}{\mu_1}\cdot\dfrac{\mu_1-\lambda}{\mu_2+\alpha},\\ r_0=\dfrac{\mu_1-\lambda}{\mu_1}\cdot\dfrac{\mu_2}{\mu_2+\alpha}.\end{cases} \qquad (9)$$

Parameter κ_1 is defined by equality

$$\kappa_1=\frac{[r_1+r_2]\lambda}{r_0}=\frac{\lambda}{\mu_2}\cdot\frac{\lambda\mu_2+\alpha\mu_1}{(\mu_1-\lambda)}.$$

First order asymptotic i.e. the proven theorem, only defines the mean asymptotic value κ_1/σ of a number of calls in an orbit in prelimit situation of nonzero values of σ. For more detailed research of a number *i*(*t*) of calls in an orbit let's consider the second order asymptotic.

SECOND ORDER ASYMPTOTIC

Let's substitute the following in the system (2)

$$H_n(u)=\exp\left(j\frac{u}{\sigma}\kappa_1\right)H_n^{(2)}(u),$$

we will get the system

$$-H_0^{(2)}(u)(\lambda+\alpha+\kappa_1)+j\sigma\frac{dH_0^{(2)}(u)}{du}+$$
$$+\mu_1H_1^{(2)}(u)+\mu_2H_2^{(2)}(u)=0,$$
$$H_1^{(2)}(u)\left(\left(e^{ju}-1\right)\lambda-\mu_1\right)+H_0^{(2)}(u)\lambda+$$
$$+\kappa_1e^{-ju}H_0^{(2)}(u)-j\sigma e^{-ju}\frac{dH_0^{(2)}(u)}{du}=0,$$
$$H_2^{(2)}(u)\left(\left(e^{ju}-1\right)\lambda-\mu_2\right)+\alpha H_0^{(2)}(u)=0. \qquad (10)$$

Let's make substitutions as shown below

$$\sigma=\varepsilon^2,\quad u=\varepsilon w,\quad H_n^{(2)}(u)=F_n^{(2)}(w,\varepsilon),$$

then we will get the system

$$-F_0^{(2)}(w,\varepsilon)(\lambda+\alpha+\kappa_1)+j\sigma\frac{\partial F_0^{(2)}(w,\varepsilon)}{\partial w}+$$

$$+\mu_1 F_1^{(2)}(w,\varepsilon)+\mu_2 F_2^{(2)}(w,\varepsilon)=0,$$
$$F_1^{(2)}(w,\varepsilon)\left(\left(e^{jw\varepsilon}-1\right)\lambda-\mu_1\right)+F_0^{(2)}(w,\varepsilon)\lambda+$$
$$+\kappa_1 e^{-jw\varepsilon}F_0^{(2)}(w,\varepsilon)-j\sigma e^{-jw\varepsilon}\frac{\partial F_0^{(2)}(w,\varepsilon)}{\partial w}=0,$$
$$F_2^{(2)}(w,\varepsilon)\left(\left(e^{jw\varepsilon}-1\right)\lambda-\mu_2\right)+\alpha F_0^{(2)}(w,\varepsilon)=0. \quad (11)$$

Theorem 2. (Second order asymptotic) In the context of Theorem 1 the following equation is true

$$\lim_{\sigma\to 0} Me^{jw\sqrt{\sigma}\left(i(t)-\frac{\kappa_1}{\sigma}\right)}=e^{\frac{(jw)^2}{2}\kappa_2},$$

where parameter κ_2 is defined by the following expression

$$\kappa_2=\frac{\lambda^3\mu_2^2+\lambda^3\alpha\mu_2+\alpha\mu_1^2\lambda^2-\lambda^3\alpha\mu_1}{\mu_2^2(\mu_1-\lambda)^2}+\kappa_1.$$

Proof. Let's substitute the following expansion into the system (5)

$$F_n^{(2)}(w,\varepsilon)=\Phi_2(w)\{r_n+j\varepsilon wf_n\}+o\left(\varepsilon^2\right), \qquad (12)$$

then we will get

$$-\Phi_2(w)\{r_0+j\varepsilon wf_0\}(\lambda+\alpha+\kappa_1)+$$
$$+j\varepsilon\frac{d\Phi_2(w)}{dw}r_0+\mu_1\Phi_2(w)\{r_1+j\varepsilon wf_1\}+$$
$$+\mu_2\Phi_2(w)\{r_2+j\varepsilon wf_2\}=o\left(\varepsilon^2\right),$$
$$\Phi_2(w)\{r_1+j\varepsilon wf_1\}(j\varepsilon w\lambda-\mu_1)+$$
$$+\Phi_2(w)\{r_0+j\varepsilon wf_0\}(\lambda+\kappa_1(1-j\varepsilon w))-$$
$$-j\varepsilon\frac{d\Phi_2(w)}{dw}r_0=o\left(\varepsilon^2\right),$$
$$\Phi_2(w)\{r_2+j\varepsilon wf_2\}(j\varepsilon w\lambda-\mu_2)+$$
$$+\alpha\Phi_2(w)\{r_0+j\varepsilon wf_0\}=o\left(\varepsilon^2\right).$$

Transforming the last system

$$\Phi_2(w)\{r_0(-\lambda-\alpha-\kappa_1)+\mu_1 r_1+\mu_2 r_2+$$
$$+j\varepsilon w[f_0(-\lambda-\alpha-\kappa_1)+\mu_1 f_1+\mu_2 f_2]\}+$$
$$+j\varepsilon\frac{d\Phi_2(w)}{dw}r_0=o\left(\varepsilon^2\right),$$
$$\Phi_2(w)\{\mu r_1+r_0(\lambda+\kappa_1)+$$
$$+j\varepsilon w[-\mu f_1+r_1\lambda+f_0(\lambda+\kappa_1)-\kappa_1 r_0]-$$
$$-j\varepsilon\frac{d\Phi_2(w)}{dw}r_0=o\left(\varepsilon^2\right),$$
$$j\varepsilon w[-\mu_2 f_2+r_2\lambda+\alpha f_0]=o\left(\varepsilon^2\right).$$

Then

$$j\varepsilon w\Phi_2(w)[f_0(-\lambda-\alpha-\kappa_1)+\mu_1 f_1+\mu_2 f_2]+$$
$$+j\varepsilon\frac{d\Phi_2(w)}{dw}r_0=o\left(\varepsilon^2\right),$$
$$j\varepsilon w\Phi_2(w)[-\mu_1 f_1+r_1\lambda+\mathbf{f_0}(\lambda+\kappa_1)-\kappa_1 r_0]-$$
$$-j\varepsilon\frac{d\Phi_2(w)}{dw}r_0=o\left(\varepsilon^2\right)$$
$$j\varepsilon w[-\mu_2 f_2+r_2\lambda+\alpha f_0]=o\left(\varepsilon^2\right).$$

Let's divide equation of the system by ε, we will get

$$\Phi_2(w)[f_0(-\lambda-\alpha-\kappa_1)+\mu_1 f_1+\mu_2 f_2]+\frac{d\Phi_2(w)}{wdw}r_0=0,$$
$$\Phi_2(w)[-\mu_1 f_1+r_1\lambda+\mathbf{f_0}(\lambda+\kappa_1)-\kappa_1 r_0]-$$
$$-\frac{d\Phi_2(w)}{wdw}r_0=0$$
$$-\mu_2 f_2+r_2\lambda+\alpha f_0=0.$$

Take note that the scalar function $\Phi_2(w)$ is defined in the following form

$$\Phi_2(w)=\exp\left\{\frac{(jw)^2}{2}\kappa_2\right\},$$

then

$$[f_0(-\lambda-\alpha-\kappa_1)+\mu_1 f_1+\mu_2 f_2]-\kappa_2 r_0=0,$$
$$[-\mu_1 f_1+r_1\lambda+\mathbf{f_0}(\lambda+\kappa_1)-\kappa_1 r_0]+\kappa_2 r_0=0,$$
$$-\mu_2 f_2+r_2\lambda+\alpha f_0=0.$$

We have

$$f_0(-\lambda-\alpha-\kappa_1)+\mu_1 f_1+\mu_2 f_2=\kappa_2 r_0,$$
$$-\mu_1 f_1+\mathbf{f_0}(\lambda+\kappa_1)=(\kappa_1-\kappa_2)r_0-r_1\lambda,$$
$$-\mu_2 f_2+\alpha f_0=-r_2\lambda.$$

By summing equations of the system (11) we have

$$\kappa_1 e^{-j\varepsilon w}\left(1-e^{j\varepsilon w}\right)F_0^{(2)}(w,\varepsilon)-$$
$$-j\varepsilon e^{-j\varepsilon w}\left(1-e^{j\varepsilon w}\right)\frac{\partial F_0^{(2)}(w,\varepsilon)}{\partial w}-$$
$$-\left(1-e^{j\varepsilon w}\right)\left[F_1^{(2)}(w,\varepsilon)+F_2^{(2)}(w,\varepsilon)\right]\lambda=o\left(\varepsilon^2\right).$$

Transforming the last system

$$\kappa_1 e^{-j\varepsilon w}F_0^{(2)}(w,\varepsilon)-j\varepsilon e^{-j\varepsilon w}\frac{\partial F_0^{(2)}(w,\varepsilon)}{\partial w}-$$
$$-\left[F_1^{(2)}(w,\varepsilon)+F_2^{(2)}(w,\varepsilon)\right]\lambda=o\left(\varepsilon^2\right).$$

Let's substitute the expansion (12), we will get the following equation

$$\kappa_1(1-j\varepsilon w)\Phi_2(w)\{r_0+j\varepsilon wf_0\}-$$
$$-j\varepsilon(1-j\varepsilon w)\left[\frac{d\Phi_2(w)}{dw}\{r_0+j\varepsilon wf_0\}+\Phi_2(w)j\varepsilon f_0\right]-$$
$$-\Phi_2(w)[r_1+j\varepsilon wf_1+r_2+j\varepsilon wf_2]\lambda=o\left(\varepsilon^2\right).$$

By applying previously obtained equations we have

$$j\varepsilon w\Phi_2(w)[\kappa_1(f_0-r_0)-(f_1+f_2)\lambda]-$$
$$-j\varepsilon\frac{d\Phi_2(w)}{dw}r_0=o\left(\varepsilon^2\right).$$

Consider ε→0

$$\frac{d\Phi_2(w)}{dw}r_0=w\Phi_2(w)[\kappa_1(f_0-r_0)-(f_1+f_2)\lambda],$$

then

$$\frac{d\Phi_2(w)}{\Phi_2(w)}=\frac{[\kappa_1(f_0-r_0)-(f_1+f_2)\lambda]}{r_0}wdw.$$

Let's denote

$$\frac{\kappa_1 f_0-(f_1+f_2)\lambda}{r_0}-\kappa_1=$$
$$=-\left(\frac{(f_1+f_2)\lambda-\kappa_1 f_0}{r_0}+\kappa_1\right)=j^2\kappa_2,$$

where

$$\kappa_2 = \frac{(f_1 + f_2)\lambda - \kappa_1 f_0}{r_0} + \kappa_1 .$$

Then, considering that $\Phi_2(0) = 1$ we have

$$\Phi_2(w) = \exp\left\{\frac{(jw)^2}{2}\kappa_2\right\}.$$

Let's find κ_2, by expressing

$$f_1 = f_{\mathbf{0}}\frac{(\lambda + \kappa_1)}{\mu_1} - \frac{(\kappa_1 - \kappa_2)r_0}{\mu_1} + r_1\frac{\lambda}{\mu_1},$$

$$f_2 = \frac{\alpha}{\mu_2}f_0 + r_2\frac{\lambda}{\mu_2}.$$

Then

$$(\kappa_2 - \kappa_1)r_0 = (f_1 + f_2)\lambda - \kappa_1 f_0 =$$

$$= f_{\mathbf{0}}\left[\frac{(\lambda + \kappa_1)\lambda}{\mu_1} + \frac{\alpha\lambda}{\mu_2} - \kappa_1\right] -$$

$$- \frac{(\kappa_1 - \kappa_2)r_0\lambda}{\mu_1} + r_1\frac{\lambda^2}{\mu_1} + r_2\frac{\lambda^2}{\mu_2}.$$

Let's consider this expression separately

$$\frac{(\lambda + \kappa_1)\lambda}{\mu_1} + \frac{\alpha\lambda}{\mu_2} - \kappa_1 =$$

$$= \frac{\lambda(\lambda\mu_2 + \alpha\mu_1) - \kappa_1\mu_2(\mu_1 - \lambda)}{\mu_1\mu_2} =$$

$$= \frac{\lambda(\lambda\mu_2 + \alpha\mu_1) - \lambda(\lambda\mu_2 + \alpha\mu_1)}{\mu_1\mu_2} = 0.$$

Then

$$(\kappa_2 - \kappa_1)r_0\frac{\mu_1 - \lambda}{\mu_1} = r_1\frac{\lambda^2}{\mu_1} + r_2\frac{\lambda^2}{\mu_2},$$

and

$$\kappa_2 = \frac{\mu_1}{\mu_1 - \lambda}\left[\frac{r_1}{r_0}\frac{\lambda^2}{\mu_1} + \frac{r_2}{r_0}\frac{\lambda^2}{\mu_2}\right] + \kappa_1 =$$

$$= \frac{\lambda^3\mu_2^2 + \lambda^3\alpha\mu_2 + \alpha\mu_1^2\lambda^2 - \lambda^3\alpha\mu_1}{\mu_2^2(\mu_1 - \lambda)^2} + \kappa_1 .$$

We have found that the parameter κ_2 equals

$$\kappa_2 = \frac{\lambda^3\mu_2^2 + \lambda^3\alpha\mu_2 + \alpha\mu_1^2\lambda^2 - \lambda^3\alpha\mu_1}{\mu_2^2(\mu_1 - \lambda)^2} + \kappa_1. \qquad (13)$$

Second order asymptotic i.e. the proven theorem 2, shows that the asymptotic probability distribution of a number $i(t)$ of calls in an orbit is Gaussian with mean asymptotic κ_1/σ and dispersion κ_2/σ. Then, with the following prelimit distribution in mind

$$P(i) = P_0(i) + P_1(i) + P_2(i),\ i \geq 0, \qquad (14)$$

we could build an approximation for said distribution and in particular the $P^{(2)}(i)$ approximation

$$P^{(2)}(i) = (L(i + 0.5) - L(i - 0.5))(1 - L(- 0.5))^{-1}, \qquad (15)$$

where $L(x)$ is the normal distribution function with parameters κ_1/σ and κ_2/σ.

Gaussian approximation (15), as will be shown below, is fairly applicable at low values $\sigma < 0{,}05$ and gives relative error at $\sigma > 0{,}05$. Moreover, prelimit distribution (14) is asymmetrical whilst the Gaussian approximation (15) is built upon the basis of symmetrical normal distribution.

NUMERICAL ALGORITM FOR SOLVING SYSTEM (1)

Let's write down system (1) at $i = 0$, $i = 1$ and $i \geq 2$, then we will have three systems

$$-(\lambda + \alpha)P_0(0) + \mu_1 P_1(0) + \mu_2 P_2(0) = 0,$$
$$-(\lambda + \mu_1)P_1(0) + \lambda P_0(0) + \sigma P_0(1) = 0,$$
$$-(\lambda + \mu_2)P_2(0) + P_0(0)\alpha = 0. \qquad (16)$$

$$-(\lambda + \sigma + \alpha)P_0(1) + \mu_1 P_1(1) + \mu_2 P_2(1) = 0,$$
$$-(\lambda + \mu_1)P_1(1) + \lambda[P_1(0) + P_0(1)] +$$
$$+2\sigma P_0(2) = 0,$$
$$-(\lambda + \mu_2)P_2(1) + P_0(1)\alpha + P_2(0)\lambda = 0. \qquad (17)$$

$$-(\lambda + i\sigma + \alpha)P_0(i) + \mu_1 P_1(i) + \mu_2 P_2(i) = 0,$$
$$-(\lambda + \mu_1)P_1(i) + \lambda[P_1(i-1) + P_0(i)] +$$
$$+(i+1)\sigma P_0(i+1) = 0,$$
$$-(\lambda + \mu_2)P_2(i) + P_0(i)\alpha + P_2(i-1)\lambda = 0, i \geq 2. \quad (18)$$

Let's consider $P_0(0) = 1$. Using the third and the first equations of the system (16) we could write down

$$P_2(0) = \frac{\alpha}{\lambda + \mu_2}, \quad P_1(1) = \frac{1}{\mu_1}\{(\lambda + \alpha)P_0(0) - \mu_2 P_2(0)\}.$$

Using the second equation of the system (16) we could write down

$$P_0(1) = \frac{1}{\sigma}\{(\lambda + \mu_1)P_1(0) - \lambda P_0(0)\}.$$

Using the third and the first equations of the system (17) we could write down

$$P_2(1) = \frac{1}{\lambda + \mu_2}\{\alpha P_0(1) + \lambda P_2(0)\},$$

$$P_1(1) = \frac{1}{\mu_1}\{(\lambda + \alpha + \sigma)P_0(1) - \mu_2 P_2(1)\}.$$

Further at $2 \leq i \leq N$ the recurrent procedure is implemented by the following equations

$$P_0(i) = \frac{1}{i\sigma}\{(\lambda + \mu_1)P_1(i-1) - \lambda P_0(i-2) - \lambda P_0(i-1)\},$$

$$P_2(i) = \frac{1}{\lambda + \mu_2}\{\alpha P_0(i) + \lambda P_2(i-1)\},$$

$$P_1(i) = \frac{1}{\mu_1}\{(\lambda + \alpha + i\sigma)P_0(i) - \mu_2 P_2(i)\}.$$

By normalizing the obtained results we have found the solution $P_n(i)$ of system (1) for all $0 \leq i \leq N$ .Suggested numerical algorithm is fairy effective as it allows finding the solution $P_n(i)$ for large values (up to thousands) of N.

APPROXIMATION ACCURACY $P^{(2)}(i)$ AND ITS APPLICATION AREA

Approximation accuracy $P^{(2)}(i)$ will be defined by using Kolmogorov equation

$$\Delta = \max_{0\le i\le N}\left|\sum_{v=0}^{i}\left(P(v)-P^{(2)}(v)\right)\right|$$

For range between distributions $P(i)$ and $P^{(2)}(i)$, where distribution $P(i)$ is defined by using numerical algorithm and the approximation $P^{(2)}(i)$ is built upon the basis of the second asymptotic and the obtained Gaussian distribution. Tables 1-5 contain values for this range Δ for various values of rate λ and σ.We consider $\mu_1 = 1$ and $\mu_2 = 2$ for all Tables. Let's consider α = 1.

Table 1: Kolmogorov range

	λ = 0,5	λ = 0,6	λ = 0,7
σ = 1	0,092	0,108	0,123
σ = 0,5	0,066	0,079	0,092
σ = 0,1	0,064	0,039	0,045
σ =0,05	0,026	0,028	0,032

Table 2: Kolmogorov range

	λ = 0,8	λ = 0,9	λ = 0,95
σ = 1	0,116	0,163	0,174
σ = 0,5	0,106	0,123	0,131
σ = 0,1	0,052	0,060	0,064
σ =0,05	0,037	0,042	0,045

Analysis of values of Gaussian approximation tabulated in tables 1-2 lets us make the following conclusions.The approximation accuracy naturally increases with the deterioration of parameter σ value. With increasing values of rate λ (intensity of the incoming flow) the Gaussian approximation accuracy decreases. Let's say that the approximation error is allowed if the Kolmogorov range Δ < 0,05 and the second order approximation is allowed for fairly small values of σ parameter, to be precise σ < 0,05. Consider α = 10.

Table 3: Kolmogorov range

	λ = 0,5	λ = 0,6	λ = 0,7
σ = 1	0,053	0,063	0,072
σ = 0,5	0,039	0,046	0,053
σ = 0,1	0,018	0,021	0,024
σ =0,05	0,013	0,014	0,017

Table 4: Kolmogorov range

	λ = 0,8	λ = 0,9	λ = 0,95
σ = 1	0,083	0,094	0,100
σ = 0,5	0,060	0,068	0,072
σ = 0,1	0,027	0,030	0,032
σ =0,05	0,019	0,021	0,023

Considering λ = 0,8, $\mu_1 = 1$, $\mu_2 = 2$, by changing values of parameters α and σ and by numerically solving the probability distribution system (1), we could find Kolmogorov range between Gaussian approximation of probability distribution of a number of calls in an orbit and the numerical distribution.

Table 5: Kolmogorov range

	σ = 0,2	σ = 0,1	σ =0,03	σ =0,01
α = 1	0,072	0,052	0,028	0,016
α = 3	0,058	0,041	0,022	0,013
α = 5	0,049	0,035	0,019	0,011

Analysis of values tabulated in tables 3-5 shows that the accuracy of Gaussian approximation greatly increases while increasing α, and therefore the area of applicability increases too. The area of applicability doubles in size and is applicable at $\sigma \le 0{,}05$. Density diagrams of probability distributions and distribution function diagrams of a number of calls in an orbit are shown in figures 2-4. The dotted line represents designated density of asymptotical distribution probabilities.

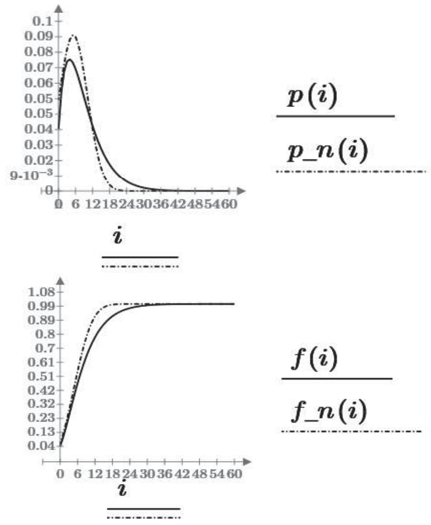


Figure 2: λ = 0,8 σ = 1 α = 1 $\mu_1 = 1$ $\mu_2 = 2$ Δ = 0,116

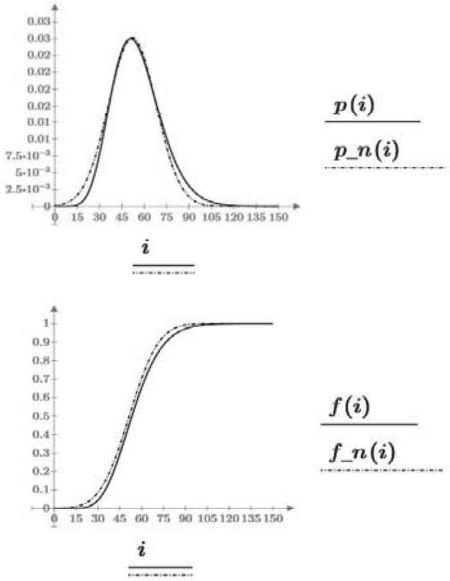


Figure 3: λ = 0,8 σ = 0,1 α = 1 $\mu_1 = 1$ $\mu_2 = 2$ Δ = 0,052

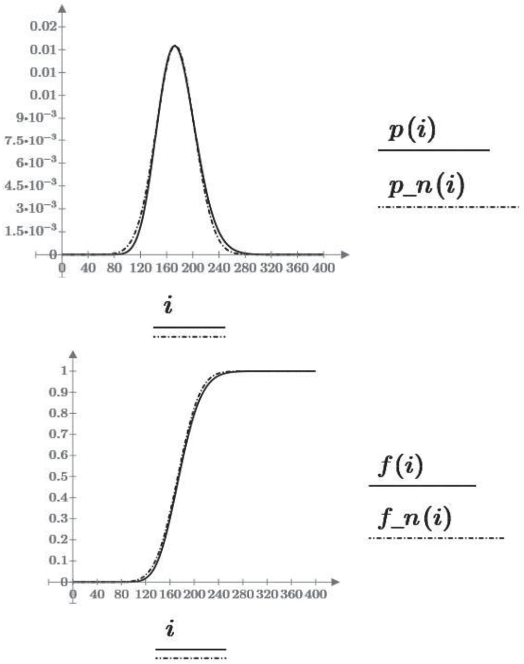


Figure 4: $\lambda = 0{,}8$ $\sigma = 0{,}03$ $\alpha = 1$ $\mu_1 = 1$ $\mu_2 = 2$ $\Delta = 0{,}016$

CONCLUSIONS

In this paper we have considered retrial queue with two-way communication. To research the system in question we have found first and second order asymptotics of a number of calls in an orbit in asymptotic condition of a low rate of retrials. Based on the obtained asymptotics we have built the Gaussian approximation of a probability distribution of a number of calls in an orbit. We have defined the conditions of applicability of the obtained approximation depending on values of system-defining parameters. As criteria we have chosen the Kolmogorov range assuming that the allowed approximation error is less than 0,05. By analyzing the obtained results we can make the conclusion that the accuracy of Gaussian approximation increases while decreasing values of σ parameter, increasing values of λ parameter and/or increasing values of α parameter
The results obtained in this paper are planned to be generalized for the case of correlated incoming flow and random time of serving in retrial queues with two-way communication.

REFERENCES

Artalejo, J.R. and Gomez-Corral, 2008. "A. Retrial Queueing Systems: A Computational Approach," Springer, Berlin.

Artalejo, J.R. and Resing J.A.C. 2010. Mean value analysis of single server retrial queues. *Asia-Pacific Journal of Operational Research* 27, 335-345.

Artalejo, Jesús R. and Tuan, P.D. 2012. Markovian retrial queues with two way communication. *Journal of industrial and management optimization*, 8 (4). 781-206.

Choi, B.D., Choi, K.B. and Lee, Y.W. 1995. M/G/1 retrial queueing systems with two types of calls and finite capacity. Queueing Systems 19, 215-229.

Falin, G.I. 1979. Model of coupled switching in presence of recurrent calls. *Engineering Cybernetics Review* 17, 53-59.

Flajolet, P. and Sedgewick, R. 2009. "Analytic Combinatorics," Cambridge University Press, Cambridge.

Martin, M. and Artalejo, J.R. 1995. Analysis of an M/G/1 queue with two types of impatient units. *Advances in Applied Probability* 27, 840-861.

Nazarov A.A., Chernikova Ya.E. 2014. The Accuracy of Gaussian Approximations of Probabilities Distribution of States of the Retrial Queueing System with Priority of New Customers. Communications in Computer and Information Science. Switzerland: Springer, 2014. Vol. 487. Information Technologies and Mathematical Modelling. P. 325-333. DOI: 10.1007/978-3-319-136714_37

Nazarov A., Izmaylova Y. 2016. Asymptotic Analysis Retrial Queueing System M|GI|1 with Hyper Exponential Distribution of the Delay Time in the Orbit and Exclusion of Alternative Cuctomers. Inform. Tehnol. and Mathem. Modelling. Queue-ing Theory and Applic. Springer. P. 292-302. DOI 10.1007/978-3-319-44615-8_26

Nazarov, A. and Paul S. 2016. A Number of Customers in the System with Server Vacations. *Communications in Computer and Information Science*. Springer, vol. 601, 334-343. DOI: 10.1007/978-3-319-30843-2_35

AUTHOR BIOGRAPHIES

ANATOLY NAZAROV was born in Tomsk region, Russia and went to the Tomsk State University, where he obtained his doctor's degree in 1986. Since 2000 Nazarov is the head of the Department of Probability Theory and Mathematical Statistics of Tomsk State University. He is leading a large research group in the field of queueing theory in TSU. His e-mail address is: nazarov.tsu@gmail.com and his Web-page can be found at http://www.fpmk.tsu.ru/node/2083.

SVETLANA PAUL was born in Krasnoyarsk region, Russia and went to the Tomsk State University, where she obtained her degree in 2008. Since 2009 she works as an associate professor Department of Probability Theory and Mathematical Statistics of Tomsk State University. She is the member of research group in the field of queueing theory in TSU. Her e-mail address is: paulsv82@mail.ru and her Web-page can be found at http://www.fpmk.tsu.ru/node/2290.

IRINA GUDKOVA received her M.Sc. degree in applied mathematics and Cand.Sc. degree in applied mathematics and computer sciences from the Peoples' Friendship University of Russia (RUDN University) in 2009 and 2011, respectively. She is currently an Associate Professor with the Applied Probability and Informatics Department, RUDN University. She has co-authored multiple research works. Her current research interests include mathematical modelling and performance analysis of 4G/5G networks, smart cities, spectrum sharing, multicast services, radio access, teletraffic theory, and queuing theory. Her e-mail address is: gudkova_ia@pfur.ru.

MODELLING OF VERTICAL HANDOVER FROM UNTRUSTED WLAN NETWORK TO LTE

Grebeshkov Alexander*, Zaripova Elvira§, Roslyakov Alexander*, Samouylov Konstantin§

* Department of Automatic Telecommunication
Povolzhskiy State University
of Telecommunications and Informatics

23 Leo Tolstoy st., Samara 443010,
Russian Federation
Email: grebeshkov-ay@psuti.ru,
arosl@mail.ru

§ Department of Applied Probability
and Informatics
Peoples' Friendship University of Russia
(RUDN University)
6 Miklukho-Maklaya st.,
Moscow 117198, Russian Federation
E-mail: zaripova_er@rudn.university,
samuylov_ke@rudn.university

KEYWORDS

Vertical handover, 3GPP access network, LTE, modelling, simulation, untrusted WLAN.

ABSTRACT

Nowadays, operators offer telecommunication services at a very high level. For the support of the best quality of service it would be reasonable to implement a handover to another radio access network cell in the same territory, a so-called vertical handover (VHO), by transferring the user connection and IP session from the current access network to a new network. In this paper, we analyze vertical handovers from an untrusted WLAN network to an LTE network. This type of handover from an untrusted WLAN network is the most difficult because of the many nodes of an LTE network related to authorization and resource allocation. The paper includes a synthesis of the basic procedures for the session setup and LTE resource allocation. Unlike other papers, we represent a complicated VHO procedure as a sequence of at least 40 signalling messages from the discovery of a new network for the VHO to its full completion at the target LTE network. We also propose a vertical handover time estimation method for the said VHO procedure. We perform numerical experiments on realistic data.

INTRODUCTION

Modern wireless access networks can be characterized as heterogeneous and ubiquitous with overlapping and seamless mobile and IP–connectivity. Enhanced quality of service (QoS) for the delivery of multimedia services can be made through a handoff to the best network from the point of view of such criteria as increasing received signal strength (RSS), decreasing handover latency, increasing available bandwidth, optimal power consumption, satisfaction of user preferences with some decision schemes (Ahmed et al., 2014).

Handover process is categorized into vertical handover (VHO) and horizontal handover (HHO). Horizontal handover occurs in homogeneous network. This type of handover is well researched. A vertical handover occurs when user equipment (UE) switches between the cells of two different networks located on the same territory but which support different radio access technologies. During a VHO process, the UE can fetch a new physical connection and a new IP session. In (Abdoulaziz et al., 2012), the VHO time (or latency) is discussed as one of the main parts of a handover process because a mobile terminal may leave the WLAN area before the VHO to a 3GPP network will be completed. In (Yu et al. 2013) there is a case with lost or buffering packets due to a time out in the WLAN–3GPP VHO. Since VHO time has a clear impact on QoS and mobility, an analytical estimation of VHO time should be useful for reasonability analysis of VHO decision making in real time.

This paper is organized as follows: the “VHO standardizations and specifications” section is devoted to the analysis of current VHO standards and technical specifications. The “WLAN–LTE VHO procedure design” section includes information on a step–by–step time-dependent VHO procedure. The ”Mathematical modelling of VHO procedures” section includes an analytical model of the discussed VHO procedure. A numerical example for initial realistic data is shown in the ”Numerical experiment” section.

VHO STANDARDIZATIONS AND SPECIFICATIONS

One of the main technical specifications which considers VHO architecture, principles and scenarios is the (3GPP TS 23.402 2016). This standard has been designed for non-3GPP access network interconnection with an Evolved 3GPP Packet Switched (EPS) domain. This technical specification describes an IP–based protocol for evolved UMTS Terrestrial Radio Access Networks (E–UTRAN). EPS may provide the UE with assistance data and policies about available networks for VHO by establishing secure communication in the form of an IP tunnel between the UE and the EPS core. This means that the security aspects of a 3GPP access and non–3GPP (non–compliant access with 3GPP specification) need to be taken into consideration for the VHO procedure and VHO time estimation. Some security aspects of VHOs are described in (3GPP TS 33.222

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

2016) where an authentication mechanism was proposed for the UE. This mechanism is appropriable for an HTTP client and server which can authenticate each other based on a shared key, generated during the bootstrapping procedure. The shared key can be obtained as a master key for the generation of transport layer security (TLS) session keys to protect initial signalling messages between the UE and the access network discovery and selection function (ANDSF). The ANDSF is a source of information about a network for VHOs. In (3GPP TS 33.402 2016), there are additional descriptions of an IPsec-based double-stack mobile IPv6 (DSMIPv6) protocol and security associations with the Internet Key Exchange version 2 (IKEv2). The IPsec security association is established between the UE and the 3GPP node which acts as a home agent (HA).

In the content of the VHO, the IEEE 802.21 procedure (IEEE Std 802.21–2008 2008) acts as a Media Independent Handover Function (MIHF). The MIHF provides abstracted services concerning VHO policy and signaling to higher layers (from Level 3 OSI and higher) with technology-specific protocol entities. The MIHF communicates with the lower layers (Level 2 OSI) as a logical part of the mobility-management protocol stack through technology-specific interfaces. Generally, the IEEE 802.21 MIH architecture and protocol are more abstracted than the 3GPP technical specification and no detailed procedure of a VHO between WLAN and 3GPP (LTE) networks is specified.

The efforts of the IETF are aimed at IP–mobility, authentication and security procedures concerning VHO. RFC 5996 (The Internet Engineering Task Force RFC 5996 2010) and RFC 4555 (The Internet Engineering Task Force RFC 4555 2006) contain specifications of an IKEv2 protocol that allows the IP addresses associated with IKEv2 and tunnel mode IPsec security associations to change. RFC 4793 (The Internet Engineering Task Force RFC 4793 2007) describes a certificate-based authentication method of the client host followed by Extensible Authentication Protocol (EAP) authentication of the user. One of the key specifications in the context of RFC 5555 (The Internet Engineering Task Force RFC 5555 2009) declares that the IPv6 protocol is not widely deployed. It is a reasonable option to extend MIPv6 capabilities to allow dual stack mobile nodes to request that their HA tunnel IPv4/IPv6 packets be addressed to their home addresses, as well as their IPv4/IPv6 care-of address (CaO). Extensions are defined for the binding update and binding acknowledgement. Finally, RFC 6611 (The Internet Engineering Task Force RFC 6611 2012) defines the MIPv6 bootstrapping procedures. It enables the assignment of home agents by utilizing the Dynamic Host Configuration Protocol (DHCP) v6 and the Authentication, Authorization, and Accounting (AAA) protocol.

Due to the design of the VHO time estimation method, it is important to form a WLAN–LTE VHO procedure that is associated with real network performance values. In the context of procedure design, the 3GPP interworking architecture for 3GPP and non–3GPP access networks is more preferable than IEEE 802.21 MIH. Our model fully complies with the technical specifications of 3GPP.

WLAN-LTE VHO PROCEDURE DESIGN

As shown above, one of the substantial elements in a 3GPP–based VHO is the ANDSF, which provides an inter–system mobility policy, network access discovery information with a list of prioritized networks for VHO as well as a WLAN selection policy and rules. The ANDSF is responsible for delivering information on discovered access networks in response to UE requests. For example, the ANDSF may list the prioritized networks for VHO in response to the UE request so that the 3GPP (LTE) has a maximal priority of 1, WLAN (Wi-Fi) has a priority of 2, while WLAN (WiMAX) has a priority of 3. There is a document (3GPP TS 23.002 2016) that specified the S2a interface for trusted non–3GPP IP access, and the S2c interface for untrusted non–3GPP access. Trusted and untrusted non-3GPP access networks support the IP protocol and use radio/wireless or fixed access technology whose specification and standardization is out of the scope of 3GPP. “Trusted” means full support of 3GPP–based authentication. For interaction between the UE and the ANDSF there is the S14 interface.

The discovery part of the procedure using the ANDSF was simulated in (Xenakis et al. 2016) but without a detailed description of VHO execution after decision making and with no general mathematical model for the establishment of a new connection. In (Triantafyllopoulou et al. 2012), a new network discovery algorithm was proposed for the VHO decision stage but with no mathematical modelling of the VHO procedure execution phase.

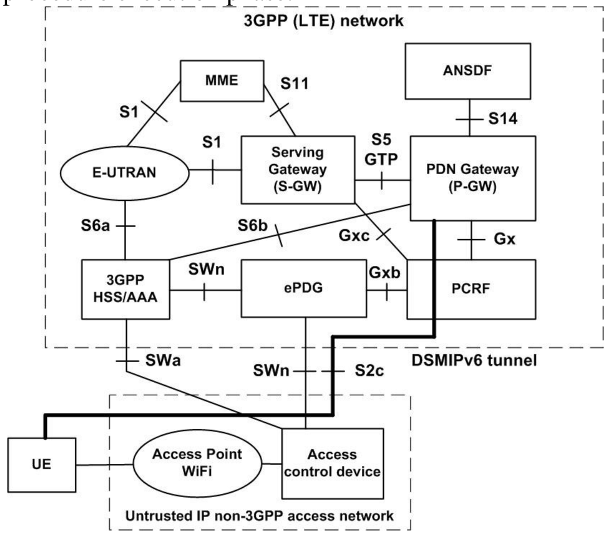


Figure 1: Architecture for WLAN – LTE VHO

For the following research, let’s suppose that the UE receives a list of prioritized networks for a VHO i.e. the decision making process has been finished and the results have been placed into a list. The VHO procedure will be based on host–based mobility, where the VHO is initiated by the UE (a mobile terminal) in accordance with (Boccardi et al. 2014) within a 5G device-centered network architecture. The architecture used for VHO time estimation is shown in Fig. 1. All interfaces in Fig. 1 are specified by 3GPP. The Access point and access control device belonging to the non-3GPP access network were described in (Gast 2005) and both support connection and information exchange between WLAN and 3GPP. A mathematical model of this architecture and VHO procedure will provide at least an upper bound for VHO time estimation.

On the side of the 3GPP network, there is an E-UTRAN subsystem and an EPS. The E-UTRAN in the context of this work supports a physical connection between the UE and 3GPP network. The home subscriber server (HSS) is combined with an AAA server and supports the registration, authentication and authorization of the UE in the 3GPP network. The Mobility Management Entity (MME) is responsible for Serving Gateway (S-GW) and Packet Data Network (PDN) Gateway (P-GW) selection, roaming, authentication, dedicated bearer establishment and the transfer of information between the MME and HSS/AAA.

The Evolved Packet Data Gateway (ePDG) can be used for the decapsulation and encapsulation of packets for IPSec tunnels, tunnel authentication and authorization, care–of–addresses (CoAs) for associating this mobile node with visited mobile networks including CoA for S2c. The S-GW is responsible for packet routing and forwarding, transport level packet marking in the uplink and the downlink. The P-GW provides PDN connectivity to the UE using non-3GPP access networks. The Home network Policy and Charging Rules Function (hPCRF) supports UE serving policy decision making, charging control of service data flow and the IP bearer resources.

The scheme of a step-by-step VHO procedure based on the 3GPP specifications can be divided into (#) phases. Phase A is presented in Fig. 2. Step (1) is the initial stage of the procedure. The UE sends a message via a non-trusted 3GPP IP access network with an ANDSF server host name. The ANDSF server name is public. Steps (2)–(3) include a special request–respond message with a pre–shared key before TLS tunnel establishment. Step (4) is a TLS finish message which is a part of the handshake procedure.

Step (5) is a request from the UE to the ANDSF to retrieve information on discovered networks Step (6) is the ANDSF's response with information on the available access networks, mobility rules and ePDG configuration information. The UE turns on a radio interface, measures access network characteristics (e.g. RSS) and selects a preferable access network (e.g. 3GPP LTE) for the VHO.

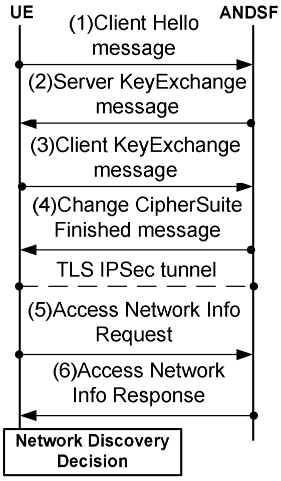

UE
ANDSF
(1)Client Hello message
(2)Server KeyExchange message
(3)Client KeyExchange message
(4)Change CipherSuite Finished message
TLS IPSec tunnel
(5)Access Network Info Request
(6)Access Network Info Response
Network Discovery Decision

Figure 2: WLAN – LTE VHO Procedure Phase A

Phase B is presented in Fig. 3. Step (7) includes the UE state with the initial request for attachment to the 3GPP ePDG.

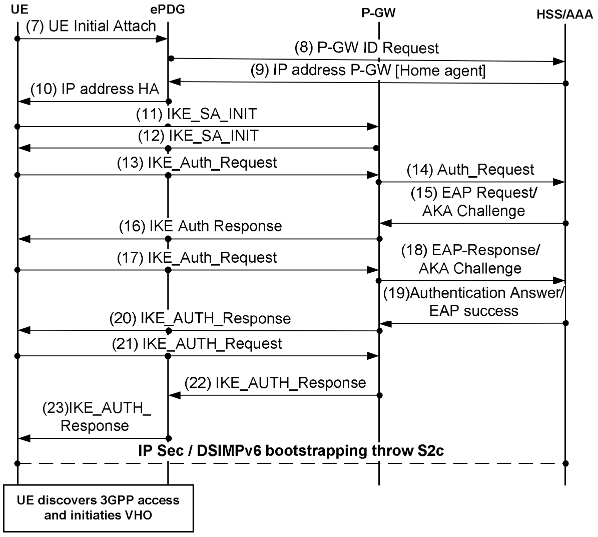


Figure 3: WLAN – LTE VHO Procedure Phase B

Step (7) includes the UE action of sending an Initial Attach message to the ePDG. Step (8) is where the ePDG sends the request from step (7) for P-GW identification to the 3GPP HSS/AAA. Step (9) includes the selection of the P-GW closest to the ePDG by the HSS/AAA; the IP address of the P-GW or the HA is sent by the HSS/AAA to the ePDG. Step (10) includes the further transmission of the IP address or HA to the UE by the ePDG. Steps (11) and (12) include the initiation of an IKEv2 exchange using a cryptographic algorithm and the successful result of this exchange.

Step (13) includes the authorization procedure to obtain an IPv6 network prefix and to protect DSMIPv6 signaling for future communication. Steps (14) and (15) are responsible for UE authentication in the 3GPP with an Authentication and Key (AKA) 3GPP protocol and DSMIPv6 with security support. Step (16) includes the transmission of security parameters to the UE for the

EAP procedure to restart the IKEv2. Step (17) is needed for the IKEv2 parameters to be checked by the UE and the generation of a UE response message as an EAP message for the P-GW and HSS/AAA. Steps (18) and (19) include the final part of the authorization procedure, where the HSS/AAA generates an Authentication Answer of EAP success. Step (20) is the translation of an EAP success message to the UE.

Step (21) includes the process of generating the Master Session Key (MSK) on the side of the UE and the generation of authentication parameters with the MSK. These parameters will be sent to the P-GW from the UE. Steps (22) and (23) describe the reception of the assigned IP address for the UE. After step (23), there is a secure IPsec tunnel for an S2c interface and the UE has authorization for VHO initialization via 3GPP.

Phase C is presented in Fig. 4. Steps (24) - (26) are the attachments/connections at the physical (L1) and channel (L2) levels to the eNode (LTE base station). Steps (27) - (30) are the sequential requests for a VHO IP–session establishment for the UE in the 3GPP network. Steps (31) - (33) are the responses containing a session grant and parameters. As a result, there is an establishment of a radio and access bearer.

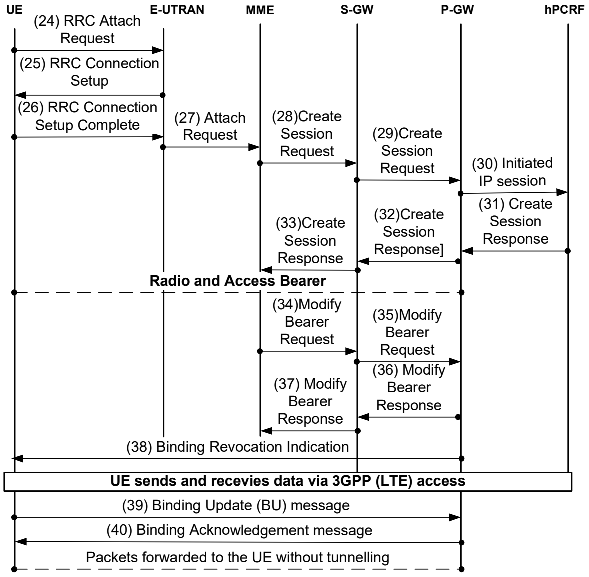

UE
E-UTRAN
MME
S-GW
P-GW
hPCRF
(24) RRC Attach Request
(25) RRC Connection Setup
(26) RRC Connection Setup Complete
(27) Attach Request
(28)Create Session Request
(29)Create Session Request
(30) Initiated IP session
(31) Create Session Response
(32)Create Session Response]
(33)Create Session Response
Radio and Access Bearer
(34)Modify Bearer Request
(35)Modify Bearer Request
(36) Modify Bearer Response
(37) Modify Bearer Response
(38) Binding Revocation Indication
UE sends and recevies data via 3GPP (LTE) access
(39) Binding Update (BU) message
(40) Binding Acknowledgement message
Packets forwarded to the UE without tunnelling

Figure 4: WLAN – LTE VHO Procedure Phase C

Steps (34) - (37) are needed to modify packet connection and then to tunnel the packets from the untrusted non-3GPP IP network to the 3GPP (LTE) access network and EPS, routing packets to the S-GW for the radio and access bearer. We must note that steps (36) and (38) could either occur simultaneously or after a specific time interval.

Step (38) is initialized by the P-GW approx. after a 1 ms expiration delay in accordance with the P-GW delay from (Nikaein and Krco, 2011). This step includes the initial message from the P–GW to the UE for the de-registration of the DSMIPv6 binding.

Steps (39)–(40) include the final messages for the de-registration of the UE DSMIPv6 binding. After step (40), the reception of packets by the UE without tunneling has been realized. Finally, the UE sends and receives data packets only through the eNode (LTE) and EPS system. Formally, the VHO is over and from this moment the tunnel is no longer needed.

MATHEMATICAL MODELLING OF VHO PROCEDURES

In this section we offer a method for VHO procedure time estimation. In the previous section, nine functional entities of VHO procedures were described: UE (I), ANDSF (II), ePDG (III), E-UTRAN (IV), MME (V), S-GW (VI), P-GW (VII), hPCRF (VIII) and HSS/AAA (IX) as well as the 40 signaling messages that are transmitted between them.

In this section we have built a mathematical model for analysing the sojourn time Δ from the first initiating message until the final message of VHO completion.

For preliminary performance measures, we used a well-known class of queueing networks for open, closed or mixed models with various service disciplines. The most noted paper, (Bassket, et al, 1975), on the BCMP method was presented by Baskett, Chandy, Muntz and Palacios. We will base our method on BCMP.

Let's use two subsets of network nodes: $M_1 = \{\text{I}\}$ and M_2, which contains the other eight nodes II-IX. By Basharin-Kendall notation, the first node is of type $M \mid M \mid inf$, while the nodes in subset M_2 are of type $M \mid M \mid 1 \mid inf$ with the FCFS service discipline. External arrivals are Poisson with a rate of λ_0. The service rates on the nodes is μ_i, where $i \in (M_1 \cup M_2)$. The average sojourn time for a customer is equal to μ_1^{-1} in the first node and $(\mu_i - \lambda_i)^{-1}$ in the FCFS nodes, where λ_i is the total incoming intensity rate to the i th node.

During the VHO procedure, messages (1)–(40) move from one node to another as shown in Fig. 5, where the numbers on the arrows correspond to the index numbers of the signalling messages.

The steady state condition for the queueing network is:

$$\lambda_0 < \min\left(\frac{\mu_2}{3};\frac{\mu_3}{3};\frac{\mu_4}{2};\frac{\mu_5}{3};\frac{\mu_6}{4};\frac{\mu_7}{10};\mu_8;\frac{\mu_9}{3}\right) \quad (1)$$

Using the approach of (Raad, et al. 2013; Gaidamaka and Zaripova 2014; Samouylov, et al. 2016) the average sojourn time Δ_{ident} with identical customers can be estimated using formula (2):

$$\Delta_{ident} = 12\mu_1^{-1} + \frac{3}{\mu_2 - 3\lambda_0} + \frac{3}{\mu_3 - 3\lambda_0} + \frac{2}{\mu_4 - 2\lambda_0} + \\ + \frac{2}{\mu_5 - 3\lambda_0} + \frac{3}{\mu_6 - 4\lambda_0} + \frac{10}{\mu_7 - 10\lambda_0} + \frac{1}{\mu_8 - \lambda_0} + \\ + \frac{3}{\mu_9 - 3\lambda_0}. \qquad (2)$$

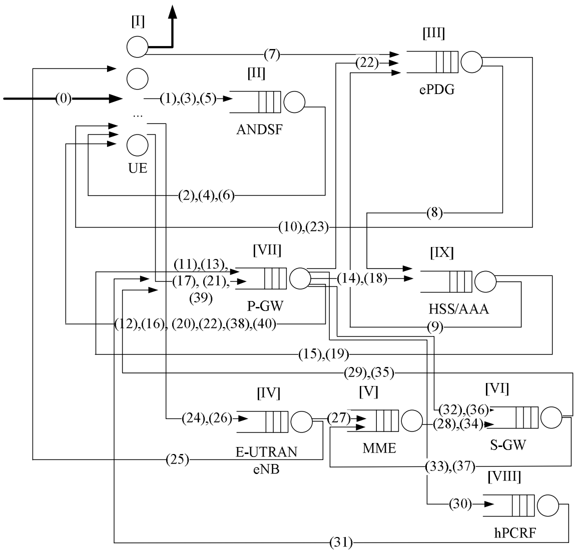


Figure 5: Nine-Node Open Queueing Network

The first addendum $12\mu_1^{-1}$ shows the total time at the first node within the VHO procedure, i.e. the 12 times that the signaling messages have gone through the UE. The second addendum $\frac{3}{\mu_2 - 3\lambda_0}$ corresponds to the total time interval on the second ANDSF node, the VHO procedure has gone through the ANDSF node 3 times.

We simplify our VHO procedure and forward messages (36) and (38) simultaneously from node (VII) P-GW. You can observe changes in the fifth and sixth addendums, where we estimated additional load from signaling messages (36) and (37).

Messages (24), (25) and (27) have different service time. For this case, the average sojourn time Δ can be estimated by formula (3) for heterogeneous customers.

$$\Delta = 10\mu_1^{-1} + T_{24} + T_{26} + \frac{3}{\mu_2 - 3\lambda_0} + \frac{3}{\mu_3 - 3\lambda_0} + \\ + \frac{2}{\mu_4 - 2\lambda_0} + \frac{1}{\mu_5 - 3\lambda_0} + T_{27} + \frac{3}{\mu_6 - 4\lambda_0} + \\ + \frac{10}{\mu_7 - 10\lambda_0} + \frac{1}{\mu_8 - \lambda_0} + \frac{3}{\mu_9 - 3\lambda_0}. \qquad (3)$$

NUMERICAL EXPERIMENT

The proposed mathematical model allows us to estimate the VHO time of the proposed method. To illustrate this estimation method, we used input data from (Nikaem and Krco 2011; Cardona, et al. 2013; Prados-Garzon et. al. 2015; Granlund et. al. 2015). The amount of transactions per second depends on the provider, network configuration and many other different parameters. We assume that a maximum of 10 % of subscribers are in need of a vertical handover on the same territory as the cell. Each request for a VHO generates the 40 signaling messages that have been described above. Average service time intervals for our preliminary analysis are shown in Table 1. The signaling messages service times differ from each other because of their functionality and different length.

Table 1: Average Service Time

Nodes	Average service time, μ_i^{-1}, ms	Ref.
I - UE	77.5 for (24) 28.5 for (26) 2 for other steps	(Nikaem and Krco, 2011)
II - ANDSF	70	(estimated as HSS/ AAA)
III - ePDG	2	(estimated as P-GW)
IV - eNB	4	(Cardona, et al., 2013)
V - MME	15 for (27) 1 for other steps	(Cardona, et al., 2013) (Prados-Garzon et. al, 2015)
VI - S-GW	2	(Nikaem and Krco, 2011)
VII - P-GW	2	(Nikaem and Krco, 2011)
VIII - hPCRF	70	(estimation as HSS / AAA)
IX - HSS / AAA	70	(Granlund et. al., 2015)

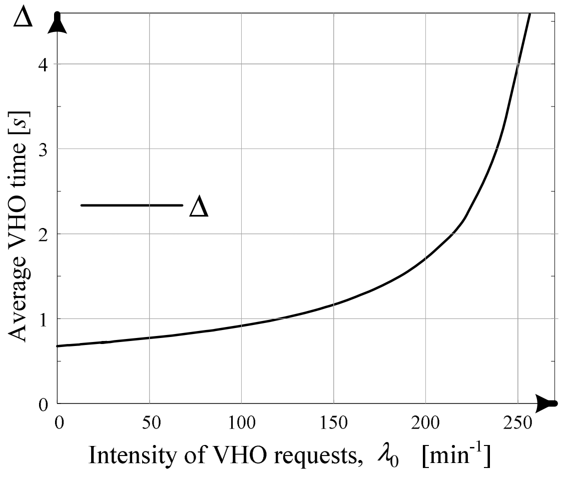


Figure 6: Average VHO Time

There is a dependence of the VHO time on the intensity of VHO requests. It should be noted that it is not necessary to use an ANDSF node when the VHO intensity λ_0 is equal to 0, since the total VHO time will be no more than 462 ms in the beginning. VHO time can be estimated by summing the average service times of the signalling messages. Increasing the intensity of VHO requests affects the average VHO time (see Fig. 6).

The ANDSF node could help select the optimal network for the handover. Therefore this estimation method is an upper bound for VHO time.

CONCLUSION

In this paper, a VHO time estimation method and a mathematical model for a non-trusted IP 3GPP access network to a 3GPP LTE was proposed. This procedure covers VHO phases from network discovery decision to the completion of the VHO procedure with physical and IP–connection re-establishment in the 3GPP network.

At the next stage of research, the scheme of a VHO from a 3GPP (LTE) network to WLAN will be discussed. Another issue at the next stage of research is the modelling of a probabilistic VHO procedure. In this future scheme, the UE has a choice of a given probability to initiate a VHO procedure to 3GPP LTE, to initiate a VHO procedure to WiMAX, or to remain in the current WLAN network.

The aim of future research is implementation of another analytical methods and models to vertical handover time estimation. The first method is for queueing network with given variation coefficients for service times. This method could be implemented for any service times distribution using Kramer and Langenbach-Beltz approximate formula for sojourn time estimation. The second method is for multiphase queuing system with background traffic. This approximate method divides incoming flow into foreground and background traffic, so we could consider several types of traffic.

ACKNOWLEDGEMENTS

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008) and RFBR (research projects No. 15-07-03051 and No. 16-07-00766).

REFERENCES

Abdoulaziz, I.H.; L. Renfa; and Z. Fanzi. 2012. "Handover Necessity Estimation for 4G Heterogeneous Networks." *International Journal of Information Sciences and Techniques*, Vol.2 (Jan.), 1-13.

Adnan, M.; H.Zen; and A.-K.Othman. 2013. "Vertical Handover Decision Processes for Fourth Generation Heterogeneous Wireless Networks." *Asian Journal of Applied Scieneces*, Vol.01 (Dec.), 229-235.

Ahmed, A.; L.M. Boulahia; and D. Gaïti. 2013. "Enabling Vertical Handover Decisins in Heterogeneous Wireless Networks: A State–of–the–Art and A Classification." *IEEE Communications surveys and tutorials*, Vol.16 (Secon Quarter 2014), 776–811.

Baskett F., Chandy K. M., Muntz R. R., Palacios F. G. "Open, Closed, and Mixed Networks of Queues with Different Classes of Customersc *Journal of the ACM*. Vol. 22. No 2. 1975. Pp. 248–260.

Boccardi, F.; R.W. Heath, Jr.; A. Lozano; T. L. Marzetta; and P. Popovski. 2014. "Five Disruptive Technology Directions for 5G." *IEEE Communications Magazine*, Vol.52 (Feb), 74-80.

Gaidamaka Yu., Zaripova E. 2014. "Session setup delay estimation methods for IMS-based IPTV services". *Lecture Notes in Computer Science* 8638, pp. 408-418.

Cardona, N.; J.F. Monserrat; and J. Cabrejas. 2013. "Enabling Technologies for 3GPP LTE-Advanced Networks". In *LTE-Advanced and Next Generation Wireless Networks 2013*, G. de la Roche, A.A. Glazunov and B. Allen (Eds.). John Wiley and Sons Ltd, Chichester, United Kingdom, 3-34.

Gast, M.S. 2005. *802.11 Wireless Networks The Definitive Guide.* O'Reilly, Sebastopol, CA.

Granlund, D.; P. Holmlund; and C. Åhlund. 2015. "Opportunistic Mobility Support for Resource Constrained Sensor Devices in Smart Cities." *Sensors*, Vol.15 (Mar.), 5112-5135

IEEE Std 802.21-2008. 2008. *IEEE Standard for Local and metropolitan area networks – Media Independent Handover Services.* The Institute of Electrical and Electronics Engineers, NY, USA (Nov).

Nikaein, N., and S. Krco. 2011. "Latency for Real-Time Machine-to-Machine Communication in LTE-Based System Architecture". In *Proceedings of the 2011 17^th^ European Wireless Conferense – Sustainable Wireless Technologies* (Vienna, Austria, Apr.27-29). IEEE, 1-6.

Prados-Garzon, J.; J.J. Ramos-Munoz; P. Ameigeiras, P. Andres-Maldonado; J.M. Lopez-Soler. 2015. "Latency evaluation of a virtualized MME". In *Proceedings of the 2016 Wireless Days* (Toulouse, France, Mar. 23-25). IEEE, 1-3.

Raad, A.; Yu. Gaidamaka; and A. Pshenichnikov 2013. "Session initiation model of IPTV service using IMS platform". *Electrosvyaz*. No. 10. Pp. 46-51.

Samouylov, K.; Yu. Gaidamaka; and E. Zaripova. 2016. "Analysis of business process execution time with queueing theory models". *CCIS* 638. Springer International Publishing Switzerland, 315-326.

The Internet Engineering Task Force RFC 4555. 2006. *IKEv2 Mobility and Multihoming Protocol (MOBIKE).* The Internet Society. (Jun).

The Internet Engineering Task Force RFC 4793. 2007. *The EAP Protected One-Time Password Protocol (EAP-POTP).* The IETF Trust. (Feb).

The Internet Engineering Task Force RFC 5555. 2009. *Mobile IPv6 Support for Dual Stack Hosts and Routers.* The IETF Trust. (Jun).

The Internet Engineering Task Force RFC 5996. 2010. *Internet Key Exchange Protocol Version 2 (IKEv2).* The IETF Trust. (Sep).

The Internet Engineering Task Force RFC 6611. 2012. *Mobile IPv6 (MIPv6) Bootstrapping for the Integrated Scenario.* The IETF Trust. (Sep).

Triantafyllopoulou, D.: T. Guo; and K. Moessner. 2012. "Energy Efficient ANDSF-assisted Network Discovery for non-3GPP Access Networks". In *Proceedings of the 2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks* (Barcelona, Spain, Sep.17-19). IEEE, Piscataway, N.J., 297-301.

Xenakis, D.; N. Passas; L. Merakos; and C. Verikoukis. 2016. "ANDSF-Assisted Vertical Handover Decisions in the IEEE 802.11/LTE-Advanced Network." *Computer Networks*, Vol.106 (Sep), 91–108.

Yu, F.R.; L. Ma; and V.C.M. Leung. 2013. "Support of Node Mobility between Networks". In *Multihomed Communication with SCTP (Stream Control Transmission Protocol) 2013*, V.C.M Leung; E.P. Ribeiro; A. Wagner; and J. Lyengar (Eds.). CRC Press, Boca Raton, 81-98.

3GPP TS 23.002. 2016. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Network architecture (Release 14).* 3GPP, Sophia Antipolis Cedex, France (Sep).

3GPP TS 23.402. 2016. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Architecture enhancements for non-3GPP accesses (Release 14).* 3GPP, Sophia Antipolis Cedex, France (Dec).

3GPP TS 33.222. 2016. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Generic Authentication Architecture (GAA); Access to network application functions using Hypertext Transfer Protocol over Transport Layer Security (HTTPS) (Release 13).* 3GPP, Sophia Antipolis Cedex, France (Jan.).

3GPP TS 33.402. 2016. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3GPP System Architecture Evolution (SAE); Security aspects of non-3GPP accesses (Release 14).* 3GPP, Sophia Antipolis Cedex, France (Dec).

AUTHOR BIOGRAPHIES

ALEXANDER YU. GREBESHKOV received his Ph.D. degree from the Moscow Technical University of Communications and Informatics. Since 1996, he has worked at the Povolzhskiy State University of Telecommunications and Informatics. He is currently an Associate Professor at the Department of Automatic Telecommunications. His research interests are network management, decision making methods, performance analysis and modelling of NGNs and post-NGNs. He has published more than 40 scientific papers and 5 books. His e-mail address is: grebeshkov-ay@psuti.ru.

ELVIRA R. ZARIPOVA received her B.Sc. and M.Sc. degrees in applied mathematics at RUDN University in 2001 and 2003, respectively. In 2015, she received her Ph.D. degree in applied mathematics and computer sciences from Peoples' Friendship University of Russia (RUDN University). Since 2003, Elvira Zaripova works at the Telecommunication Systems Department of RUDN University. She is currently an Associate Professor at the Department of Applied Probability and Informatics. Her current research interests lie in the area of performance analysis of radio resource management techniques in LTE networks on which she has published several papers in refereed journals and conference proceedings. Her e-mail address is: zaripova_er@rudn.university.

PROF. ALEKSANDR V. ROSLYAKOV received his Ph.D. degree from the Bonch-Bruevich Saint - Petersburg State University of Telecommunications and Doctor of Sciences degree from the Povolzhskiy State University of Telecommunications and Informatics. In 2008, he became head of the Department of Automatic Telecommunications of PSUTI. His research interests are the performance analysis of VPNs, call-centres, SS#7 networks, NGNs, teletraffic theory, and Internet of Things. He has written more than 100 scientific and technical papers and 17 books. His e-mail address is: arosl@mail.ru. and his web-page can be found at http://www.roslyakov-av.ru.

PROF. KONSTANTIN E. SAMOUYLOV received his Ph.D. degree from the Moscow State University and Doctor of Sciences degree from the Moscow Technical University of Communications and Informatics. During 1985–1996, he held several positions at the Faculty of Sciences at Peoples' Friendship University of Russia (RUDN University), where he became head of the Telecommunication Systems Department in 1996. From 2014, he became head of the Department of Applied Probability and Informatics. During the last two decades, Konstantin Samouylov has been conducting research projects for the Helsinki and Lappeenranta Universities of Technology and several institutes of Russian Academy of Sciences. His current research interests are performance analysis of 5G networks (LTE, M2M), teletraffic theory, signalling network (SIP) planning, and cloud computing. He has written more than 150 scientific and technical papers and 5 books. His e-mail address is: samuylov_ke@rudn.university.

Modeling and simulation of reliability function of a homogeneous hot double redundant repairable system

Vladimir Rykov[1,3], Dmitry Kozyrev[1,2], Elvira Zaripova[1]
[1]Peoples Friendship University of Russia (RUDN University),
6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
[2]V.A.Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia
[3]Gubkin Russian State University of oil and gas, 65 Leninsky Prospekt, Moscow, 119991, Russia
Emails: rykov_vv@rudn.university, kozyrev_dv@rudn.university, zaripova_er@rudn.university

KEYWORDS

System reliability function, mathematical modeling and simulation, redundant systems, markovization method

ABSTRACT

Calculation of a system reliability function is one of the principal problems in reliability theory. It is well known that even for the simplest double redundant repairable system this function has not been analytically calculated yet in case when elements reliability and recovery functions have general distributions.

In current paper we study the reliability function and the mean time to failure of a homogeneous hot standby repairable system with exponential distribution of its elements life time and general distribution of their repair time. With the help of the developed general discrete-event simulation model we extend the previous studies to a general case of a homogeneous hot double redundant repairable system with general distributions of both life and repair time lengths of elements.

It is shown that the results of exact analytical calculation and simulation results have close agreement.

INTRODUCTION AND MOTIVATION

Calculation of a system reliability function is one of the principal problems in reliability theory. It is well known that even for the simplest double redundant repairable system this function has not been analytically calculated yet in case when elements' reliability and recovery functions have general distributions.

First this problem was discussed in the book by Gnedenko, Belyaev, Solov'ev [1] for the case of homogeneous warm double redundant repairable system with the help of renewal theory methods. The same problem for the system with heterogeneous elements by the same methods was discussed also in [2]. In paper [3] multi-dimensional alternative processes have been used for studying the complex reliability systems. A system has been considered in case of no limitations on the number of repair facilities. When there is enough repair units, the components of the process describing the system behavior become independent allowing to calculate the system characteristics. However, in the case of limited number of repair facilities the problem of calculation of reliability function has not been solved yet, and it is considered in the current paper for the simplest case of a homogeneous hot standby system with only one repair unit.

On the other hand in series of works by B.V. Gnedenko, A.D. Solov'ev [4], [5], [6] and others it was shown that in case of "quick" restoration the system life time distribution becomes asymptotically insensitive to the shapes of its elements' life and repair time distributions and in scale of the system mean life time it tends to the exponential one.

In papers [7], [8], [9] a homogeneous cold standby double redundant system has been considered in the case when one of the input distributions (either of life or repair time lengths) is exponential. For these models explicit expressions for stationary probabilities have been obtained which show their evident dependence on the non-exponential distributions in the form of their Laplace-Stiltjes transforms. At that, the numerical investigations, performed by V.Rykov and D.Kozyrev and presented at the Eighth International Workshop on Simulation (Vienna, September 21st – 25th, 2015), show that this dependence becomes vanishingly small under "quick" restoration. However the problem of the reliability function calculation was not considered.

In current paper we study the reliability function and the mean time to failure of a homogeneous hot standby repairable system with exponential distribution of its elements' life time and general distribution of their repair time. With the help of the developed general discrete-event simulation model we extend the previous studies to a general case of a homogeneous hot double redundant repairable system with general distributions of both life and repair time lengths of elements.

PROBLEM SETTING AND NOTATIONS

Consider a hot standby repairable system with one repair unit. The elements of the system (units) have exponentially distributed times to failure with parameter α and general repair time distribution $B(x)$. Through-

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

out the paper we will use a generalization of Kendall's notation [11] for queuing systems. In this notation the symbols $\langle GI_n|GI|m\rangle$ stand for a closed system, i.e. a system where the flow of customers is generated by a finite number n of sources that is shown by index in the first position. Symbol GI means "General Independent" and in the first position of this notation it denotes the general distribution of independent life times of the elements of the system and in the second one — the general distribution of their independent repair times. These symbols can be substituted by M for exponential ($exp(\cdot)$), Erlang $E(\cdot,\cdot)$, Gnedenko-Weibull ($GW(\cdot,\cdot)$) with appropriate parameters or any other symbol describing the distribution of life and/or repair time. Finally, the last factor m denotes the number of repair units in the system. In the current paper we consider a simple hot double redundant model, namely $\langle M_2|GI|1\rangle$ and study its reliability function (RF) under different distributions.

The cumulative distribution functions (CDF) of the random life time A and random repair time B are denoted respectively by $A(x)$ and $B(x)$. We suppose the existence of the corresponding probability density functions (PDF), which are denoted by $a(x) = A'(x)$ and $b(x) = B'(x)$. The mean time between failures, the mean service (repair) time, the failure and repair hazard functions are denoted as follows:

$$a = \int_0^\infty (1 - A(x))dx, \text{ and } b = \int_0^\infty (1 - B(x))dx.$$

and

$$\alpha(x) = \frac{a(x)}{1 - A(x)}, \text{ and } \beta(x) = \frac{b(x)}{1 - B(x)}.$$

Define also the moment-generating functions (m.g.f.) of life A and repair B times, the Laplace-Stiltjes transforms (LST) of their distributions by the following expressions:

$$\tilde{a}(s) = \int_0^\infty e^{-sx} a(x)dx$$

and

$$\tilde{b}(s) = \int_0^\infty e^{-sx} b(x)dx, \; Re[s] \geq 0.$$

Life and repair times are assumed independent. The "up" (working) states of each unit and the system will be marked by 0 and the "down" (failed) states – by 1. Under considered assumptions the system behavior can be described by a random process which takes values from the system state space $E = \{0, 1, 2\}$. Denote by $J(t)$ the random process, describing the system behavior: $J(t) = i$ if the system is in state i. At that the system state subset $E_0 = \{0, 1\}$ represents its working (up) states, and the subset $E_1 = \{2\}$ represents the system failure (down) state. The process $J = \{J(t), t \geq 0\}$ represents the system states, and takes the values from E_0 if at least one of the system components is in up state and takes the value 1 otherwise. The corresponding system state probabilities are denoted by $\pi_0(t)$, $\pi_1(t; x)$, $\pi_2(t)$.

We are interested in studying the system lifetime T, which is the duration of time when at least one unit is working. Thus the system lifetime T can be represented as follows: $T = \inf\{t : J(t) = 2\}$ and the system reliability function as

$$R(t) = \mathbf{P}\{T \leq t\} = \mathbf{P}\{J(\tau) \in E_0, \; \tau \in (0, t)\} \qquad (1)$$

ANALYTICAL RESULTS

Using a standart method of comparing the state probabilities at time instants t and $t + \Delta$, we develop the following system of Kolmogorov forward partial differential equations for these probabilities:

$$\begin{aligned}
&\frac{d}{dt}\pi_0(t) = -2\alpha\pi_0(t) + \int_0^t \pi_1(t, u)\beta(u)du, \\
&\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right)\pi_1(t; x) = -(\alpha + \beta_1(x))\pi_1(t; x), \\
&\frac{d}{dt}\pi_2(t) = \alpha \int_0^t \pi_1(t; u)du.
\end{aligned}$$

with boundary and initial conditions

$$\pi_1(t, 0) = 2\alpha\pi_0(t), \quad \pi_0(0) = 1.$$

Theorem. Laplace transform of the reliability function for the considered system has the following form:

$$\tilde{R}(s) = \frac{1}{s} - \tilde{\pi}_2(s) = \frac{s + \alpha + 2\alpha(1 - \tilde{b}(s + \alpha))}{(s + \alpha)[s + 2\alpha(1 - \tilde{b}(s + \alpha))]} \qquad (2)$$

which coincides with the results obtained before in [7], [9].

Corollary. The mean time to failure of the system under consideration equals

$$m = \mathbf{E}[T] = \tilde{R}(0) = \frac{1 + 2(1 - \tilde{b}(\alpha))^2}{2\alpha(1 - \tilde{b}(\alpha))} \qquad (3)$$

SIMULATION RESULTS

In this section we present the results of simulation of a homogeneous two-unit cold standby repairable system $\langle GI_2|GI|1\rangle$ with one repair unit and general distributions of both life and repair times of its elements.

A. General simulation model

We perform the simulation using the discrete event modeling method. We consider the functioning of the system being modeled as a sequence of operations being performed across entities (events). The simulation model is specified graphically as a process flowchart (see Fig. 1).

In order to ensure the precise understanding and reproducibility of the simulation model, we present an algorithm for a simulation process which is represented in the form of pseudocode (see Algorithm 1).

B. Comparison of analytical solution and simulation results for $\langle GI_2|GI|1\rangle$

In this section we perform the comparison of analytical (where possible) and simulation results for the reliability of the considered system. As a model parameter

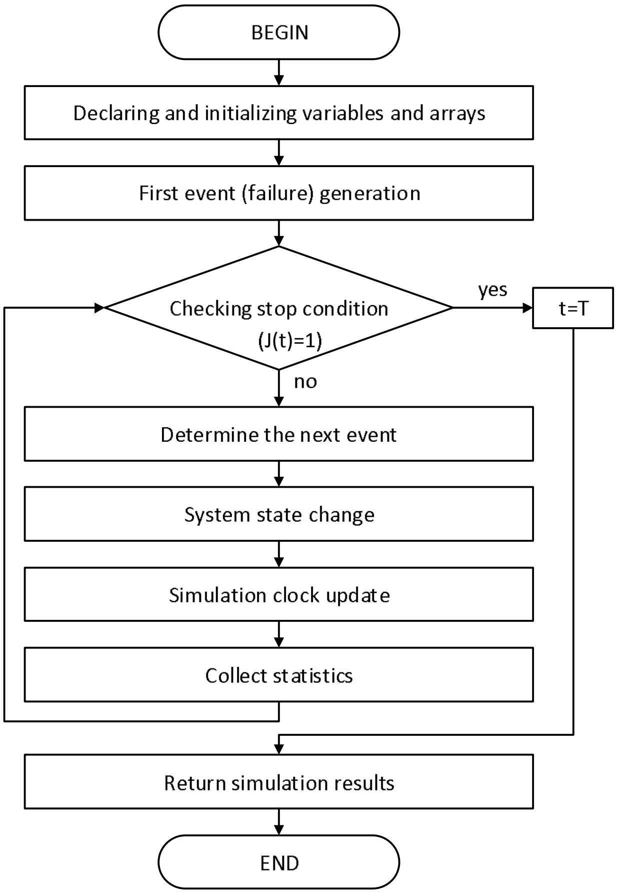


Fig. 1. Flowchart of the discrete-event simulation model

```
Input: $a$, $b$, $T$, '$GI^{(1)}$', '$GI^{(2)}$'
$a$ – mean life time of an element, $b$ – mean repair time, $T$ – maximum simulation time, '$GI^{(1,2)}$' denote CDFs of life and repair time lengths, respectively.
Output: empirical reliability function $R_{empir}$, system mean time to failure $ET$
begin
    double $t := 0.0$;
    int $i := 0$; $j := 0$;
    double $t$_nextfail:= 0.0;
    double $t$_nextrepair:= 0.0;
    int $k := 1$;
    array $r[\,] := [0, 0, 0]$;
    $s := df\_Exp(\frac{1}{a})$;
    $t\_nextfail := t + s$;
    while $t < T$ do
        if $i = 0$ then
            $t\_nextrepair := \infty$;
            $j := j + 1$; $t := t\_nextfail$;
        else if $i = 1$ then
            $s_1 := df\_GI1(..)$;
            $s_2 := df\_GI2(..)$ $t\_nextfail := t + s_1$;
            $t\_nextrepair := t + s_2$;
            if $t\_nextfail < t\_nextrepair$ then
                $j := j + 1$; $t := t\_nextfail$;
            else
                $j := j - 1$; $t := t\_nextrepair$
            end
        end
        else
            $i = 2$; $t\_nextfail := \infty$;
            $j := j - 1$; $t := t\_nextrepair$;
        end
        if $t > T$ then
            $t := T$
        end
        $r[k] := [t, i, j]$;
        $i := j$; $k := k + 1$;
    end
    Evaluate duration of overall time spent in working states;
    Calculate estimates of the reliability function and the system mean life time.
end
```

Algorithm 1: Pseudocode for the simulation process of $\langle GI_2|GI|1\rangle$ model

we consider the value $\rho = \frac{\mathbf{E}[A]}{\mathbf{E}[B]}$, which can be interpreted as a relative rate of system recovery [12]. It will be shown that as $\rho \to \infty$ the sensitivity of the model to shapes of input distributions becomes negligible. Distributions that we've used in our experiments include, but are not limited to the following ones: Exponential (Exp(α)), Erlang (E(k, α)), Gnedenko-Weibull (GW) and Pareto (P). The simulation time has been chosen equal to $T = 10000$ and number of replications equal to 100 and 1000 (for more precise simulation results).

For illustrative purposes we conduct this comparison graphically at Figure 2 and Figure 3 where results of simulation are represented for different distributions $GI^{(1)}$, $GI^{(2)}$ and $\rho = \frac{5}{5} = 1$. In all cases the parameters of distributions have been chosen so that the value of $\mathbf{E}[B]$ remained fixed ($\mathbf{E}[B] = 5$) and the mean time to failure of an element would ascend $\mathbf{E}[A] = \rho\mathbf{E}[B]$ according to the values of ρ. Instead of parameters of distributions the coefficient of variation c (the ratio of the standard deviation to the mean) is indicated in parentheses in the legends of all figures.

Figure 4 depicts the curves of the system reliability function $R(t)$ under rare failures of system elements. As it can be seen from the figure, the differences between both simulation and analytical curves become indistinguishable very quickly. At a relatively small value of $\rho = 10$ the reliability function curves for are already very close to each other for all considered special cases of the $\langle GI_2|GI|1\rangle$ model. The observed behavior is fairly expected, as it was proved by B.V. Gnedenko and A.D. Solov'ev [5],[6]. What is more important and interesting — is that we can assess the rate of this convergence and it will be done in the full-text version of the current paper with the means of quantiles for the given reliability level.

CONCLUSION

The problem of analytical calculation and simulation assessment of reliability function for a homogeneous hot

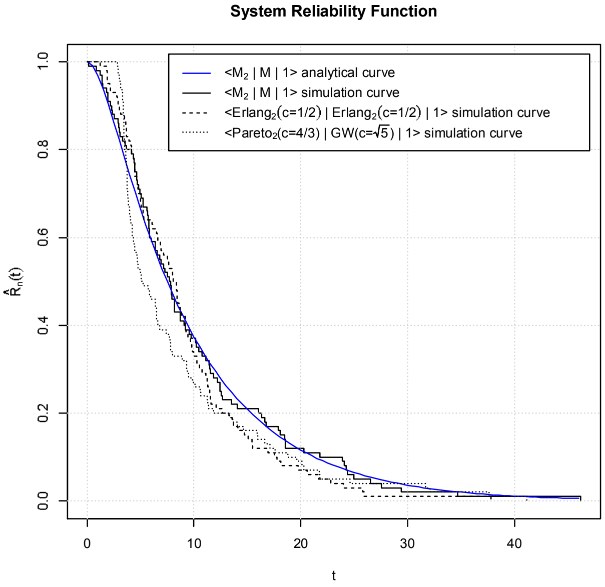


Fig. 2. Empirical and analytical values of $R(t)$ versus time t (values based on 100 replications), $\rho = 1$

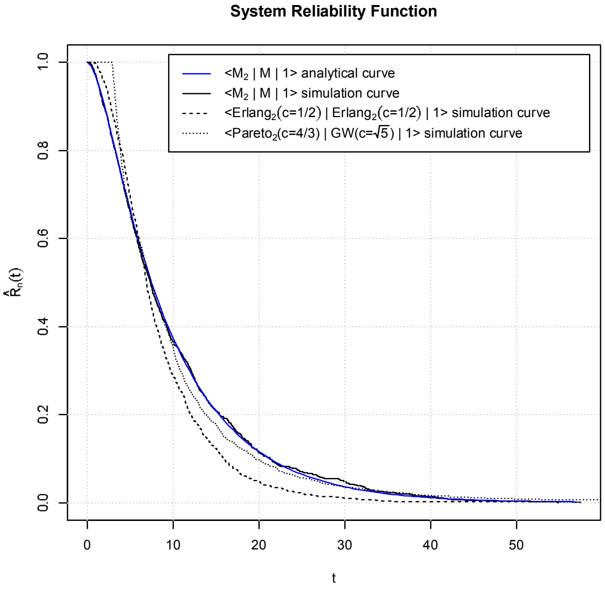


Fig. 3. Empirical and analytical values of $R(t)$ versus time t (values based on 1000 replications), $\rho = 1$

double redundant repairable system has been considered. Explicit representation of this function in terms of Laplace transform has been found. It was shown that under rare failures of system elements the reliability function is approximated with exponential distribution with appropriate mean life time. Also analysis of the obtained results shows that the results of exact analytical calculation (where possible) and simulation results have close agreement.

ACKNOWLEDGMENTS

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008), and by the Russian Foundation for Basic Research according to the research projects No. 17-07-00142 and No. 17-01-00633.

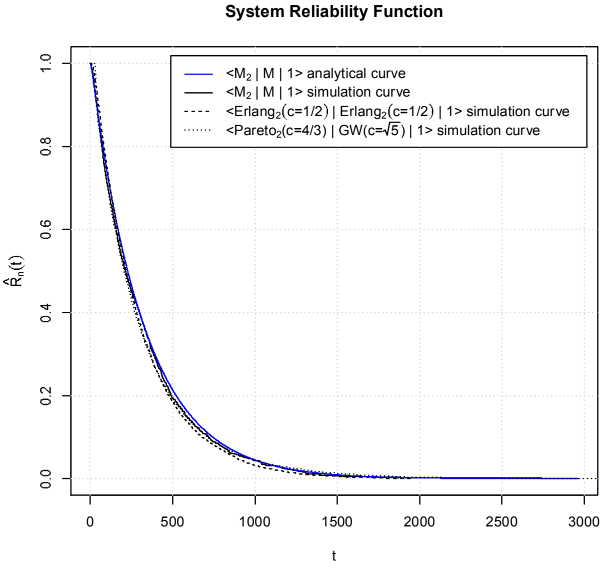


Fig. 4. Empirical and analytical values of $R(t)$ versus time t (values based on 1000 replications) under rare failures of system elements, $\rho = 10$

References

[1] B.V. Gnedenko, Yu.K. Belyaev, and A.D. Solovyev, Mathematical Methods of Reliability Theory, translation edited by R. E. Barlow // Academic Press, New York, 1969.

[2] S.K. Srinivasan, M.N. Gopalan. Probabilistic Analysis of a Two-Unit Syatem witha Warm Standby ans a Single Repair Facility. // Oper. Res. Vol 21, No 3 (May-June 1973), pp. 748-754.

[3] V. Rykov. Multidimensional Alternative Processes as Reliability Models. Modern Probabilistic Methods for Analysis of Telecommunication Networks. (BWWQT 2013) Proceedings. Eds: A.Dudin, V.Klimenok, G.Tsarenkov, S.Dudin. Series: CCIS 356. Springer, 2013. P.147-157.

[4] B.V. Gnedenko. On cold double redundant system. // Izv. AN SSSR. Texn. Cybern. No. 4 (1964), P. 312. (In Russian)

[5] B.V. Gnedenko. On cold double redundant system with restoration. // Izv. AN SSSR. Texn. Cybern. No. 5 (1964), P. 111 - 118. (In Russian)

[6] A.D. Solov'ev. On reservation with quick restoration. // Izv. AN SSSR. Texn. Cybern. No. 1 (1970), P. 5671. (In Russian)

[7] V. Rykov, Tran Ahn Ngia. On sensitivity of systems reliability characteristics to the shape of their elements life and repair time distributions. // Vestnik PFUR. Ser. Mathematics. Informatics. Physics. No.3 (2014), P. 65-77. (In Russian)

[8] D.Efrosinin, V.Rykov. Sensitivity Analysis of Reliability Characteristics to the Shape of the Life and Repair Time Distributions. In: Information Technologies and Mathematical Modelling. . Eds by Alexander Dudin, Anatoly Nazarov, Rafael Yakupov, Alexander Gortsev (13-th International Scientific Conference ITMM 2014 named after A.F. Terpugov, Anzhero-Sudzhensk, Russia, November 20-22 2014. Proceedings). Communication in Computer and Information Science, Vol. 487, pp. 101-112.

[9] Dmitry Efrosinin, Vladimir Rykov and Vladimir Vishnevskiy. Sensitivity of Reliability Models to the Shape of Life and Repair Time Distributions. (9-th International Conference on Availability, Reliability and Security (ARES 2014), p.430-437. Published in CD: 978-I-4799-4223-7/14, 2014, IEEE. DOI 10.1109/ARES 2014.65.

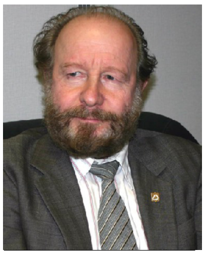
[10] Petrovsky I. G. Lections on theory usual differential equations. M.-L.: GITTL. 1952. 232p. in Russian

[11] D.G. Kendall. Stochastic processes occurring in the theory

of queues and their analysis by the method of embedded Markov chains. Annals of Math. Stat., 1953, Vol. 24, P. 338-354.

[12] D.V. Kozyrev. Analysis of Asymptotic Behavior of Reliability Properties of Redundant Systems under the Fast Recovery // Bulletin of Peoples Friendship University of Russia. Series ‘’Mathematics. Information Sciences. Physics” No.3 (2011), pp.49–57. (In Russian)

AUTHOR BIOGRAPHIES



VLADIMIR V. RYKOV obtained his D.Sc. in Phys.-Math. from the Moscow State University in 1990. Since then he has worked as full professor at the Applied Mathematics and Computer Modeling Department of the Gubkin Russian State Oil and Gas University and full professor at the Department of Applied Probability and Informatics, RUDN University. His main interests include controllable queueing systems and reliability theory. He is full member of International Academy of Informatization since 1994. Recent Affiliations: (1) Moscow Mathematical Society, Probability Theory Section since 1968; (2) Russian Actuarial Society since 1995; (3) Kappa Mu Epsilon, National Mathematics Honor Society, Michigan Epsilon Chapter since 2001. He is editor-in-chief of the electronic journal “Reliability: Theory and Applications”, a member of the editorial board of the journal ”Automatic Control and Computer Sciences” and a member of the doctoral council of the Gubkin Russian State Oil and Gas University in ”Mathematical modeling”. Author of more than 250 scientific works, including 13 monographs, and 5 translated books.
His e-mail address is *vladimir_rykov@mail.ru*

DMITRY V. KOZYREV received his Ph.D. degree in Physics and Mathematics in 2013. Since 2007 he has been working at the Department of Probability Theory and Mathematical Statistics of RUDN University and since 2013 he is an associate professor at the Department of Applied Probability and Informatics, RUDN University, and a senior research fellow at V.A.Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia. His current research activities include modeling and simulation of repairable systems, reliability assessment and performance analysis, sensitivity analysis, Reliable Internet of Things.
His e-mail address is *kozyrev_dv@rudn.university*

ELVIRA R. ZARIPOVA received her B.Sc. and M.Sc. degrees in applied mathematics at RUDN University in 2001 and 2003, respectively. In 2015, she received her Ph.D. degree in applied mathematics and computer sciences from Peoples Friendship University of Russia (RUDN University). Since 2003, Elvira Zaripova works at the Telecommunication Systems Department of RUDN University. She is currently an Associate Professor at the Department of Applied Probability and Informatics. Her current research interests lie in the area of performance analysis of radio resource management techniques in LTE networks on which she has published several papers in refereed journals and conference proceedings.
Her e-mail address is: *zaripova_er@rudn.university.*

MODELING AND RESPONSE TIME ANALYSIS FOR WEB BROWSING UNDER INTERRUPTIONS IN LTE NETWORK

Evgeny Mokrov*, Eduard Sopin*†, Ekaterina Markova*, Dmitry Poluektov*,
Irina Gudkova*†, Pavel Masek‡, Jiri Hosek‡

*Peoples' Friendship University of Russia (RUDN University)

Applied Informatics and Probability Theory Department
Miklukho-Maklaya St 6, Moscow, 117198, Russian Federation
{ evmokrov, sopin_es, markova_ev, gudkova_ia}@pfur.ru
poluektov_dmitri@mail.ru

†Institute of Informatics Problems, Federal Research Center "Computer Science and Control"

Vavilova str., Moscow, 119333, Russia
{sopin_es, gudkova_ia}@pfur.ru

‡ Brno University of Technology

Department of Telecommunications
Technicka 3082/12, Brno, Czech Republic
{hosek, masekpavel}@feec.vutbr.cz

KEYWORDS

Interruption, queuing system, unreliable server, processor sharing, probability distribution, queue sojourn time, service sojourn time, web browsing.

ABSTRACT

The modern development of the telecommunication technologies is closely related with the rapid data traffic growth problem under the conditions of limited frequency resources for mobile networks. One of the possible solutions to this problem is the utilization of the Licensed Shared Access (LSA) framework, which is developed with the assistance of ETSI. The LSA spectrum is shared between the owner (incumbent) and LSA licensee (e.g., mobile network operator). At any time, LSA spectrum could be used by incumbent or mobile network operator but not together at once. In this connection, if the incumbent needs its frequency, then LSA band becomes unavailable for mobile operator. This leads to service interruptions for mobile operator users. Service interruption has a significant impact both on quality of service (QoS) and quality of experience (QoE). In this regard, we propose the model of the LTE network cell with LSA framework as a finite capacity queuing system with unreliable server. The important QoS parameters of the model are moments of queue and service sojourn time. We derived formulas and provide numerical analysis of these performance measures.

INTRODUCTION

According to Cisco Systems analysts (Cisco Systems 2015), the volume of traffic generated in mobile networks will be about 24.3 exabytes per month by the year 2019. Due to the rapid growth of data traffic, the problem of frequency resources shortage appears. In this regard, mobile operators have to search the additional resources to provide services for users with the required level of Quality of Service (QoS). One of the solutions proposed by ETSI is the usage of Licensed Shared Access (LSA) framework (ETSI TR 103 113. 2013; ETSI TR 103 154 2014; ETSI TR 103 235. 2015).

Key players of LSA are incumbent, national regulator and LSA licensee (ETSI TR 103 113. 2013). LSA framework enables controlled access to the shared spectrum between owner (so called incumbent) and limited number of LSA licensees (e.g., telecommunication operators). The incumbent has absolute priority to decide whether the mobile operator will gain access to utilize the frequency bands. The usage rules are based on agreements between key players, according to requirements for QoS and Quality of Experience (QoE) parameters.

There are different models of mobile networks with LSA framework. From the viewpoint of the mobile network operator who uses the LSA framework, there are models with one frequency band, belonging to the incumbent and the mobile operator (Borodakiy et al. 2014; Gudkova et al. 2016), and models with two frequency bands, which also consider the part of the spectrum, belonging only to the mobile operator (Gudkova et al. 2015; Samouylov et al. 2016).

In the paper we propose a model of LTE network with LSA framework and shutdown policy (Borodakiy et al. 2014; Gudkova et al. 2015), which allows coordinate interference between incumbent and mobile operator. We consider only one frequency band with tolerant to delays traffic. We simulate a queuing system with unreliable server, finite buffer and Processor sharing (PS) discipline. Queuing systems with PS discipline have been widely studied and are used to describe the bandwidth sharing in computer and communication systems (Fredj et al. 2001; Al-Begain et al. 2003; Kleinrock 1964; Kleinrock 1967; Yashkov 1987; Yashkov and Yashkova. 2007; Morrison 1985). However, almost all of these works offer analysis for infinite capacity models. An important characteristic of these models is average mean sojourn time (Zhen and Knessl 2009; Rege and Sengupta 1985), to calculate

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

which Nunez-Queija (Nunez-Queija 2000; Nunez-Queija 2001) has derived the Laplace transform in a system where the on-periods and the off-periods form an alternating renewal process. Analysis of finite buffer PS queuing systems with waiting could be found only in a few works (Knessl 1990; Zhen and Knessl 2013; Dudin et al. 2015). However, to make the mathematical model applicable for analysis of wireless network characteristics, we introduced a threshold on number of simultaneously served customers, since there are limited frequency-time resources in the network. Earlier, we provided analysis of the moments of customer sojourn time in the queue for two types of arrival processes: Poisson arrivals (Samouylov et al. 2016) and Markov Modulated Poisson Process (MMPP) arrivals (Samouylov et al. 2017). In this work, we continue to study the main QoS parameters of model with a Poisson distribution, namely, we analyze moments of queue and service sojourn time.

The paper is organized as follows. We consider a model of LTE network with LSA framework and tolerant to delays traffic as a queuing system with unreliable server, finite buffer, PS discipline, and threshold on the number of simultaneously served customers, also in this section we propose formulas for calculating the main performance measures of this model. Then we conduct a numerical analysis of the QoS parameters by the example of web browsing in wireless networks under interruptions. Finally, we conclude the paper.

MATEMATICAL MODEL FOR SYSTEM WITH UNRELIABLE SERVER

Queuing Model

We consider a single-server queuing system with finite capacity r and unreliable server with exponentially distributed on- and off-period durations with rates α and β respectively. Customers arrive according to a Poisson process with rate λ. Upon the arrival of a customer, it is placed in the queue. The server works according to PS discipline serving a maximum of N customers at once $0 < N \leq r$, the service rate is exponential with parameter μ. Customers are served according to, but no more than N. It means, if there are n customers in the system, only $n *= \min\{n, N\}$ of them are served, and the rest $\max\{n - n*, 0\}$ customers wait in the queue. If upon the arrival of a customer, the system is already full, then the customer is lost.

Behavior of the queue with unreliable server can be described by a Markov process $\xi(t)$ with generator matrix $\mathbf{Q}$, described in (Samouylov et al. 2016)

Stationary distribution vector of $\xi(t)$ $\mathbf{q}^T = (q_{00}, q_{10}, \dots, q_{r0}, q_{01}, q_{11}, \dots, q_{r1})$ can be acquired as the solution of system of equilibrium equations $\mathbf{q}^T\mathbf{Q} = \mathbf{0}, \mathbf{q}^T\mathbf{1} = 1$, where $\mathbf{1}$ is vector of ones with appropriate size $(2r + 2)$ (Samouylov et al. 2016).

Queue and Service Sojourn Time

In this paragraph, we study average response time and average direct download delay. For the considered Web browsing scenario, average response time corresponds to the average time from the moment when a UE requests to browse a webpage to the moment when the first element of the requested page appears on the screen and calculated as queue sojourn time. Average direct download delay corresponds to the average time from the moment when the first element of the requested webpage appears on the screen to the moment when the page is completely rendered and displayed by the browser. Average direct download delay is calculated as service sojourn time.

First let us study queue sojourn time. It has a phase type cumulative distribution function (CDF) $F(x)$ (Asmussen, 2003). To find it, let us introduce a continuous-time Markov chain $\chi(t)$ that describes the behavior of the queue from arrival to start of service of a particular customer. The Markov chain has an absorbing state ω, which is reached when the considered customer enters service. To take into account the influence of the order of customers arriving during off-periods on sojourn time, we added a third component to the Markov chain $\chi(t)$. Besides the absorbing state, there are also states (n, k, s), $k = 1, \dots, r - N, n = k + N, \dots, r, s = 0,1$ to calculate queue sojourn time Here, as in the previous section, s indicates the on-periods $(s = 1)$ and the off-periods $(s = 0)$ and n is the number of customers in the systems. The second component $k > 1$ indicates that the customer has $(N + k)$-th spot in the queue.

Taking into account the order of states, generator matrix $\mathbf{A}$ of $\chi(t)$ becomes

$$\mathbf{A} = \begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{0} & 0 \end{bmatrix}, \tag{1}$$

where $\boldsymbol{c}$ is an exit vector to the absorbing state,

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{00} & \mathbf{C}_{01} \\ \mathbf{C}_{10} & \mathbf{C}_{11} \end{bmatrix}, \mathrm{c} = \begin{bmatrix} \mathbf{0} \\ \mathbf{c_0} \end{bmatrix}. \tag{2}$$

Blocks $\mathbf{C}_{01}$ and $\mathbf{C}_{10}$ are diagonal matrices, $\mathbf{C}_{01} = \alpha\mathbf{I}, \mathbf{C}_{10} = \beta\mathbf{I}$, while $\mathbf{C}_{00}$ and $\mathbf{C}_{11}$ have the following form:

$$\mathbf{C}_{11} = \begin{bmatrix} \mathbf{A}_1 & & & & \\ \mathbf{B}_2 & \mathbf{A}_2 & \cdots & \mathbf{0} & \\ & \mathbf{B}_3 & & & \\ \vdots & & \ddots & \vdots & \\ & & & \mathbf{A}_{\mathrm{r-N-1}} & \\ \mathbf{0} & & \cdots & \mathbf{B}_{\mathrm{r-N}} & \mathbf{A}_{\mathrm{r-N}} \end{bmatrix}, \tag{3}$$

$$\mathrm{C}_{00} = \mathrm{diag}(\mathrm{D_k}), \mathrm{k} = 1, \dots, \mathrm{r} - \mathrm{N}, \tag{4}$$

Matrices $\mathbf{A}_k$, $\mathbf{D}_k, k = 1, \dots, r - N$ and $\mathbf{B}_k$, $k = 2, \dots, r - N$ have the same structure as the ones in (Samouylov et al. 2016). The order of matrices $\mathbf{A}_k$ and $\mathbf{D}_k$ is $(r + 1 - N - k) \times (r + 1 - N - k)$. The order of matrices $\mathbf{B}_k$ $(r+1-N-k)\times(r+2-N-k)$. Vector $\mathbf{c}_0$ is

$$\mathbf{c}_0^{\mathrm{T}} = [\mu \quad \cdots \quad \mu \quad 0 \quad \cdots \quad 0], \tag{5}$$

where there are exactly $(r-N)$ rows of μ. That corresponds to the order of the matrix $\mathbf{A}_1$.

According to PASTA property (Wolff 1982), stationary distribution of $\xi(t)$ and stationary distribution of the Markov chain $\xi(t_n - 0)$, embedded at the moments t_n, $n = 1,2,\ldots$ just before the arrival, are equal. Consequently, blocking probability is given by $\pi = q_{r0} + q_{r1}$.

Proposition 1. The initial probability distribution $\boldsymbol{\psi}$ of a Markov chain $\chi(t)$ has the following form:

$$\boldsymbol{\psi}(\mathrm{n},\mathrm{k},\mathrm{s}) = \begin{cases} \frac{\mathrm{q(n-1,s)}}{1-\pi}, \text{if } \mathrm{k = n - N, N < n \le r} \\ \quad 0, \text{ otherwise} \end{cases}. \tag{6}$$

Proof. Initial distribution $\boldsymbol{\psi}$ equals to the distribution of a Markov chain $\xi(t_n + 0)$ embedded at the moments just after the arrival with the addition of appropriate j component.

Proposition 2. The mean queue sojourn time in the Markov chain $\xi(t)$ is

$$\mathrm{m_q} = \boldsymbol{\psi}\mathrm{u}, \tag{7}$$

where nonnegative vector $\mathbf{u} = -\mathbf{C}^{-1}\mathbf{1}$ is a unique solution of the system of equations $\mathbf{Cu} = -\mathbf{1}$, $\mathbf{1}$ is a vector of ones.

Proof. CDF $F(x)$ of sojourn time is equal to CDF of time before absorption in the Markov chain $\chi(t)$ with the initial distribution $\boldsymbol{\psi}$. Therefore, Laplace-Stieltjes Transform (LST) of CDF $F(x)$ is $F^*(s) = \boldsymbol{\psi}(s\mathbf{I} - \mathbf{C})^{-1}\mathbf{c}$ (Asmussen, 2003) and the average value is $m_q = -\boldsymbol{\psi}\mathbf{C}^{-1}\mathbf{1}$.

Denote $\mathbf{E}_k = -\mathbf{D}_k^{-1}$, put vector $\mathbf{u}$ in form $\mathbf{u} = (\mathbf{u}_{10}, \mathbf{u}_{20}, \ldots, \mathbf{u}_{\mathrm{r-N},0}, \mathbf{u}_{11}, \mathbf{u}_{21}, \ldots, \mathbf{u}_{r-N,1})$ and after some simplifications we derive the following recurrent algorithm for the solution of the system of equations $\mathbf{Cu} = -\mathbf{1}$, that decreases computational complexity:

```
$\mathrm{m} = 0;$
$\mathbf{u} = -(\mathbf{A}_1 + \alpha\cdot\beta\cdot\mathbf{E}_1)^{-1}\cdot(\mathbf{1}_{[(r-N)\times 1]} +$
    $+\alpha\cdot\mathbf{E}_1\cdot\mathbf{1}_{[(r-N)\times 1]});$
$\mathrm{m} = \mathrm{m} + \mathbf{u}\cdot q_{N1}/(1-\pi);$
$\mathbf{u1} = \mathbf{E}_1\cdot(\mathbf{1}_{[(r-N)\times 1]} + \beta\cdot\mathbf{u});$
$for\ k = 2{:}\,(r-N)$
$\mathrm{m} = \mathrm{m} + \mathbf{u1}_{\mathrm{N}}\cdot q_{N1}/(1-\pi);$
$for\ k = 2{:}\,(r-N)$
$begin$
    $\mathbf{u} =\text{-}\,(A_k - \alpha\cdot\beta\cdot\mathbf{E}_k)^{-1}\cdot(\mathbf{1}_{[(r-N-k+1)\times 1]} -$
        $-\alpha\cdot\mathbf{E}_k\cdot\mathbf{1}_{[(r-N-k+1)\times 1]} + \mathbf{B}_k\cdot\mathbf{u});$
    $\mathrm{m} = \mathrm{m} + \mathbf{u}_1\cdot q_{N+k,0}/(1-\pi);$
    $\mathbf{u1} = (\text{-}\,\mathbf{E}_k\cdot(\mathbf{1}_{[(r-N-k+1)\times 1]} + \beta\cdot\mathbf{u});$
    $\mathrm{m} = \mathrm{m} + \mathbf{u1}_1\cdot q_{N+k,1}/(1-\pi);$
$end$
```

Thus, the presented algorithm lets us calculate queue sojourn time. Note that having LST of sojourn time CDF, we can obtain not only mean value, but also its variance and higher-order moments.

Also in this paper we calculate service sojourn time as the difference between the system sojourn time and queue sojourn time, thus service sojourn time can be represented as

$$\mathrm{m_s} = \mathrm{m} - \mathrm{m_q}, \tag{8}$$

where m_q is the queue sojourn time derived above and m is the system sojourn time, derived in (Samouylov et al. 2016).

NUMERICAL ANALYSIS FOR WEB BROWSING PERFORMANCE MEASURES

The spectrum allocated for specific applications, for example, cordless cameras, portable video links, mobile video links, and terrestrial or aeronautical telemetry could be used for the shared access between an incumbent and a mobile operator according to the ETSI recommendation (ETSI TR 103 113. 2013). In this paper we consider the case of aeronautical telemetry. In this connection, let us assume the server on-period duration α^{-1} is equal to time when airport (incumbent) does not need a frequency band for telemetry. The server off-period β^{-1} duration is equal to time when the airport needs a frequency band for communicating airplanes with air traffic control (ATC).

Let users are stationary and perform web browsing (ITU-T G.1030. 2014). We assume the average exponential webpage size is equal to b.

The waiting time before the download starts is the average time from the moment when user requests to browse a webpage to the moment when the first element of the requested page appears on the screen and corresponds to queue sojourn time within our queuing system. The value of average response time is limited to a threshold T_r which is equal to 2 seconds according to (ITU-T G.1010. 2001). The average direct download delay is the average time from the moment when the first element of the requested webpage appears on the screen to the moment when the page is completely rendered and displayed by the browser and corresponds to service sojourn time.

According to ITU recommendation (ITU-T M.2370. 2015), web browsing is about 9% of the total mobile traffic. Therefore, our numerical experiment does not consider the overall downlink channel peak bit rate and deals with the part of it that corresponds to web browsing.

The average time when airplane flights over the cell, i.e. the average time when the server is off, approximately is equal to 60 s. The average download time, defined as the sum of the average response time and average direct download delay, is significantly less than 60 s and limited to a threshold T_d of 4 seconds (ITU-T G.1010. 2001). In this connection interruptions under LSA operation have practically no influence on the webpage download time, therefore, we consider for the numerical example the channel availability and unavailability time periods that describe interruptions owing to the higher priority applications

Table 1 summarizes the initial data for the considered example. This data is the same as in (Samouylov et al. 2016) to ensure that the results acquired in this work usefully complement the ones obtained previously.

Table 1: System parameters

Parameter description	Notation	Value
Overall downlink channel peak bit rate	C_1	See Tab. 2
Average time when channel is available	α^{-1}	10 s
Average time when channel is unavailable	β^{-1}	2 s
Number of UEs within the cell	N_{UEs}	20 ÷ 70
Average webpage size	b	See Tab. 3
Threshold on the webpage response time	T_r	2 s
Threshold on the webpage download time	T_d	4 s
Average time when UE reads the webpage	Δ_R	30 s

From this table, considering that web browsing only takes 9% of the total channel, average service time can be obtained as $\mu^{-1} = \frac{b}{0.09 \cdot C_1}$. The queue length and number of servers can be acquired as $r = \mu * T_r$ and $N = \mu * (T_d - T_r)$ correspondingly. The arrival rate can be acquired as $\lambda = N_{UEs}/(\Delta_R + T_d)$, since each customer reads page on average Δ_R s after it read the previous one and it take an average of T_d s for it to download a page.

Table 2. System parameters: downlink channel peak bit rate (Motorola White paper, 2009)

Antenna technology	Downlink channel peak bit rate		
	5 MHz	10 MHz	20 MHz
MIMO 2x2	43 Mbps	86 Mbps	173 Mbps
MIMO 4x4	82 Mbps	163 Mbps	326 Mbps

Table 3. System parameters: average webpage size (HTTP Archive, 2016)

Caching	URLs		
	All	Top 1000	Top 100
Without cache (total size)	2296 KB	2017 KB	1257 KB
With cache (without images)	839 KB	890 KB	493 KB

Figures 1 and 2 depict the scenario with different downlink channel peak bit rates according to the table 2, and Figures 3 and 4 depict the scenario with different average webpage sizes (see Table 3).

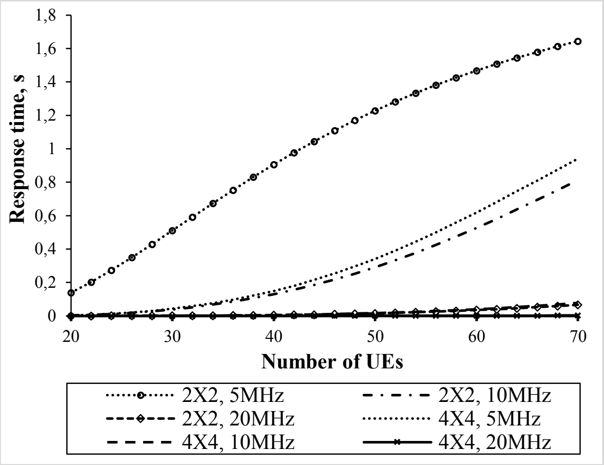


Figure 1: Average response time vs. downlink channel peak bit rate C1 (for top 100 URL webpage with cache)

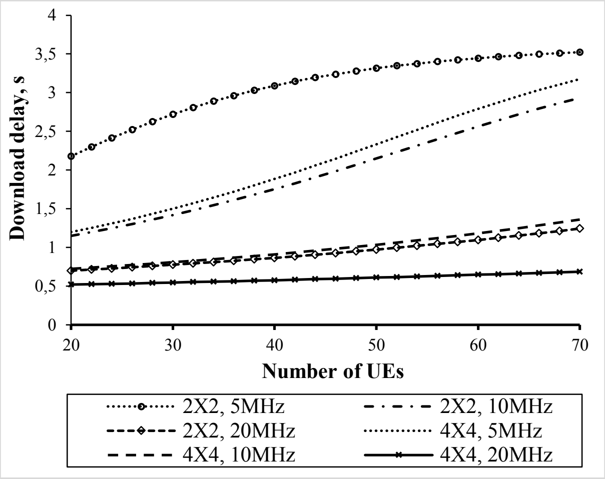


Figure 2: Average direct download delay vs. downlink channel peak bit rate C1 (for top 100 URL webpage with cache)

On plots we can see, that for 20 Mhz MIMO, browsing top 100 URL webpages with caching (size 493 KB) both delay and response time always lie within the threshold of 1.5 s and 0.5 s respectively. That also corresponds to the plot in (Samouylov et al. 2016), where download time for these parameters was shown to be within the threshold of 2 s as well, but now we see, that the page is downloaded much faster, compared to its download delay. Cases of 5 MHz MIMO 4x4 and 10 MHz MIMO 2x2 satisfy the threshold of 1 s response time and up to 3 s of delay. 5 MHz MIMO 2x2 initially have quite a large delay of more than 2 s, that further goes up, rendering it unusable for large number of UEs and since the same growth can be seen for response time it can be concluded that even it can’t be used in systems with large number of UEs.

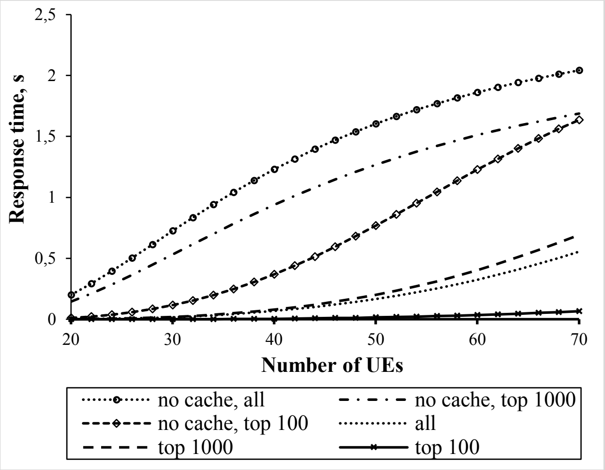


Figure 3: Average response time vs. average webpage size b (for 20 MHz MIMO 2 x 2)

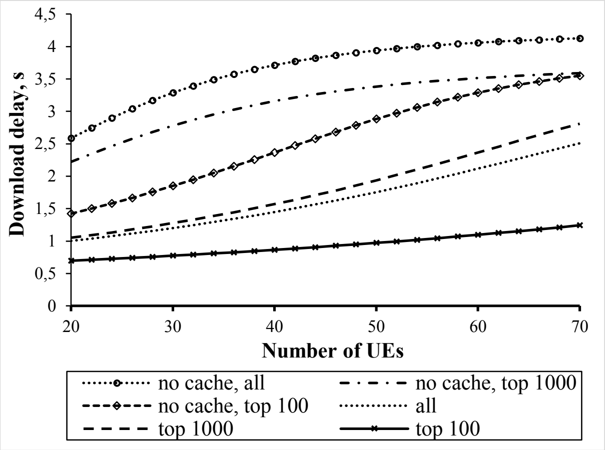


Figure 4: Average direct download delay vs. average webpage size b (for 20 MHz MIMO 2 x 2)

For cells with 20 MHz bandwidth and MIMO 2x2, 100 URL webpages (with cache) will be downloaded almost instantly after a short delay of less than1 s, while all pages (with cache) can take up to 0.5 s to download after a minimum of 1 s delay. For more than 50 UEs the delay will already exceed the preferable threshold. For pages without cash the delay alone would exceed the preferable threshold at more than 30 users for top 100 URL and for more than 50 users even the acceptable threshold can be breached due to growth in response time. For all URL it is impossible to contain the delay within the preferable threshold and the acceptable threshold can only be maintained for less than 30 UEs due to the steep growth of response time.

CONCLUSION

This paper is the continuation of (Samouylov et al. 2016) where a Markov chain based method for the analysis of mean sojourn time in the finite capacity queuing system with an unreliable server and PS discipline was proposed. In this paper, we analyzed service and queue sojourn time in the same system under the same initial data with using a different Markov chain. This lets us get more insight into the webpage download time, studied at the previous article, giving it to us as the response time and delay. Thus, we can acquire time that a page would take to actually be downloaded and displayed in the considered system and time a customer have to wait without downloading any data. We use the developed method for calculating the system and queue sojourn time for the considered queuing system to analyze the page response time and delay and their dependencies on some QoE system influence factors like channel pick bit rate, webpage size, caching.

In future work, average waiting time during the download due to interruptions, i.e. the average sum of time periods within the webpage download time when the download was interrupted due to the channel failures can be studied as well as the average waiting time during the webpage direct download due to interruptions can be studied. This would let us get insight into the dependency of average download time on interruptions. In addition, the acquired parameters can be obtained for different distributions of the arrival flow, such as MAP for example.

ACKNOWLEDGEMENT

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008), RFBR according to the research projects No. 16-37-00421 mol_a and No. 17-07-00142 a. The BUT work was partially supported by National Sustainability Program under grant LO1401, infrastructure of the SIX Center was used.

REFERENCES

Cisco Systems. 2014. Cisco visual networking index: Global Mobile Data Traffic Forecast Update, 2014–2019: usage: White paper.

ETSI TR 103 113. 2013. Electromagnetic compatibility and Radio spectrum Matters (ERM); System Reference document (SRdoc); Mobile broadband services in the 2 300 MHz - 2 400 MHz frequency band under Licensed Shared Access regime.

ETSI TR 103 154. 2014. Reconfigurable Radio Systems (RRS); System requirements for operation of Mobile Broadband Systems in the 2 300 MHz - 2 400 MHz band under Licensed Shared Access (LSA).

ETSI TR 103 235. 2015. Reconfigurable Radio Systems (RRS); System architecture and high level procedures for operation of Licensed Shared Access (LSA) in the 2 300 MHz - 2 400 MHz band.

Gudkova, I.; E. Markova E.; P. Masek P.; S. Andreev; J. Hosek; N. Yarkina N.; Samouylov K., and Y. Koucheryavy. 2016. “Modeling the utilization of a multi-tenant band in 3GPP LTE system with Licensed Shared Access”. In: *8th International Congress on Ultra Modern Telecommunications and Control Systems ICUMT*, 179-183.

Samouylov, K.; I. Gudkova; E. Markova; and N. Yarkina N. 2016. “Queuing model with unreliable servers for limit power policy within Licensed Shared Access framework”. *Lecture Notes in Computer Science* 9870, 404-413.

Borodakiy, V.Y.; K.E. Samouylov; I.A. Gudkova; D.Y. Ostrikova; A.A. Ponomarenko; A.M. Turlikov; and S.D. Andreev. 2014. “Modeling unreliable LSA operation in 3GPP LTE cellular networks”. In: *6th International Congress on Ultra Modern Telecommunications and Control Systems ICUMT*, 490–496.

Gudkova I.A.; K.E. Samouylov; D.Y. Ostrikova; E.V. Mokrov; A.A. Ponomarenko-Timofeev; S.D. Andreev; and Y.A. Koucheryavy. 2015. “Service failure and interruption probability analysis for Licensed Shared Access regulatory framework”. In: *7th International Congress on Ultra Modern Telecommunications and Control Systems ICUMT*-2015, 123-131.

Ben Fredj, S.; T. Bonald; A. Proutiere; G. Regnie; and J.W. Roberts. 2001. “Statistical bandwidth sharing: a study of congestion at flow level”. In *ACM SIGCOMM 2001*, 111-122.

Al-Begain, K.; I. Awan; and D.D. Kouvatsos. 2003. “Analysis of GSM/GPRS cell with multiple data service classes”. *Wireless Personal Communications* 25, 41-57.

Kleinrock, L. 1964. “Analysis of a time-shared processor”. *Naval Research Logistics Q.,* 11, 59-73.

Kleinrock, L. 1967. “Time-shared systems: a theoretical treatment.” *J. ACM* 14, No.2, 242-261.

Yashkov, S.F. 1987. “Processor-sharing queues: some progress in analysis”. *Queueing Systems* 2, No.1, 1-17.

Yashkov, S.F. and A.S. Yashkova. 2007. “Processor sharing: a survey of the mathematical theory”. *Automation and Remote Control*, No.9, 1662-1731.

Morrison, J.A. 1985. “Response-time distribution for a processor-sharing system.” *Applied Mathematics* 45, No.1, 152-167.

Knessl, C. 1990. “On finite capacity processor-shared queues.” *Applied Mathematics* 50, No.1, 264-287.

Zhen, Q. and C. Knessl. 2009. “On sojourn times in the finite capacity M/M/1 queue with processor sharing.” *Operations Research Letters* 37, No.6, 447-450.

Rege, K. and B. Sengupta. 1985. “Sojourn time distribution in a multiprogrammed computer system.” *AT&T* 64, No.5, 1077-1090.

Nunez-Queija, R. 2000. “Sojourn times in a processor sharing queue with service interruptions”. *Queueing Systems* 34, No.1, 351-386.

Nunez-Queija, R. 2001. “Sojourn times in non-homogeneous QBD processes with processor-sharing”. *Stochastic Models* 17, No.1, 61-92.

Zhen, Q. and C. Knessl. 2013. “Asymptotic analysis of spectral properties of finite capacity processor shared queues”. *Studies in Applied Mathematics* 131, No.2, 179-210.

Dudin, A.; C.S. Kim; S. Dudin; and O. Dudina. 2015. “Priority Retrial Queueing Model Operating in Random Environment with Varying Number and Reservation of Servers”. *Applied Mathematics and Computations* 269, 674-690.

Samouylov, K.; V. Naumov; E. Sopin E; I. Gudkova; and S. Shorgin S. 2016. “Sojourn time analysis for processor sharing loss system with unreliable server”. *Lecture Notes in Computer Science* 9247, 284-297.

Samouylov, K.; E. Sopin E; I. Gudkova “Sojourn time analysis for processor sharing loss queuing system with service interruptions and MAP arrivals”. *Distributed Computer And Communication Networks: control, computation, communications (DCCN-2017), in print.*

Asmussen, S. 2003. *Applied Probability and Queues.* Springer, New York.

Wolff, R.W. 1982. “Poisson arrivals see time averages”. *Operational Research* 30, No.2, 223-231.

Motorola. 2009. “Realistic LTE Performance – From Peak Rate to Subscriber Experience”. White paper. http://www.apwpt.org/downloads/realistic_lte_experience_wp_motorola_aug2009.pdf

HTTP Archive. 2016. Interesting Stats. http://httparchive.org/interesting.php?a=All&l=Apr%2015%202016

AUTHOR BIOGRAPHIES

EVGENY MOKROV is a postgraduate student at the Department of Applied Informatics and Probability Theory of Peoples' Friendship University of Russia (RUDN University). He received a master's degree from the PFUR in 2015. Currently, his current research focuses on performance analysis of LSA framework. His e-mail address is: evmokrov@ sci.pfu.edu.ru.

EDUARD SOPIN received his B.Sc. and M.Sc. degrees in applied mathematics from the Peoples’ Friendship University of Russia (RUDN University) in 2008 and 2010, respectively. In 2013, he received his PhD degree in applied mathematics and computer science. Since 2009, Eduard Sopin works at the Telecommunication Systems Department of RUDN University, now he is an associate professor at the Department of Applied Probability and Informatics of RUDN University. His current research interests lie in the area of performance analysis of modern wireless networks and cloud/fog computing. His e-mail address is sopin_es@pfur.ru.

EKATERINA MARKOVA received her B.Sc. and M.Sc. degrees in applied mathematics from the Peoples’ Friendship University of Russia in 2009 and 2011, respectively. In 2015, she received her Ph.D. degree in applied mathematics and computer sciences from the RUDN University. Since 2012, she works at the Telecommunication Systems Department of RUDN University, she is an Associate Professor at the Department of Applied Probability and Informatics of RUDN University. Her current research interests lie in the area of performance analysis of radio resource management techniques in LTE networks. Her e-mail address is: markova_ev@pfur.ru .

DMITRY POLUEKTOV is a master student at the Department of Applied Informatics and Probability Theory of Peoples' Friendship University of Russia (RUDN University). He received a bachelor's degree from the RUDN University in 2016. Currently, his current research focuses on performance analysis of 4G/5G networks and LSA framework. His e-mail address is: poluektov_dmitri@mail.ru.

IRINA GUDKOVA received her M.Sc. degree in applied mathematics and Cand.Sc. degree in applied mathematics and computer sciences from the Peoples' Friendship University of Russia (RUDN University) in 2009 and 2011, respectively. She is currently an Associate Professor with the Applied Probability and Informatics Department, RUDN University. She has co-authored multiple research works. Her current research interests include mathematical modelling and performance analysis of 4G/5G networks, smart cities, spectrum sharing, multicast services, radio access, teletraffic theory, and queuing theory. Her e-mail address is: gudkova_ia@pfur.ru.

PAVEL MASEK received his BS and MS degrees from the Department of Telecommunication, Brno University of Technology, Czech Republic, in 2011 and 2014, respectively. He is currently pursuing his PhD degree in teleinformatics at the same university. He has publications on a variety of networking-related topics in internationally recognized venues including those published in the IEEE Communications Magazine, as well as several technology products. His primary research interest lies in the area of wireless networks - M2M/H2H communication, cellular networks, heterogeneous networking, and data offloading techniques. His e-mail address is: masekpavel@feec.vutbr.cz

JIRI HOSEK is an Associate Professor and Head of WISLAB research group (http://wislab.cz) at Department of Telecommunications, Brno University of Technology, Czech Republic. Jiri deals mostly with industry-oriented R&D projects in the area of future mobile networks, Internet of Things and home automation services. Jiri (co-) authored more than 70 research works on networking technologies, wireless communications, quality of service, quality of experience and IoT applications including those published in the IEEE Communications Magazine. Jiri is an experienced speaker regularly participating and actively presenting his research work on premier international conferences and workshops. His e-mail address is: hosek@feec.vutbr.cz

ON AN EXACT SOLUTION OF THE RATE MATRIX OF QUASI-BIRTH-DEATH PROCESS WITH SMALL NUMBER OF PHASES

Rama Murthy Garimella
International Institute of Information Technology,
Hyderabad, India,
Email: rammurthy@iiit.ac.in

Rumyantsev Alexander
Institute of Applied Mathematical Research
of the Karelian Research Centre RAS;
Petrozavodsk State University,
Petrozavodsk, Russia
Email: ar0@krc.karelia.ru

KEYWORDS

QBD Process, Matrix-Analytic Method, Exact Solution, Internet-of-Things

ABSTRACT

A new method of obtaining exact solution for the rate matrix R in the Matrix-Analytic method in case of the phase state of dimension two is proposed. The method is based on symbolic solution of the determinental polynomial equation, and obtaining a linear matrix equation for the unknown rate matrix R by Cayley–Hamilton theorem. The method is applied to analyze the Energy-Performance tradeoff of an Internet-of-Things device. A new randomized regime switching scheme is proposed, which, as it is shown by means of numerical experiment, provides significant decrease of energy consumption of the system under study.

INTRODUCTION

Univariate polynomial equations naturally arise in many branches of mathematics. An exact symbolic solution of such an equation (in terms of arithmetic operations and radicals) is known to exist for polynomials of order less or equal to four. Moreover, the nonexistence of such a solution for a polynomial equation of order greater or equal to five was established in Abel–Ruffini theorem. The solvability concept of an arbitrary monic polynomial was provided by Galois group theory.

A generalization of polynomial equations in which the coefficients and the argument are matrices (matrix polynomial equations) has been studied in a number of works, and the detailed theory of uni-variate matrix polynomial equations was developed [15]. In the research area of Queueing Theory, the matrix quadratic equation

$$R^2A^{(2)} + RA^{(1)} + A^{(0)} = 0 \qquad (1)$$

was first used to find a solution of a QBD process (by means of Complex Analysis-based method) in late 60's. Wallace [34] and Evans [10] showed that in the case of QBD process, matrix geometric solution exists for the equilibrium distribution where the rate matrix R is a minimal nonnegative solution of matrix quadratic equation (1). M. Neuts generalized the result to arbitrary $G/M/1$-type Markov processes and showed [24] that R is the minimal non-negative solution of matrix power series equation

$$\sum_{i=0}^{\infty} R^i A^{(i)} = 0. \qquad (2)$$

Thus, the analysis is essentially reduced to obtaining the matrix R. In general, a few iterative procedures, claimed to be numerically stable, are used [24, 18, 6] (see also the comparison of iterative procedures [17]). Alternatively, the spectral decomposition-based methods were suggested [8, 21] (which required eigenanalysis of a matrix polynomial), some of them utilizing special structure of the model to dramatically decrease the computational complexity [9].

As opposed to probabilistic-based iterative procedures, the Jordan canonical form representation of rate matrix R has been suggested as a closed form/analytic solution of (2). H.R. Gail et al. [13] suggested Spectral Analysis method based on Jordan canonical form to analyze G/M/1- and M/G/1-type Markov chains. G. Rama Murthy [26] successfully proposed method for computing the Jordan form representation of R (i.e computing eigenvalues and generalized eigenvectors of R). He thus showed the relationship between classical Complex Analysis method and the iterative procedures.

However, an explicit formula for R in terms of the given matrices $A^{(0)}, A^{(1)}, A^{(2)}$ for the QBD process is, to the best of our knowledge, not available in general. The explicit expression for R was obtained for a number of special cases, in particular, when either $A^{(0)}$, or $A^{(2)}$ is a rank-one matrix [18, 14]. A natural question that remained was whether there are other cases where R can be expressed explicitly. In this paper we answer such a question when $A^{(0)}, A^{(1)}, A^{(2)}$ and, consequently, R are 2x2 matrices.

We apply the proposed algorithm of the exact solution of (1) to suggest a new simple approach to improve energy efficiency of battery-powered Internet-of-Things (IoT) devices. IoT is an intensively studied field, where energy efficiency is one of the dominant areas. The realization of many IoT applications relies on the wireless network of small battery-powered devices (such as wireless sensors, transmitter nodes, actuators etc.), with a typical battery lifetime varying from days to several years.

Proceedings 31st European Conference on Modelling and
Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi,
Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

Given that, a typical IoT device is expected to work in energy saving mode, though providing a required quality of service (QoS). Recent research covers various aspects of energy-efficient IoT network design, such as efficient network architecture [31, 32], implementation of efficient routing/clustering algorithms [7, 5, 29, 33] (see also a recent survey [1]). Our application is focused on peer-to-peer type wireless networks, which consist of basic low-price low-powered devices (nodes) which use no (or only basic) centralized management and possess strict energy consumption restrictions. Under this assumptions, implementation of the aforementioned sophisticated routing schemes and network architecture seems irrelevant. Instead, we propose a simple randomized regime switching scheme, which, once implemented at each node, provides significant decrease of energy consumption of the system under study.

We also note, that the proposed randomized management approach may be applied to systems, where cost effectiveness and service elasticity is important, such as high-performance and cloud-based computing systems [22], as well as teletraffic systems (e.g. on-demand content servers), where the operational cost (e.g. energy cost, or cloud service cost), as well as the system speed, is to be adopted to the working conditions. The approach is suitable for heavy load conditions, since the centralized management (which could become a bottleneck once implemented) is unnecessary.

This research paper is organized as follows. First, we present an algebraic approach on obtaining the matrix R. Next, we apply this approach to solve the optimization problem related to Energy-Performance tradeoff in the field of Internet-of-Things. We illustrate the approach with simulation results.

EXACT SOLUTION FOR RATE MATRIX OF A QBD PROCESS

The QBD process is a continuous time Markov process $\{(X(t), Y(t)), t \geqslant 0\}$ with countable state space $E := \{(0, j), j = 1, \dots, m_0, (i, j), i \geqslant 1, j = 1, \dots, m\}$, where the *phase* variable $Y(t)$ may take one of m (or m_0 for boundary states) values and *level* variable $X(t)$ is increased/decreased by at most one at each transition. The state space E can be partitioned into *levels* with level $n \geqslant 1$ being the subset $\{(n, j), j = 1, \dots, m\} \subset E$. The infinitesimal generator matrix of a QBD process has the following block-tridiagonal representation [18]

$$Q = \begin{pmatrix} A^{0,0} & A^{0,1} & 0 & 0 & \dots \\ A^{1,0} & A^{1,1} & A^{(0)} & 0 & \dots \\ 0 & A^{(2)} & A^{(1)} & A^{(0)} & \dots \\ 0 & 0 & A^{(2)} & A^{(1)} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (3)$$

where $A^{(i)}, i = 0, 1, 2$ are square matrices of order m, $A^{0,0}$ is a square matrix of order m_0 and $A^{1,0}, A^{0,1}$ are possibly rectangular matrices. Note that the matrix $A := A^{(0)} + A^{(1)} + A^{(2)}$ necessarily satisfies the balance condition

$$A\mathbf{1} = 0, \quad (4)$$

where $\mathbf{1}$ is the vector of ones of corresponding dimension. Recall also that $A^{(0)} \geqslant 0, A^{(2)} \geqslant 0$ componentwise, and $A^{(1)}_{i,j} \geqslant 0$ for $i \neq j$, whereas $A^{(1)}_{i,i} \leqslant 0$. Note that it readily follows, that for any $i = 1, \dots, m$,

$$|A^{(1)}_{i,i}| = \sum_{j=1}^{m} \left[A^{(0)}_{i,j} + A^{(2)}_{i,j} + (1 - \delta_{i,j}) A^{(1)}_{i,j} \right], \quad (5)$$

where $\delta_{i,j}$ is the Kronecker delta function.

To establish the stability conditions, it is necessary to find a solution of the following system

$$\begin{cases} \alpha A & = 0 \\ \alpha \mathbf{1} & = 1. \end{cases} \quad (6)$$

where the stochastic vector α may be interpreted as the distribution of the phase $Y(t)$ at high levels $X(t)$. Given that, the stability follows from the Neuts ergodicity condition, see [16]: the QBD process $\{X(t), Y(t)\}, t \geqslant 0$ is positive recurrent iff

$$\rho := \alpha A^{(0)} \mathbf{1} / \alpha A^{(2)} \mathbf{1} < 1. \quad (7)$$

Provided (7) holds, there exists a vector of limiting probabilities $\pi = (\pi_0, \pi_1, \dots)$, which consists of the vectors π_k of equilibrium probabilities of level $k \geqslant 0$ (i.e. $\pi_k = (\pi_{k,1}, \dots, \pi_{k,m}), k \geqslant 1$), and is the unique solution of the following system

$$\begin{cases} \pi Q & = 0 \\ \pi \mathbf{1} & = 1. \end{cases} \quad (8)$$

It is shown in [24], that the vector π can be level-wise computed by the celebrated matrix-geometric solution

$$\pi_k = \pi_{k-1} R, \quad k \geqslant 1, \quad (9)$$

where the square matrix R of order m is the minimal nonnegative solution of the system (1), which exists under the stability assumptions. The initial vectors π_0, π_1 are obtained by the following linear system of equations

$$(\pi_0, \pi_1) \begin{pmatrix} A^{0,0} & A^{0,1} \\ A^{1,0} & A^{1,1} + RA^{(2)} \end{pmatrix} = 0, \quad (10)$$

$$\pi_0 \mathbf{1} + \pi_1 (I - R)^{-1} \mathbf{1} = 1. \quad (11)$$

Thus, obtaining the matrix R is crucial for performance analysis of the stable system.

Intuitively, the probability $\pi_k \mathbf{1}$ of the system persistence at a particular level $k \geqslant 0$ is related to the speed of the geometrical decrease in (9). Indeed, the Spectral radius $\eta = sp(R)$ (the greatest eigenvalue of R) is the so-called *caudal characteristic* of the QBD. It is true, that since the system is stable, then η is real, positive, and $\eta < 1$ (which is related to the Perron–Frobenius theorem, see [18]). Thus, the probabilistic behavior of the QBD is defined by the values ρ, and η, for details see [18].

Note that in general the equation (1) has to be solved numerically. However, under some restrictions on the matrices $A^{(0)}$ and $A^{(2)}$, there is a possibility to obtain the explicit solution. The following theorems provide an algebraic formula for the matrix R of a discrete-time Markov chain (an extension to a continuous-time process is straightforward).

Theorem 1 *[18] Assume that $A^{(2)} = cr$, where c is a column vector, and r is a row vector s.t. $r\mathbf{1} = 1$. Then $G = \mathbf{1}r$.*

Note that the matrix G is a minimal nonnegative solution of a matrix equation $G = A_2 + A_1 G + A_0 G^2$, and the matrix R can be easily obtained by the well-known duality $R = A_0(I - A_1 - A_0 G)^{-1}$ [19].

Theorem 2 *[18] Assume that $A^{(0)} = cr$, where c is a column vector, and r is a row vector s.t. $r\mathbf{1} = 1$. Then $R = c\xi$, where $\xi = r(I - A^{(1)} - \eta A^{(2)})^{-1}$ and $\eta = \xi c$, where $\eta = \mathrm{sp}(R)$.*

Note that in the latter theorem the value η has to be computed in advance, see [18].

However, the restriction of the Theorems 1, 2 is the requirement of matrices $A^{(0)}$ or $A^{(2)}$ to be of rank one. A general exact solution of the matrix quadratic equation (1) can be obtained by obtaining the Jordano canonical form. This approach is based on the following lemma.

Lemma 1 *The polynomial matrix $A(\xi) := A^{(0)} + \xi A^{(1)} + \xi^2 A^{(2)}$ allows the following factorization:*

$$A(\xi) \equiv (\xi I - R)(\xi A^{(2)} + RA^{(2)} + A^{(1)}). \qquad (12)$$

The proof follows by direct expansion of the r.h.s. of (12) and substitution of (1). This result reported in [11] was first applied to QBD processes in [26].

Remark 1 *It follows from Lemma 1, that $\det A(\xi) = \det(\xi I - R)\det(\xi A^{(2)} + RA^{(2)} + A^{(1)})$. Thus, the eigenvalues of R are the zeroes of the determinantal polynomial $\det A(\xi)$. Note that there are exactly m zeroes of the determinantal polynomial $\det A(\xi)$ which are strictly inside unit circle and these are the eigenvalues of R (provided $\eta < 1$), see e.g. [26]. Hence, the remaining $r \leqslant m$ zeroes are the zeroes of $\det(\xi A^{(2)} + RA^{(2)} + A^{(1)})$, and are outside the unit circle (a more detailed discussion of the number and location of eigenvalues of R for the case $\eta \leqslant 1$ can be found in [23]).*

The following lemma proved in [26] enables determination of left eigenvectors of rate matrix, R.

Lemma 2 *u is a left eigenvector of rate matrix R corresponding to eigenvalue ξ if and only if*

$$uA(\xi) = \mathbf{0}. \qquad (13)$$

Thus, using Lemma 1 and Lemma 2, if rate matrix R is diagonalizable, it can be obtained by the spectral representation

$$R = TDT^{-1}, \qquad (14)$$

where D is the diagonal matrix of m eigenvalues inside the unit disk, and columns of T consist of right eigenvectors of rate matrix. A more general discussion of Jordan canonical form method of obtaining the rate matrix R is provided in [27]. It should be noted that the suggested approach is a closed form method, in contrast to widely used iterative numerical procedures.

EXACT ALGEBRAIC FORMULA FOR R OF THE QBD PROCESS WITH TWO PHASES

Let $m = 2$, that is, the phase state of the QBD process has exactly two states at each level. Observe that the determinantal polynomial $\det A(\xi)$ is of degree four. Note also, that $\xi = 1$ is always the root of $\det A(\xi)$ by the balance condition (4). Rewrite

$$\det A(\xi)/(\xi - 1) = a_3\xi^3 + a_2\xi^2 + a_1\xi + a_0. \qquad (15)$$

Denote by $\xi_i, i = 1, 2, 3$ the zeroes of the polynomial (15). By the Remark 1, two of these zeroes are inside the unit disk, denote them ξ_1, ξ_2. Consider the characteristic polynomial of the matrix R in a monic form

$$\det(\xi I - R) = (\xi - \xi_1)(\xi - \xi_2) = \xi^2 + b_1\xi + b_0, \qquad (16)$$

where $b_0 = \xi_1\xi_2 = \det R$, $b_1 = -(\xi_1 + \xi_2) = -\mathrm{Trace}(R)$ are (real) scalars, since R is a nonnegative matrix. By Perron–Frobenius theorem, $\eta = \mathrm{sp}(R)$ is simple, real eigenvalue, and $\eta \in (0, 1)$, provided (7) holds. W.o.l.o.g. let $\xi_1 = \eta$. Then ξ_2 is also real, which makes ξ_3 also real. Given that, it is easy to obtain

$$b_1 = \frac{a_2}{a_3} + \xi_3, \qquad (17)$$

$$b_0 = -\frac{a_0}{a_3\xi_3}. \qquad (18)$$

Hence, it is necessary to obtain ξ_3. Note that it can be done by the celebrated trigonometric solution of the cubic equation (15) by the following substitutions:

$$p = \frac{3a_3a_1 - a_2^2}{9a_3^2}, \; q = \frac{2a_2^3 - 9a_3a_2a_1 + 27a_3^2a_4}{27a_3^3}. \qquad (19)$$

Then [35]

$$\xi_3 = -\frac{a_2}{3a_3} + 2\sqrt{-p}\cos\left(\frac{1}{3}\cos^{-1}\left(\frac{q}{2p\sqrt{-p}}\right)\right). \qquad (20)$$

By Cayley-Hamilton theorem, we obtain

$$R^2 = -b_1R - b_0I, \qquad (21)$$

which, by substitution into (1), leads to the following system of linear equations:

$$R\left[A^{(1)} - b_1A^{(2)}\right] - b_0A^{(2)} + A^{(0)} = \mathbf{0}. \qquad (22)$$

Thus, if $A^{(1)} - b_1 A^{(2)}$ is invertible, then

$$R = \left[b_0 A^{(2)} - A^{(0)}\right] \left[A^{(1)} - b_1 A^{(2)}\right]^{-1}. \qquad (23)$$

Lemma 3 *The matrix $A^{(1)} - b_1 A^{(2)}$ is invertible.*

Proof: Since the eigenvalues are distinct, then R is diagonalizable, hence, R has a spectral representation (14), which is equivalent to

$$R = \eta E_1 + \xi_2 E_2, \qquad (24)$$

where E_1, E_2 are the so-called residue matrices, with $E_1 + E_2 = I$. Note, that by Perron–Frobenius theorem, $E_1 \geqslant \mathbf{0}$ componentwise [18], which provides a componentwise inequality

$$R \geqslant \xi_2 E_1 + \xi_2 E_2 = \xi_2 I. \qquad (25)$$

Recall, that from (4), the equality $A(1)\mathbf{1} = \mathbf{0}$ holds. Considering the expansion (12) at $\xi = 1$ provides $(I - R)(A^{(2)} + RA^{(2)} + A^{(1)})\mathbf{1} = \mathbf{0}$. However, since by (16) $\det(I - R) = b_0 \neq 0$, then the matrix $(I - R)$ is nonsingular, hence, $(A^{(2)} + RA^{(2)} + A^{(1)})\mathbf{1} = \mathbf{0}$, which provides (considering the non-trivial case $A^{(2)} \neq \mathbf{0}$)

$$|A^{(1)}_{i,i}| > \sum_{j=1}^{m} \left[\xi A^{(2)}_{i,j} + (RA^{(2)})_{i,j} + (1-\delta_{i,j}) A^{(1)}_{i,j}\right], \quad (26)$$

for any real $\xi \in (-1, 1)$. Now taking $\xi = \eta$ in (26) and noting, that (25) provides $RA_2 \geqslant \xi_2 A_2$ componentwise, yields that $[A^{(1)} - b_1 A^{(2)}]$ is a strictly diagonally dominant matrix and hence is nonsingular. ■

Thus, the suggested procedure for exact computation of the rate matrix R is as follows.

1. Obtain the maximal eigenvalue ξ_3 by (20).
2. Obtain b_1 by (17) and b_0 by (18).
3. Obtain R by (23).

Alternatively, ξ_1, ξ_2 can be computed using Lemma 1. Further, left eigenvectors can be computed by Lemma 2. Finally, R is obtained by (14). Note that this approach can be generalized to an arbitrary $G/M/1$-type Markov process.

Now we consider the case $\xi_2 = 0$, i.e. the rate matrix R is singular. Thus, following the suggested procedure, we obtain $b_0 = 0$ and $b_1 = -\eta$. Hence $R = [-A^{(0)}][A^{(1)} + \eta A^{(2)}]^{-1}$, which corresponds to the known result presented by Theorem 2.

Remark 2 *Note that algebraically the matrices R and $A^{(0)}$ have the same rank. Since $\det A(\xi)$ is the polynomial of power $2m$, and eigenvalues of R are zeroes of the aforementioned polynomial, in the following cases (provided $\eta < 1$) the zeroes can be explicitly computed, and thus the rate matrix R can be theoretically computed in closed form:*
i) $m \leqslant 4$, if $A^{(0)}$ is rank-one matrix,
ii) $m \leqslant 3$ if $A^{(0)}$ is of rank two.
We continue with $m = 2$ and a full rank $A^{(0)}$ below.

RANDOM SWITCHING FOR POWER SAVING

Now we turn to a practical application of the result. Heavy restrictions on energy consumption of a battery-powered IoT device result in typical application of asynchronous information transmission/receive modes with various data rates, and equipment of the device with lightweight, low consuming logic [2]. These aspects lead to the following queueing model of a single IoT service device. Consider a queueing system with a renewal input flow of customers arriving into an (unbounded) First-Come-First-Served queue. The i.i.d. interarrival times are exponentially distributed with rate $\lambda > 0$. Each customer requires an exponentially distributed (with unit rate) amount of work to be done (say, information to be transmitted). The single server operates two speed modes (call them high and low), with rates $\mu_2 > \mu_1 > 0$. The server may switch the speed only at the arrival/departure epochs (asynchronously). Denote $c_0 < c_1 < c_2$ the energy consumption per unit time in idle (no customers in the system)/low/high modes. In order to preserve energy, the server implements the following random switching policy:

- at the task arrival epoch, given the current mode is low, switch to high mode with probability (w.p.) p_1, or remain low w.p. $1 - p_1$;
- at the task departure epoch, given the current mode is high, switch to low mode w.p. p_2, or remain high w.p. $1 - p_2$.

Let $\nu(t) \in \{0, 1 \dots\}$ be the number of customers, and $m(t) \in \{1, 2\}$ be the mode of the system at time $t \geqslant 0$. Then the following Markov process

$$\{(\nu(t), m(t)) \in \{0, 1, \dots\} \times \{1, 2\},\ t \geqslant 0\} \qquad (27)$$

is a continuous-time QBD process, with $\nu(t)$ being the *level*, and $m(t)$ being the *phase* at time t.

The infinitesimal generator matrix of the process (27) has the form (3), where we define the matrices explicitly:

$$A^{(0)} = \begin{pmatrix} (1-p_1)\lambda & p_1\lambda \\ 0 & \lambda \end{pmatrix}, \qquad (28)$$

$$A^{(1)} = \begin{pmatrix} -\lambda - \mu_1 & 0 \\ 0 & -\lambda - \mu_2 \end{pmatrix}, \qquad (29)$$

$$A^{(2)} = \begin{pmatrix} \mu_1 & 0 \\ p_2\mu_2 & (1-p_2)\mu_2 \end{pmatrix}, \qquad (30)$$

$$A^{0,0} = -\lambda I, \quad A^{0,1} = A^{(0)}, \qquad (31)$$

$$A^{1,1} = A^{(1)}, \quad A^{1,0} = A^{(2)}. \qquad (32)$$

To establish the stability criterion of the process (27), we construct the matrix

$$A := A^{(0)} + A^{(1)} + A^{(2)} = \begin{pmatrix} -p_1\lambda & p_1\lambda \\ p_2\mu_2 & -p_2\mu_2 \end{pmatrix}. \quad (33)$$

Solving the system (6) and using (28)-(30), the condition (7) provides

$$\alpha_1\mu_1 + \alpha_2\mu_2 > \lambda,$$

where

$$\alpha_1 = \frac{p_1\lambda}{p_1\lambda + p_2\mu_2},\ \alpha_2 = \frac{p_2\mu_2}{p_1\lambda + p_2\mu_2},$$

which is equivalent to

$$\lambda p_1(\lambda - \mu_2) + \mu_2 p_2(\lambda - \mu_1) < 0. \tag{34}$$

Intuitively, the condition (34) indicates a negative drift of the service process of the system under heavy load, with respect to the mode switching intensity.

Note that $p_1 = 1$ and $p_2 = 0$ corresponds to a classical $M/M/1$ service system working at the speed μ_2 (referred below as classical system), where the stability condition (34) reduces to the celebrated $\rho := \lambda/\mu_2 < 1$. Now let $\mathrm{E}\nu_0$ be the average number of customers, and $\mathrm{E}\mathcal{E}_0$ be the average energy consumption per unit time in the classical system in stationary regime. It can be readily seen, that

$$\mathrm{E}\nu_0 = \frac{\rho}{1-\rho}, \tag{35}$$

$$\mathrm{E}\mathcal{E}_0 = c_0(1-\rho) + c_2\rho, \tag{36}$$

where, recall, $1-\rho$ is the stationary idle probability of the classical system (for details on the aforementioned classical results see [3]). We may consider $\mathrm{E}\nu_0$ as the QoS parameter of the classical system.

Now we turn to the original two-mode system, i.e. we consider the non-trivial case $p_1, p_2 > 0$. Provided (34) holds, we define the matrix R following the steps of the suggested procedure of exact computation, where the coefficients in the polynomial (15) are obtained as follows:

$$\begin{aligned} a_3 &= \mu_1\mu_2(1-p_2),\\ a_2 &= -\mu_1(\lambda+\mu_2) - \lambda\mu_2(1-p_2),\\ a_1 &= \lambda(\lambda+\mu_1) + \lambda\mu_2(1-p_1),\\ a_0 &= -\lambda^2(1-p_1). \end{aligned}$$

After obtaining R from the equation (23), the equations (9)–(11) provide the stationary system state probabilities. Straightforward manipulation leads to the following system for π_1:

$$\begin{cases} \pi_1\left(\frac{1}{\lambda}A^{(2)} - R^{-1}\right)A^{(0)}\mathbf{1} &= 0,\\ \pi_1\left(\frac{1}{\lambda}A^{(2)} + (I-R)^{-1}\right)\mathbf{1} &= 1. \end{cases} \tag{37}$$

Then the value π_0 is obtained as follows:

$$\pi_0 = \frac{1}{\lambda}\pi_1 A^{(2)}. \tag{38}$$

The obtained solution allows to evaluate the average number of customers in the system as the QoS measure:

$$\mathrm{E}\nu_{p_1,p_2} = \pi_1(I-R)^{-2}\mathbf{1}, \tag{39}$$

where by notation ν_{p_1,p_2} we stress the dependence on the mode switching probabilities. The average energy consumption may be obtained as follows:

$$\mathrm{E}\mathcal{E}_{p_1,p_2} = \pi_0 c_0\mathbf{1} + \pi_1(I-R)^{-1}(c_1,c_2)^T, \tag{40}$$

where the transposed vector (c_1, c_2) is the column-vector of energy consumption in each mode. Then the following optimization problem can be solved: minimization of the average energy consumption in stationary regime, provided the controlled QoS decrease:

$$\begin{cases} \mathrm{E}\mathcal{E}_{p_1,p_2} & \to_{p_1,p_2>0} \min,\\ \mathrm{E}\nu_{p_1,p_2} & \leqslant (1+\varepsilon)\mathrm{E}\nu_0, \end{cases} \tag{41}$$

for some $\varepsilon > 0$.

Now we present the numerical investigation of the optimization problem (41), using the energy and speed values from a real-world example. We describe the ATmega328/P controller [4], widely used in applications. This low power consuming micro-controller is capable of dynamical voltage/frequency setup, which allows to significally reduce the power consumption. We use the values of power consumption for given frequency/voltage configuration extracted from [4] as follows: 0.6 mW for 1 MHz/2 V active regime, 26 mW for 8 MHz/5 V active regime, and 0.08 mW for 2 V idle regime. We assume, that $\rho < 1$, that is, the system working at 8 MHz is stable.

Note, however, that the high frequency regime is relatively expensive. Now we present two scenarios: (i) the system is stable at the low regime (with input rate λ_1); (ii): the system is unstable at the low regime (with input rate λ_2). Thus, we have the following settings:

$$\begin{aligned} &\mu_1 = 1,\ \mu_2 = 8,\\ &c_0 = 0.08,\ c_1 = 0.6,\ c_2 = 26,\\ &\lambda_1 = 0.5,\ \lambda_2 = 2.5. \end{aligned}$$

Following the procedure of obtaining the exact solution, we numerically solve the optimization problem (41) and obtain approximate optimal values p_1, p_2, as well as approximate optimal energy consumption. We vary the QoS degradation $1+\varepsilon \in (1.1, 30)$ and plot the obtained values. We perform simulation with R package [25].

It is easy to see from Fig. 1, that, as expected, if the low regime is stable, the system tends to an M/M/1 system working at low speed, with the appropriate growth of performance degradation.

On the other hand, if the system at low frequency is unstable (see Fig. 2), the system oscillates around the stability border, and the consumption tends to some average value w.r.t. the switching probabilities.

CONCLUSIONS AND FUTURE WORK

We have presented the algebraic approach for obtaining the exact value of rate matrix R in case the phase state of the QBD process has only two phases. Note, that, albeit being simplistic, such a QBD process may be used to model quite a number of recent applications, such as the IoT devices, high-performance and cloud-based servers (where processors use the Dynamic Voltage and Frequency Scaling technology), telecommunications (with two types of service, such as the Invite and non-Invite messages in SIP protocol [12]). For such types of applications, it is crucial to obtain the cost effectiveness

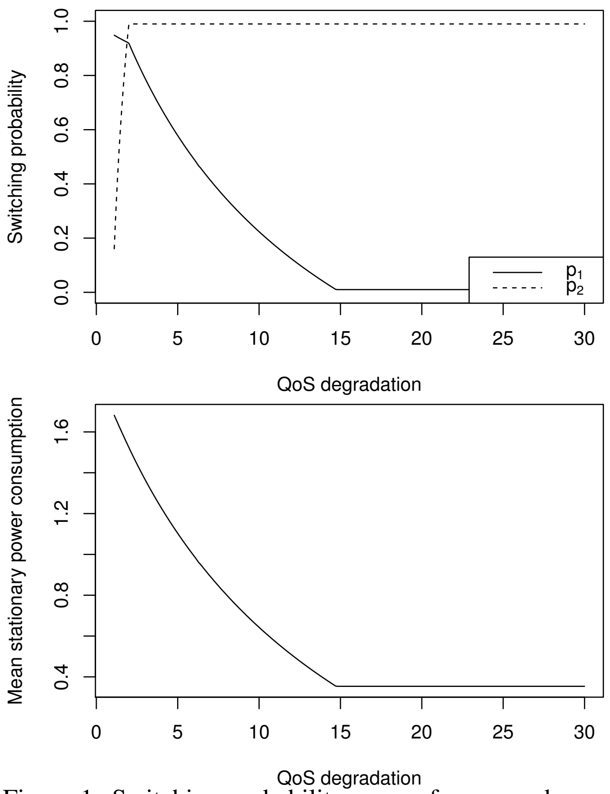


Figure 1: Switching probability vs. performance degradation (upper); average energy consumption per unit time vs. performance degradation in a system being stable at low level (input rate $\lambda_1 = 0.5, \mu_1 = 1$).

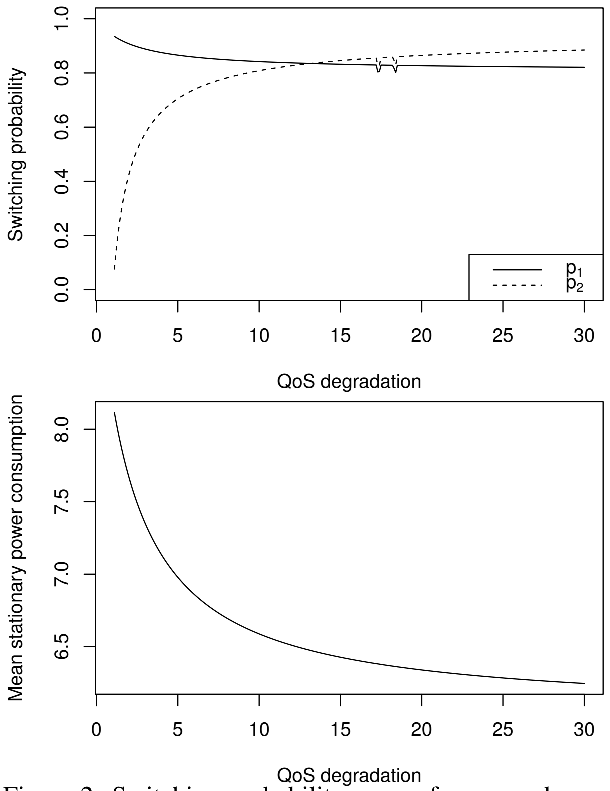


Figure 2: Switching probability vs. performance degradation (upper); average energy consumption per unit time vs. performance degradation in a system being unstable at low level (input rate $\lambda_2 = 2.5, \mu_1 = 1$).

(such as energy efficiency in the IoT) without implementation of a sophisticated control due to a significant overhead induced by such a control. We proposed a randomized scheme which may be implemented in such systems to obtain the desired cost effectiveness under controlled QoS degradation, when possible. We leave the field test of the proposed scheme for future research.

We note, that most of the results of this research paper can be generalized to arbitrary $G/M/1$-type Markov process (with two states at each level). In this case, the rate matrix R is a solution of the matrix power series equation (2). However, a detailed study is a topic of a separate discussion [28].

ACKNOWLEDGEMENTS

The authors thank Prof. Evsey Morozov for his helpful comments, that significantly improved the paper. The research of AR is partially supported by RF President's Grant No. MK-1641.2017.1, and by RFBR, grants No.15-07-02341, 15-07-02354, 16-07-00622, 15-29-07974.

REFERENCES

[1] Abbas, Z., Yoon, W.: A Survey on Energy Conserving Mechanisms for the Internet of Things: Wireless Networking Aspects. Sensors. 15, 24818–24847 (2015)

[2] Augustin, A., Yi, J., Clausen, T., Townsley, W.M.: A study of LoRa: Long range & low power networks for the internet of things. Sensors. 16, 1466 (2016)

[3] Asmussen, S. Applied Probability and Queues. 2003.

[4] ATmega328/P Datasheet Complete Rev.: Atmel-42735B-ATmega328/P_Datasheet_Complete-11/2016

[5] Bagula, A., Abidoye, A.P., Zodi, G.-A.L.: Service-Aware Clustering: An Energy-Efficient Model for the Internet-of-Things. Sensors. 16, 9–0 (2016)

[6] Bini, D.A., Latouche, G., Meini, B.: Solving matrix polynomial equations arising in queueing problems. Linear Algebra and its Applications. 340, 225–244 (2002)

[7] Chelloug, S.A.: Energy-Efficient Content-Based Routing in Internet of Things. Journal of Computer and Communications. 03, 9–20 (2015)

[8] Daigle, J. N., Lucantoni, D. M.: Queueing systems having phase-dependent arrival and service rates. Numerical Solution of Markov Chains, 161–202. Marcel Dekker Inc., New York (1991)

[9] Do, T.V., Chakka, R.: An efficient method to compute the rate matrix for retrial queues with large number of servers. Applied Mathematics Letters. 23, 638—643 (2010)

[10] Evans, R. V.: Geometric Distribution in some Two-Dimensional Queueing Systems, Operations Research. 15, 830–846 (1967)

[11] Gantmacher, Felix (1959), Theory of matrices, AMS Chelsea publishing

[12] Gaidamaka, Y.V.: Model with threshold control for analyzing a server with an SIP protocol in the overload mode. Automatic Control and Computer Sciences. 47, 211–218 (2013)

[13] Gail, H.R., Hantler, S.L., Taylor, B.A.: Spectral Analysis of M/G/1 and G/M/1 Type Markov Chains. Advances in Applied Probability. 28, 114 (1996)

[14] Gillent, F. and Latouche, G.: Semi-explicit solutions for M/PH/1-like queuing systems. European Journal of Operational Research. 13(2), 151–160, (1983)

[15] Gohberg, I., Lancaster, P., Rodman, L.: Matrix polynomials. Society for Industrial and Applied Mathematics, Philadelphia (2009).

[16] He, Q.-M.: Fundamentals of Matrix-Analytic Methods. Springer, New York (2014)

[17] Hung, T.T., Do, T.V.: Computational aspects for steady state analysis of QBD processes. Periodica Polytech. Ser. Electr. Eng. 44, 179–200 (2001)

[18] Latouche, G. and Ramaswami, V.: Introduction to Matrix Analytic Methods in Stochastic Modeling. ASA–SIAM, Philadelphia (1999)

[19] Latouche, G.: A note on two matrices occurring in the solution of quasi-birth-and-death processes. Stochastic Models. 3, 251–257 (1987)

[20] Low-Power Long Range LoRa Technology Transceiver Module. Datasheet No. DS50002346B, 2015.

[21] Mitrani, I., Chakka, R.: Spectral expansion solution for a class of Markov models: Application and comparison with the matrix-geometric method. Performance Evaluation. 23, 241–260 (1995)

[22] Mukherjee, D., Dhara, S., Borst, S., van Leeuwaarden, J.S.H.: Optimal Service Elasticity in Large-Scale Distributed Systems. ArXiv e-prints. 1703.08373, (2017)

[23] Naoumov, V., Samouylov, K. and Gaidamaka, Yu.: Multiplicative solutions of finite Markov chains. RUDN, Moscow (2015) (in Russian)

[24] Neuts, M.F. Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach. The Johns Hopkins University Press, Baltimore (1981).

[25] R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0. http://www.r-project.org/

[26] Rama Murthy, G.: Transient and equilibrium analysis of computer networks: Finite memory and matrix geometric recursions, PhD. Thesis, Purdue University, West Lafayette (1989)

[27] Rama Murthy, G., Kim, M., Coyle, E. J.: Equilibrium analysis of skip-free Markov chains: Non-linear Matrix Equations. Communications in Statistics – Stochastic Models. 4, 547–571 (1991)

[28] Rama Murthy, G. and Rumyantsev, A.: On an exact solution of the rate matrix of G/M/1-type Markov process with small number of phases. Manuscript in preparation

[29] Rani, S., Talwar, R., Malhotra, J., Ahmed, S., Sarkar, M., Song, H.: A Novel Scheme for an Energy Efficient Internet of Things Based on Wireless Sensor Networks. Sensors. 15, 28603–28626 (2015)

[30] Sanchez-Iborra, R., Cano, M.-D.: State of the Art in LP-WAN Solutions for Industrial IoT Services. Sensors. 16, 708 (2016)

[31] Sarwesh, P., Shet, N.S.V., Chandrasekaran, K.: Energy Efficient Network Design for IoT Healthcare Applications. In: Bhatt, C., Dey, N., and Ashour, A.S. (eds.) Internet of Things and Big Data Technologies for Next Generation Healthcare. pp. 35–61. Springer International Publishing, Cham (2017)

[32] Sarwesh P, N. Shekar V. Shet, Chandrasekaran K: Energy efficient network architecture for IoT applications. In: 2015 International Conference on Green Computing and Internet of Things (ICGCIoT). pp. 784–789 (2015)

[33] Vellanki, M., Kandukuri, S.P.R., Razaque, A.: Node Level Energy Efficiency Protocol for Internet of Things. Journal of Theoretical and Computational Science. 3, (2016)

[34] Wallace, V. L.: The Solution of Quasi Birth and Death Processes Arising from Multiple Access Computer Systems, PhD Thesis, University of Michigan (1969)

[35] Zwillinger, D. CRC Standard Mathematical Tables and Formulae, 31st Edition. CRC, Boca Raton, 2003, 910 pages.

AUTHOR BIOGRAPHIES

RAMA MURTHY GARIMELLA is an Associate Professor at the International Institute of Information Technology, Hyderabad, India. He is a member of Eta Kappa Nu, Phi Kappa Phi (Honor Societies in USA), IEEE Computer Society, Computer Society of India, Senior Member of Association for Computing Machinery (ACM), and Fellow of IETE (Institution of Electronics and Telecommunication Engineers). His research interests include Wireless Sensor Networks, Adhoc Wireless Networks, Multi-Dimensional Neural Networks, Performance Evaluation.

RUMYANTSEV ALEXANDER received his PhD. from Petrozavodsk State University. He is now a researcher in the Institute of Applied Mathematical Research of the Karelian Research Centre of the Russian Academy of Sciences. His research interests include Stochastic Processes, Queueing Systems, High-Performance and Distributed Computing, Multi-Core and Many-Core Systems.

SIR DISTRIBUTION IN D2D ENVIRONMENT WITH NON-STATIONARY MOBILITY OF USERS

Sergey Fedorov*, Yurii Orlov*†, Andrey Samuylov†‡, Dmitri Moltchanov†‡,
Yuliya Gaidamaka†§, Konstantin Samouylov†§, Sergey Shorgin§

*Deparment of Kinetic Equations

Keldysh Institute of Applied Mathematics

Miusskaya Sq. 4, Moscow, 125047, Russian Federation
fyodor-on@mail.ru, ov3159f@yandex.ru

†Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
Miklukho-Maklaya St 6, Moscow, 117198, Russian Federation
ov3159f@yandex.ru, {samuylov_ak, molchanov_da, gaydamaka_yuv, samuylov_ke}@rudn.university

‡Department of Electronics and Communications Engineering
Tampere University of Technology
Korkeakoulunkatu 10, Tampere, 33720, Finland
{andrey.samuylov, dmitri.moltchanov}@tut.fi

§Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences
Vavilova st. 44-2, Moscow, 119333, Russian Federation
{gaydamaka_yuv, samuylov_ke}@rudn.university,
sshorgin@ipiran.ru

KEYWORDS
D2D communications, cellular networks, Fokker-Planck equation, SIR distribution, stochastic modeling

ABSTRACT

Fifth generation (5G) cellular systems are expected to rely on the set of advanced networking techniques to further enhance the spatial frequency reuse. Device-to-device (D2D) communications is one of them allowing users to establish opportunitic direct connections. The use of direct communications is primarily determined by the signal-to-interference ratio (SIR). However, depending on the users movement, the SIR of an ative connection is expected to drastically fluctuate. In this work we develop an analytical framework allowing to predict the channel quality between two moving entities in a filed of moving interfering stations. Assuming users movement driven by Fokker-Planck equation we obtain the empirical probability density function of SIR. The proposed methodology can be used to solve problems in the area of stochastic control of D2D communications in cellular networks.

INTRODUCTION

The need for higher capacity at the air interface in next generation mobile cellular systems calls for efficient ways of utilizing available frequency bands. Together with widening the bands itselves by moving upper in the frequency band to millimeter wave spectrum, one of the trends nowadays is developing of new techniques improving the use of available spectrum, e.g., increasing the base stations density via using the micro-, pico- and femto-cells, and enabling user devices to communicate directly with one another by establishing direct device-to-device (D2D) connections. The latter is the focus of this work.

The contemprorary cellular systems, such as GSM and LTE, are interference-limited in anture, implying that the link quality between two nodes highly depends on other nearby nodes using the same frequency, as their signal will literally interfere with the link of interest. As a result, the critical metric of interest for such systems characterizing the link quality is the signal-to-interference (SIR) ratio. A given level of SIR upper bounds the maximum capacity of a link. Furthermore, whenever SIR downcrosses some threshold it becomes impossible to maintain stable connection.

In this paper, we study the evolution of SIR for a pair of devices as a function of mobility, assuming that all the devices, including those of interest and interfering ones, are mobile. We formulate the framework that enables us to describe the analytical properties of SIR as a function of movement trajectories.

The proposed methodology not only forms the basis for analysis of statistical properties of SIR but potentially allows for stochastic control SIR functional. The future cellular system may contain advanced functionality of advising users to move in a certain derection to improve the channel quality. In this case, the framework offers the prossibility of functional modeling on the moving bodies trajectories. The case when the trajectories depend on the functional is of particular interest.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

GENERATING TRAJECTORIES OF NON-STATIONARY RANDOM WALKS

Consider a process with a non-stationary probability distribution function (pdf) and a given evolution equation. The method for generating a trajectory of the random process with thse properties has been proposed in (Bosov et al. 2014). In (Orlov 2014, Orlov and Fedorov 2016) it has been extended to generating an trajectories, whose pdf is evolving according to a given kinetic equation. The software tool for generating an ensemble of random non-stationary trajectories and performing their statistical analysis has been introduced in (Orlov and Fedorov 2016, Fedorov and Orlov, 2016). Below, we describe the key points of the developed methodology.

Let pdf $f(x,t)$ be defined by a kinetic equation. Consider the Fokker-Plank equation,

$$\frac{\partial f}{\partial t}+\frac{\partial}{\partial x}(u(x,t)f)-\frac{\lambda(t)}{2}\frac{\partial^2 f}{\partial x^2}=0\,. \qquad (1)$$

This equation can be solved numerically for a given initial condition,

$$f(x,t)\big|_{t=0}=\rho(x)\,.$$

The solution of (1) is based on the time horizon T. For the sake of simplicity, choose a unit time step. At each time instant $t=1,2,\ldots,T$ a random number is generated from the cumulative distribution function (CDF)

$$F(x,t)=\int_0^x f(y,t)dy\,. \qquad (2)$$

When the solution of (1) is represented as a histogram $f_j(t)$, where j is the number of class intervals, the continuous strictly monotonic CDF has the form

$$F(x,t)=(nx-j)\cdot f_{j+1}(t)+\sum_{k=1}^{j} f_k(t), \qquad (3)$$

for $x\in[(j-1)/n;\,j/n]$, $j=1,..,n$.

The algorithm for generating a sequence of random numbers, forming a feasible trajectory of the time series for a predetermined time interval is as follows. First, generate a series of random numbers, $\{y_k\}$, of length T having uniform distribution in $[0,1]$. The series with CDF $F(x,t)$ is obtained by combining the numerical inverse transform method with sliding window of length T, that is,

$$y_k=F_N(x_k,k). \qquad (4)$$

Generating a set of samples, we obtain the set of trajectories that can be considered as an ensemble of solutions of the kinetic equation.

STATISTICAL ANALYSIS OF FUNCTIONAL ON THE ENSEMBLE OF TRAJECTORIES

Consider SIR functional defined as a function of the distance between the moving points. The positions of these points form an ensemble of trajectories of a random process. Let x_i^α be a position of point i in a certain region of a three-dimensional space, where the superscript α labels (x,y,z) coordinates . The number of points in the system is $N+2$. Let also r_{ij} be the distance between the points, calculated in the Cartesian coordinate system as

$$r_{ij}^2=\sum_{\alpha=1}^{3}(x_i^\alpha-x_j^\alpha)^2\,. \qquad (5)$$

Let $\varphi_{ij}=\varphi(r_{ij})$ be an arbitrary transformation function of distance between two points *i* and *j*. Consider the SIR functional (Hesham ElSawy et al. 2014, Samuylov et al. 2015, Gaidamaka et al. 2016) between a pair of points, for example 1 and 2, where 1 stands for the receiver of interest, and 2 is the corresponding transmitter,

$$S(\mathbf{r}_1,\mathbf{r}_2)=\frac{\varphi_{12}}{\sum_{j=3}^{N+2}\varphi_{1j}}\,. \qquad (6)$$

Here $\mathbf{r}_i$ is the vector of *i* point coordinates.

The sum in the denominator of (6) can replaced by the product of *N* means of the transformation function,

$$U(\mathbf{r}_1)=\int_V \varphi(|\,\mathbf{r}_1-\mathbf{r}'\,|)\rho(\mathbf{r}')d\mathbf{r}'\,, \qquad (7)\text{rr}$$

where $\rho(\mathbf{r}')$ is the pdf of the spatial distribution of the remaining *N* points in area *V*. For convenience, denote $\mathbf{r}=\mathbf{r}_{12}$ and $r=|\,\mathbf{r}\,|$. Then SIR in (6) can be written as

$$S=S(r)=\frac{\varphi(r)}{NU(r)}\,. \qquad (8)$$

Consider only those $\rho(r)$, for which (8) is monotonous. In this case, the pdf of *S* is written as

$$g(S)=[S^{-1}(r)]'\rho(S^{-1}(r))\,. \qquad (9)$$

Assume that φ is monotonous and bounded, e.g.,

$$\varphi(r)=1/(r^2+a^2),\ a>0\,. \qquad (10)$$

Let *V* be a three-dimensional unit sphere, and points be uniformly distyributed in V . Then, (8) is monotonous. In this case, $U(r)$ in (8) is obtained in closed from,

$$\begin{aligned}&U(r)=\\&=\frac{3a}{r}\left(arctg\frac{r}{a}+\frac{1}{2}\left(arctg\frac{1-r}{a}-arctg\frac{1+r}{a}\right)\right)- \qquad (11)\\&-\frac{3}{4r}\ln\left(\frac{(1-r)^2+a^2}{(1+r)^2+a^2}\right).\end{aligned}$$

Non-monotonous link distances $\rho(\mathbf{r}')$ may lead to non-monotonous behaviour of (8). This happens when the movement of points in V is not stationary and, thus, does not result in stationary distribution of points in V . In this case, we should analyze the dynamic change of (8) when density $\rho(\mathbf{r}')$ varies in time. In practice, we assume that the distribution of points changes over time according to the kinetic equation (1) with respect to pdf $f(r,t)$. Specific samples of points forming trajectories of non-stationary random walk process are obtained using (4) based on the solution of kinetic equation on a given time horizon.

The average interference at r is determined by

$$U(r,t)=\int_V \varphi(|\mathbf{r}-\mathbf{r}'|)f(\mathbf{r}',t)d\mathbf{r}' \,. \qquad (12)$$

The communication between these transmitter and receiver of interest shall now be studied directly using (12). We emphasize that (1) is written with respect to the pdf of the increments of coordinates, so that the trajectory of the i-th point is the trajectory of the total random increment $x_i(t)$. This allows to easily specify an arbitrary random trajectory for the functional *S*.

The average value of the SIR is given by,

$$s(t)=\frac{1}{N}\int\frac{\varphi(r)}{U(r,t)}f(r,t)dr \,. \qquad (13)$$

Below, we obtain time-series with arbitrary (in general) non-stationary distribution function.

EVOLUTION EQUATION FOR AVERAGE SIR

The evolution equation for mean SIR in (13) is given by

$$N\frac{ds}{dt}=\int_V\frac{\varphi(r)}{U(r,t)}\frac{\partial f(r,t)}{\partial t}d^3r- \qquad (14)$$
$$-\int_V\frac{\varphi(r)}{U^2(r,t)}\frac{\partial U(r,t)}{\partial t}f(r,t)d^3r \,.$$

We assume that the pdf is zero at the boundary of *V*. The second term in the right part of (14) after substituting the derivative for $\frac{\partial f}{\partial t}$ from (1) and taking into account the expression (12) is represented in the form

$$-\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}\frac{\partial U(\mathbf{r},t)}{\partial t}f(\mathbf{r},t)d\mathbf{r}=$$
$$=-\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}\left(\frac{\partial}{\partial t}\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)f(\mathbf{r}',t)d\mathbf{r}'\right)f(\mathbf{r},t)d\mathbf{r}=$$
$$=-\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}\left(\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)\frac{\partial}{\partial t}f(\mathbf{r}',t)d\mathbf{r}'\right)f(\mathbf{r},t)d\mathbf{r}=$$
$$=-\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}\left(\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)\cdot\right.$$
$$\left.\cdot\left(-div_{\mathbf{r}'}(\mathbf{u}(\mathbf{r}',t)f(\mathbf{r}',t))+\frac{\lambda(t)}{2}\Delta_{\mathbf{r}'}f(\mathbf{r}',t)\right)d\mathbf{r}'\right)f(\mathbf{r},t)d\mathbf{r}.$$

After integration by parts with taking into account that the boundary probability flow is equal to zero, the first internal integral in the last expression has the form

$$\int_V\left(\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)\left(div_{\mathbf{r}'}(\mathbf{u}(\mathbf{r}',t)f(\mathbf{r}',t))\right)d\mathbf{r}'\right)\cdot$$
$$\cdot\frac{\varphi(r)}{U^2(\mathbf{r},t)}f(\mathbf{r},t)d\mathbf{r}=$$
$$=-\int_V\left(\int_V f(\mathbf{r}',t)\mathbf{u}(\mathbf{r}',t)grad_{\mathbf{r}'}\varphi(|\mathbf{r}-\mathbf{r}'|)d\mathbf{r}'\right)\cdot$$
$$\cdot\frac{\varphi(r)}{U^2(\mathbf{r},t)}f(\mathbf{r},t)d\mathbf{r}=$$
$$=\int_V\left(\int_V f(\mathbf{r}',t)\mathbf{u}(\mathbf{r}',t)grad_{\mathbf{r}}\varphi(|\mathbf{r}-\mathbf{r}'|)d\mathbf{r}'\right)\cdot$$
$$\cdot\frac{\varphi(r)}{U^2(\mathbf{r},t)}f(\mathbf{r},t)d\mathbf{r}=$$
$$=\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}div_{\mathbf{r}}\left(\int_V f(\mathbf{r}',t)\mathbf{u}(\mathbf{r}',t)\varphi(|\mathbf{r}-\mathbf{r}'|)d\mathbf{r}'\right)f(\mathbf{r},t)d\mathbf{r}=$$
$$=\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}f(\mathbf{r},t)div_{\mathbf{r}}J(\mathbf{r},t)d\mathbf{r},$$

where

$$\mathbf{J}(\mathbf{r},t)=\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)\mathbf{u}(\mathbf{r}',t)f(\mathbf{r}',t)d^3r' \,.$$

We assume, that all integrals uniformly converge by $\mathbf{r}$ and *t*, which means that the negative signs can be taken out. In the same way we get

$$-\frac{\lambda(t)}{2}\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}\left(\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)\left(\Delta_{\mathbf{r}'}f(\mathbf{r}',t)\right)d\mathbf{r}'\right)f(\mathbf{r},t)d\mathbf{r}=$$
$$=-\frac{\lambda(t)}{2}\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}f(\mathbf{r},t)\Delta_{\mathbf{r}}\left(\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)f(\mathbf{r}',t)d\mathbf{r}'\right)d\mathbf{r}=$$
$$=-\frac{\lambda(t)}{2}\int_V\frac{\varphi(r)}{U^2(\mathbf{r},t)}f(\mathbf{r},t)\Delta_{\mathbf{r}}U(\mathbf{r},t)d\mathbf{r},$$

so that finally we have

$$\frac{\partial U}{\partial t}=\frac{\lambda}{2}\Delta U-div\mathbf{J}, \qquad (15)$$
$$\mathbf{J}=\int_V\varphi(|\mathbf{r}-\mathbf{r}'|)\mathbf{u}(\mathbf{r}',t)f(\mathbf{r}',t)d^3r' \,.$$

This means that the mean interference changes over time in the same way as its pdf, i.e., according to the diffusion equation with the same ratio as in (1).

Next, applying (1) to the first term in (14) we get

$$\int_V\frac{\varphi(r)}{U(r,t)}\frac{\partial f(r,t)}{\partial t}d^3r=$$
$$=-\int_V\frac{\varphi(r)}{U(r,t)}div(\mathbf{u}f)d^3r+\frac{\lambda}{2}\int_V\frac{\varphi(r)}{U(r,t)}\Delta fd^3r \,.$$

After integrating by parts we obtain

$$\int_V\frac{\varphi}{U}\frac{\partial f}{\partial t}d^3r=\int_V f\cdot\left(\mathbf{u}\nabla+\frac{\lambda}{2}\Delta\right)\left(\frac{\varphi}{U}\right)d^3r \,.$$

As a result, (14) takes the following form

$$N\frac{ds}{dt}=\int_V\left(\left(\mathbf{u}\nabla+\frac{\lambda}{2}\Delta\right)\left(\frac{\varphi}{U}\right)\right)f(r,t)d^3r- \qquad (16)$$
$$-\int_V\frac{\varphi}{U^2}\left(\frac{\lambda}{2}\Delta U-div\mathbf{J}\right)f(r,t)d^3r \,.$$

The latter equation is non-linear with respect to the density of the ensemble of sample trajectories and cannot be solved in closed-form. However, the structure of (16) scheds the light on various properties of SIR. The term $\mathbf{u}\nabla(\varphi/U)$ represents the influence of drift in the Fokker-Planck equation (1) on variance of SIR, and the term $(\varphi/U^2)div\mathbf{J}$ affects the mean SIR $U(r,t)$. These terms have different signes, so the total transport effect of the drift $\mathbf{u}\nabla(\varphi/U)+(\varphi/U^2)div\mathbf{J}$ after integration with density $f(r,t)$ is negligible in some situations, that can be reffered to as "zero SIR flow". This phenomenon is observed when the scalar product of drift and radius-vector $\mathbf{ur}$ is zero. In this case, SIR variance is due to diffusion effect only. However, in practice, such ideal situations are rarely observed. The reason is that for realistic environments such as malls, stadiums, etc., there is always a non-zero time-dependent drift, corresponding to, e.g., lunch breaks, other intermissions, when human flow changes. In this cases numerical simulations are of special importance.

SIMULATION OF THE SIR DISTRIBUTION

In this section we provide a numerical illustrations, showing the dependence of SIR on mobility parameters. In our statistical experiment we generate a set of time series, for which the sampling density in the window of T=1000 steps evolves on the horizon T from red curve to green curve. This evolution is depicted on Fig. 1. The red curve is the initial pdf in (1) and the green curve is its position after 1000 time steps for given drift $u(x,t)$ and diffusion $\lambda(t)$ functions. The pdfs have similar shape, but the quantitative difference is greater than the average difference between finite stationary samples of the same length. This difference is attributed to the effects of non-stationarity.

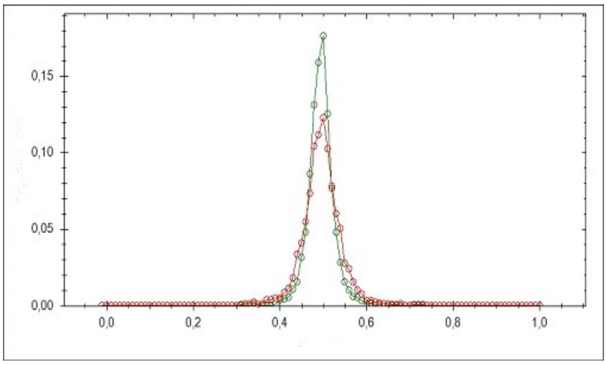


Fig. 1. Initial and final pdfs of coordinates increments.

Further, following the method of modeling of random trajectories in (4), we generated 1000 samples for each of the three dimensions in a cube with a side of 10 and perfectly reflecting boundaries. The resulting set of trajectories is interpreted as the movement path of points for witch we compute the SIR, using the function $\varphi(r)=1/(r^2+a^2)$. with $a=0{,}02$. A non-zero parameter a in this formula corresponds to the minimum distance between communicating nodes.

The SIR time series for this model are shown in Fig. 2. Here, horizontal axis represents the modeling time. As one may observe, approximately 80% of SIR values, presented on the vertical axis, are less than 0.2. At the same time, occasionally, there are extremely large outliners corresponding to extremely good spatial locations of nodes.

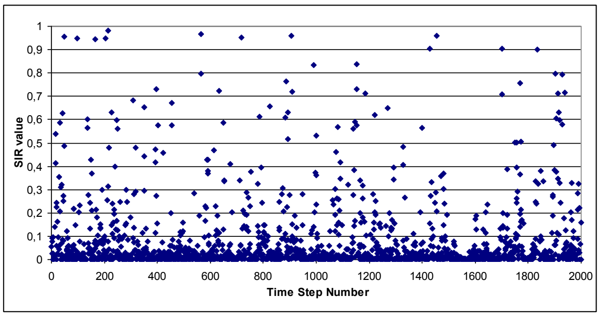


Fig. 2. SIR time series simulation.

Fig. 3 shows theSIR pdf for different values of diffusion and zero drift. For the sake of visualization, we show only the values that are less than 0.8. As parameter λ increases, the sample pdf of SIR shifts upwards, which means a shift toward lower values of SIR.

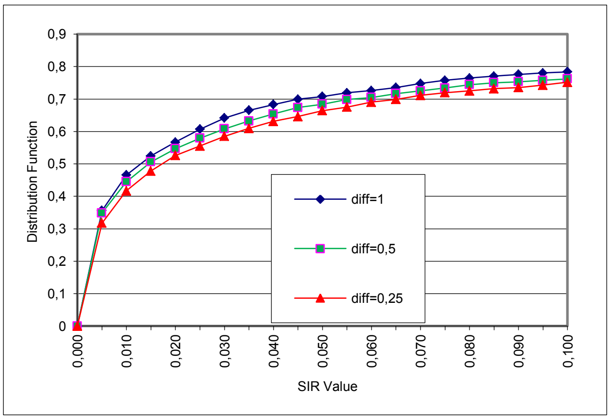


Fig. 3. SIR distribution as a function of the diffusion parameter $\lambda/2$ and zero drift speed.

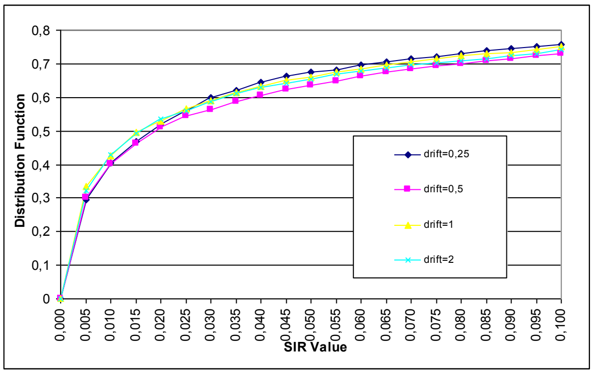


Fig. 4. The SIR distribution as a function of drift with zero diffusion.

Varying the drift speed shows another effect, illustrated in Fig. 4. One may observe the periodic dependence of sample pdf on the drift speed.

CONCLUSION

In this paper we have analyzed interference and SIR for D2D deployment as a function of random movement of nodes. We have first demonstrated that these metrics can be obtained in closed-form when the underlying mobility of users is a stationary. Further, assuming non-stationary movement of users, we have obtained kinetic equation representation of SIR showing that it cannot be solved in closed-form.

To obtain SIR properties we have debeloped a numerical simulsation methodology. In the provided numerical example we had shown that non-stationarty results in fundamentally different distributions of SIR compared to the stationary case. The resulting pdf is highly sensitive to the drift and diffusion coefficient that are used to represent the motion of users.

ACKNOWLEDGEMENTS

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008) and RFBR (projects 16-07-00766, 17-07-00845).

REFERENCES

Bosov A.D., Kalmetev R.Sh. and Orlov Yu.N. 2014. Modelirovanie nestatitsionarnogo vremenneogo ryada s zadannyimi svoystvami vyiborochnogo raspredeleniya (in Russian). *Matematicheskoe modelirovanie*, No.3, 97-107.

Gaidamaka Yu., Orlov Yu., Fedorov S., Samuylov A., Molchanov D. 2016. Simulation of Devices Mobility to Estimate Wireless Channel Quality Metrics in 5G Network. *Proc. ICNAAM*, *September 19-25, Rhodes, Greece*.

Fedorov S.L., Orlov Yu.N. 2016. Metody chislennogo modelirovaniya processov nestatsionarnogo sluchajnogo bluzhdaniya (in Russian). – Moscow: MIPT.

Hesham E.S., Ekram H., Mohamed-Slim A. 2014. Analytical Modeling of Mode Selection and Power Control for Underlay D2D Communication in Cellular Networks. *IEEE Transactions on Communications*, 62, No. 11, 4147-4161.

Orlov Yu.N. 2014. Kineticheskie metody issledovanija nestacionarnykh vremennykh riadov (in Russian). – Moscow: MIPT.

Orlov Yu.N., Fedorov S.L. 2016. Generation of non-stationary time-series trajectories on the basis of Fokker-Planck equation (in Russian). *Trudy MIPT*, 8, No. 2, 126-133.

Orlov Yu.N., Fedorov S.L. 2016. Modelirovanie ansamblya nestacionarnyh traektorij s pomoshch'yu uravneniya Fokkera-Planka (in Russian). *Zhurnal Srednevolzhskogo matematicheskogo obshchestva*, No1.

Samuylov A., Ometov A., Begishev V., Kovalchukov R., Moltchanov D., Gaidamaka Yu., Samouylov K., Andreev S. and Koucheryavy Ye. 2015. “Analytical Performance Estimation of Network-Assisted D2D Communications in Urban Scenarios with Rectangular Cells,” in *Transactions on Emerging Telecommunications Technologies*. – John Wiley & Sons.

AUTHOR BIOGRAPHIES

SERGEY L. FEDOROV was born in Moscow, Russia and went to the Moscow Institute of Physics and Technology, where he studied computer science and mathematical modeling. He worked for a three years in Dorodnitsyn Computing Centre of RAS and after that three years in Keldysh Institute of Applied Mathematics of RAS. His specialty is the mathematical modeling, and numerical methods in the field of simulation of non-stationary stochastic processes, images recognition and kinetic equations investigation. His e-mail address is: fyodor-on@mail.ru

YURII N. ORLOV was born in Mytischy, Moscow region, Russia and at 1987 finished the Moscow Institute of Physics and Technology, where he studied statistical physics and kinetic theory. After post-graduation at the Keldysh Institute of Applied Mathematics he obtained the Cand.Sc. degree in 1992 and doctor degree on mathematical physics in 2007. His interests area belongs to non-stationary stochastic processes and time-series. His e-mail address is: ov3159f@yandex.ru.

ANDREY SAMUYLOV received the Ms.C. in Applied Mathematics and Cand.Sc. in Physics and Mathematics from the RUDN University, Russia, in 2012 and 2015, respectively. Since 2015 he is working at Tampere University of Technology as a researcher, working on performance analysis of various 5G wireless networks technologies. His research interests include P2P networks performance analysis, performance evaluation of wireless networks with enabled D2D communications, and mmWave band communications. His e-mail address is: andrey.samuylov@tut.fi.

DMITRI MOLTCHANOV received the M.Sc. and Cand.Sc. degrees from the St. Petersburg State University of Telecommunications, Russia, in 2000 and 2002, respectively, and the Ph.D. degree from the Tampere University of Technology in 2006. He is a Senior Research Scientist with the Department of Electronics and Communications Engineering, Tampere University of Technology, Finland. He has authored over 100 publications. His research interests include performance evaluation and optimization issues of wired and wireless IP networks, Internet traffic dynamics, quality of user experience of real-time applications, and mmWave/terahertz communications systems. He serves as a TPC member in a number of international conferences. His e-mail address is: dmitri.moltchanov@tut.fi.

YULIYA GAIDAMAKA received the Ph.D. in Mathematics from the Peoples' Friendship University of Russia in 2001. Since then, she has been an associate professor in the university's Department of Applied Probability and Informatics. She is the author of more than 50 scientific and conference papers. Her current research focuses on performance analysis of 4G/5G networks including M2M- and D2D communications, queuing theory, and mathematical modeling of communication networks. Her e-mail address is gaydamaka_yuv@rudn.university.

KONSTANTIN SAMOUYLOV received his Ph.D. degree from the Moscow State University and Doctor of Sciences degree from the Moscow Technical University of Communications and Informatics. In 1996 he became the head of the Telecommunication Systems Department of RUDN University, Russia, and later, in 2014, he became the head of Department of Applied Informatics and Probability Theory. During the last two decades, Konstantin Samouylov has been conducting research projects for the Helsinki and Lappeenranta Universities of Technology, Moscow Central Science Research Telecommunication Institute, several Institutes of Russian Academy of Sciences and a number of Russian network operators. His current research interests are performance analysis of 4G networks (LTE, WiMAX), signalling network (SIP) planning, and cloud computing. He has written more than 150 scientific and technical papers and 5 books. His e-mail address is: samuylov_ke@rudn.university.

SERGEY YA. SHORGIN received the Doctor of Sciences degree in Physics and Mathematics in 1997. Since 2003, he is a professor, and since 2015 he is a Deputy Director of Federal Research Center “Computer Science and Control”, Russian Academy of Sciences. He is the author of more than 100 scientific and conference papers and coauthor of three monographs. His research interests include probability theory, modeling complex systems, actuarial and financial mathematics. His email address is: sshorgin@ipiran.ru.

TIME-DEPENDENT SIR MODELING FOR D2D COMMUNICATIONS IN INDOOR DEPLOYMENTS

Yurii Orlov*†, Dmitry Zenyuk*, Andrey Samuylov†‡, Dmitri Moltchanov†‡, Sergey Andreev‡,
Oxana Romashkova§, Yuliya Gaidamaka†, Konstantin Samouylov†

*Deparment of Kinetic Equations

Keldysh Institute of Applied Mathematics

Miusskaya Sq. 4, Moscow, 125047, Russian Federation
fyodor-on@mail.ru, ov3159f@yandex.ru

†Department of Applied Probability and Informatics
Peoples' Friendship University of Russia (RUDN University)
Miklukho-Maklaya St 6, Moscow, 117198, Russian Federation
ov3159f@yandex.ru, {samuylov_ak, molchanov_da, gaydamaka_yuv, samuylov_ke}@rudn.university

‡Department of Electronics and Communications Engineering
Tampere University of Technology
Korkeakoulunkatu 10, Tampere, 33720, Finland
{andrey.samuylov, dmitri.moltchanov, sergey.andreev}@tut.fi

§Department of Applied Informatics

Moscow City University
2nd Selskohoziaystvenny Pr, 4, Moscow, 129226, Russian Federation
ox-rom@yandex.ru

KEYWORDS

D2D communications, cellular network, fractal motion, SIR probability density function, time-dependence

ABSTRACT

Device-to-device (D2D) communications is expected to become an integral part of the future 5G cellular systems. The connectivity performance of D2D sessions is heavily affected by the dynamic changes in the signal-to-interference ratio (SIR) caused by random movement of communicating pairs over a certain bounded area of interest. In this paper, taking into account the recent findings on the movement of users over a landscape, we characterize the probability density function (pdf) of SIR under stochastic motion of communicating D2D pairs on planar fractals. We demonstarte that the pdf of SIR depends on the fractal dimension and the spatial density of trajectories. The proposed model can be further used to investigate time-dependent user-centric performance metrics including the link data rate and the outage time.

INTRODUCTION

Fifth-generation (5G) mobile cellular systems are aimed at 1000x rate boost compared to the current 4G mobile networks (Andrews 2012). This dramatic increase in system capacity is expected to be achieved with a combination of additional bandwidth in the millimeter-wave bands, improved physical layer techniques, and fundamentally new networking mechanisms. Direct device-to-device (D2D) communications is considered as one of the techniques allowing to drastically improve the degrees of spatial reuse in 5G systems and thus enhance the overall system capacity (Asadi 2014).

The widely-accepted metrics for performance assessment of wireless connectivity in interference-limited cellular systems is the signal-to-interference ratio (SIR) defined as a ratio between the received signal strength and the aggregate interference. Using the Shannon-Hartley theorem (Proakis 1994), SIR can then be used to estimate the theoretically achivable capacity of a link. The latter serves as an upper bound for further development of the optimal modulation and coding schemes (MCSs).

The performance of 5G systems is conventionally assessed by utilizing the tools of stochastic geometry (Bacelli 2010, Haenggi 2012). Accordingly, communicating pairs over the same operating frequency are modeled with a stationary isopropic spatial point process, such as Poisson point process. The performance is then characterized with the standard methods of geometric probability that estimate SIR as a function of path loss models and random distances between points.

Even though the described approach allows to capture the spatially-averaged probability density function (pdf) of SIR, it effectively neglects the impact of movement of users over the landscape. At the same time, the movement of users is an inherent property that may affect the time-dependent SIR behavior. That is, even when the time-averaged behavior is satisfactory, there could still be occasional outages caused by positioning of the communicating pairs with respect to each other. The frequency of outages and their durations are critical parameters that affect user satisfaction of a service.

Proceedings 31st European Conference on Modelling and Simulation ©ECMS Zita Zoltay Paprika, Péter Horák, Kata Váradi, Péter Tamás Zwierczyk, Ágnes Vidovics-Dancs, János Péter Rádics (Editors)
ISBN: 978-0-9932440-4-9/ ISBN: 978-0-9932440-5-6 (CD)

In this paper, we formulate a new framework to assess time-dependent behavior of SIR in a moving field of mutually interfering D2D pairs. The movement of transmitters is modeled as random-walk trajectories on planar fractal sets. Based on a novel technique for non-stationary random walks originally reported in (Orlov 2014, Orlov and Fedorov 2016), we demonstrate that the shape of the pdf of SIR is heavily affected by the fractal dimension and the spatial density of trajectories.

Our proposed technique is based on solving non-stationary kinetic equations for the sample distributions of device coordinates while being able to describe spurious correlations. Over smaller sample sizes, these processes can be considered as stationary random walks with long-range dependences. Such walks are characterized by the pdfs with "fat tails" and often emerge in dynamic processes on fractal sets. In practice, such random walks can correspond to movement of people in large shopping malls, where relatively slow travel of customers in front of display windows is interchanged with rapid movement between them using e.g., elevators. The analysis of SIR under this type of mobility is an important theoretical problem in 5G wireless networks (Hesham et al. 2014, Samuylov at al. 2015, Orlov et al., 2016).

RANDOM WALKS ON FRACTAL SETS

The modeling methodology proposed in this study originates from (Orlov 2014, Orlov and Fedorov 2016). Particularly, a simulation technique for random-walk trajectories on Cantor set has been proposed in (Zenyuk et al. 2013). Further generalization to a broader class of fractal sets based on iterated function systems has been developed in (Zenyuk and Orlov 2016). Below, we briefly introduce the basic definitions and outline our proposed methodology.

The simplest way to construct a one-dimensional Cantor C_λ set is to iteratively eliminate open middle intervals of length λ from the set of closed intervals. Specifically, one starts by discarding the middle interval of $[0,1]$, thus leaving two line segments; then, the middle part of the remaining segments is removed, hence leaving four line segments. The process is continued *ad infinitum*. At the *n*-th iteration, 2^{n-1} open intervals each having the length of β^{n-1} are eliminated, where $\beta = (1-\lambda)/2$. Therefore, the total length, l_n, of the removed interval is given by

$$l_n = \lambda(2\beta)^{n-1}, \quad \beta = \frac{1-\lambda}{2}, \tag{1}$$

which implies that C_λ has zero Lebesgue measure. The Hausdorff dimension of Cantor set is $-\log_\beta 2$.

Furhter, every point of Cantor set can be represented as

$$x = (1-\beta)\sum_{k=1}^{\infty} b_k \beta^k, \quad b_k \in \{0,1\}. \tag{2}$$

The popular version of C_λ is a ternary Cantor set with $\lambda = \beta = 1/3$ characterized by a closed-form expression

$$C_{1/3} = [0,1] \setminus \bigcup_{m=1}^{\infty} \bigcup_{k=0}^{3^{m-1}-1} \left(\frac{3k+1}{3^m}; \frac{3k+2}{3^m} \right). \tag{3}$$

From (2), it follows directly that the expansion of base 3 of every point in the ternary Cantor set does not contain a unit. The form (2) is also a bijection from Cantor set onto the set of all infinite binary sequences enabling an explicit synthesis of random walk on C_λ.

Consider a random process c_t in a discrete time $t \in N$,

$$c_t = (1-\beta)\sum_{k=0}^{\infty} \xi_t^{(k)} \beta^k, \tag{4}$$

where $\xi_t^{(k)}$ are the discrete random variables having Bernoulli distribution. The simplest type of a random walk is obtained, when $\xi_t^{(k)}$ are mutually independent for every fixed *t* and every $\xi_t^{(k)}$ is a time-homogeneous Markov chain. Without the loss of generality, assume that the process starts from zero, $\forall k \; P\left(\xi_1^{(k)} = 0\right) = 1$ for every *k*. We also introduce the following notaton

$$\begin{aligned} &P(\xi_t^{(k)} = 0) = u_t, \\ &P(\xi_t^{(k)} = 1) = w_t, \\ &u_t + w_t = 1. \end{aligned} \tag{5}$$

The state probabilities satisfy a system of equations

$$\begin{pmatrix} u_{t+1} \\ w_{t+1} \end{pmatrix} = \Pi^T \begin{pmatrix} u_t \\ w_t \end{pmatrix}, \quad \Pi = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}, \tag{6}$$

where p_{ij} are the transition probabilities from $\{\xi_t^{(k)} = i\}$ to $\{\xi_{t+1}^{(k)} = j\}$. If $p_{00} = p_{11} = 1$, this system has the only trivial solution $u_t = 1, w_t = 0$, which implies that $c_t \equiv 0$. Otherwise, the solution of (6) is

$$\begin{aligned} u_t &= \frac{p_{10}}{p_{01}+p_{10}} + \frac{p_{01}}{p_{01}+p_{10}}(\mathrm{Tr}\Pi - 1)^t, \\ w_t &= \frac{p_{01}}{p_{01}+p_{10}} - \frac{p_{01}}{p_{01}+p_{10}}(\mathrm{Tr}\Pi - 1)^t, \end{aligned} \tag{7}$$

and the process c_t is strictly stationary when $\mathrm{Tr}\Pi = 1$.

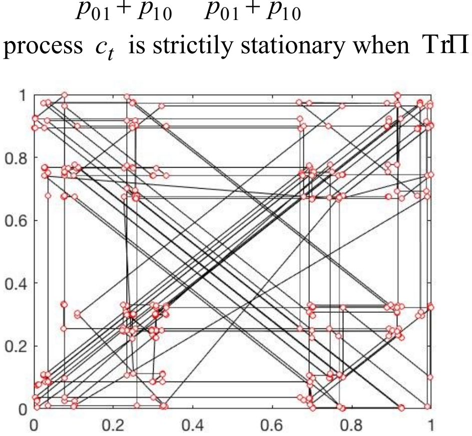


Fig. 1: Sample trajectories on 2D Cantor set.

Observe that many fractal sets, including the Cantor set, are attractors of the iterated systems of contraction mappings, see e.g., (Hutchinson 1981). Random walks on such fractal sets can be constructed in a similar way as described above. Sample paths of trajectories for several fractal sets inside a unit square are depicted in Fig. 1-3. A sample trajectory of Brownian motion is given in Fig. 4 for the sake of comparison.

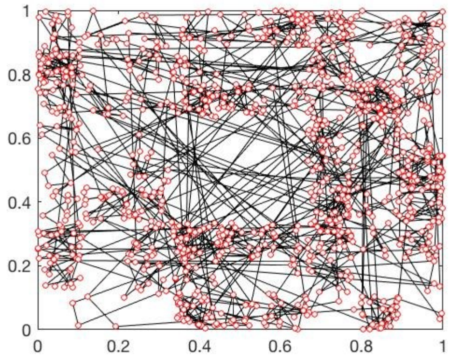


Fig. 2: Sample trajectory on Sierpinski square.

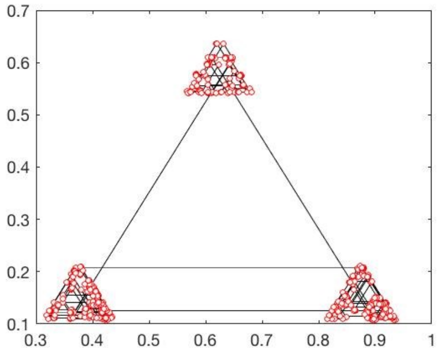


Fig. 3: Sample trajectory on Sierpinski triangle.

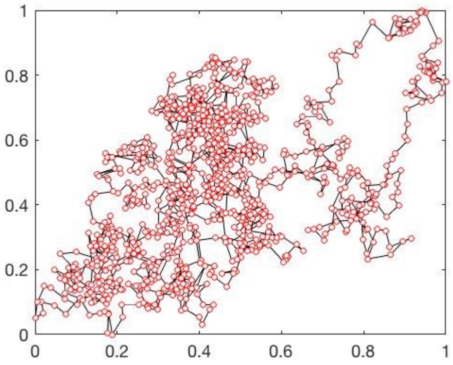


Fig. 4: Sample trajectory of Brownian motion.

As one may observe, the random walks on fractale sets allow to represent a number of practical situations when users are moving in certain indoor areas. A characteristic example is the movement of users inside shopping malls.

In the following sections, we investigate a dependence of SIR on the geometric properties of the underlying set. We demonstrate that in case of random walks on fractals, the SIR is affected not only by the Hausdorff dimension of the set, but also by its shape.

SIR DISTRIBUTION DENSITY

Our system of interest consists of $N+2$ moving devices. Their trajectories are assumed to be mutually independent. A pair $(x_i(t), y_i(t))$ defines the position of the *i*-th device at time step *t*. Squared distance between two particular devices at time step *t* equals to

$$r_{ij}^2(t) = (x_i(t) - x_j(t))^2 + (y_i(t) - y_j(t))^2 . \qquad (8)$$

The path loss value is chosen to be proportional to the inverse square of distance between the devices i.e., $\varphi_{ij} \equiv \varphi(r_{ij}) = 1/r_{ij}^2$. The SIR for every pair of devices is calculated according to

$$S(\mathbf{r}_1, \mathbf{r}_2; t) = \frac{\varphi_{12}(t)}{\sum_{j=3}^{N+2} \varphi_{1j}(t)} , \qquad (9)$$

where the sum in the denominator is the aggregate interference at the device.

The mean interference is then

$$U(\mathbf{r}, t) = \int_V \varphi(|\mathbf{r} - \mathbf{r}'|, t) \rho(\mathbf{r}') d\mathbf{r}' , \qquad (10)$$

where $\rho(\mathbf{r}')$ is the joint pdf of distances.

We now set *i*=1, *j*=2, $\mathbf{r} = \mathbf{r}_{12}$, and calculate all distances in the moving system of coordinates associated with the second device. Observe that (9) can be rewritten as

$$S \equiv S(\mathbf{r}, t) = \frac{\varphi(r, t)}{NU(\mathbf{r}, t)} . \qquad (11)$$

We are interested in the mean SIR over the ensemble of trajectories e.g., defined as,

$$q(t) = \int_V S(\mathbf{r}, t) \rho(\mathbf{r}) d\mathbf{r} . \qquad (12)$$

The time series of SIR observaions, $q(t)$, obtained over one experiment with *N*=100 is shown in Fig. 5-8. According to these illustrations, the values of SIR on different fractal sets show a similar behavior. However, as we demonstarte in what follows they have drastically different statistical properties.

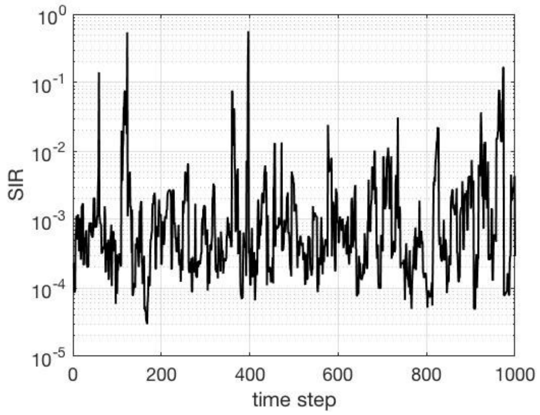


Fig. 5. Time series of SIR for 2D Cantor set.

Time series of SIR obtained in one experiment with *N*=100 is shown in Fig. 9. As one may observe, the structure of time series for random walks on fractals is different from that observed for Brownian motion.

The sample pdfs of SIR with $N = 10$ are depicted in Fig. 10. We specifically enforced a special non-uniform division of the value sets to obtain pdfs with coinciding modes. Note that this procedure does not affect the qualitative structure of the pdfs.

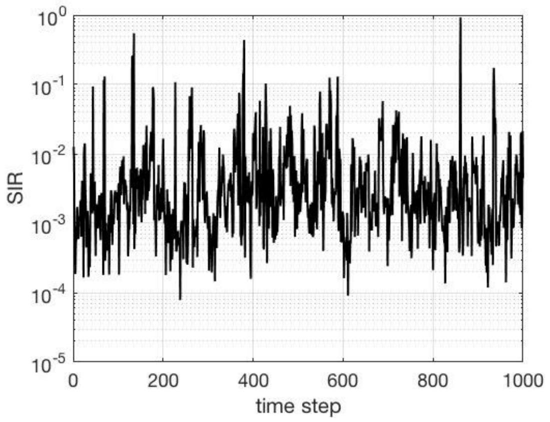


Fig. 6. Time series of SIR for Sierpinski square.

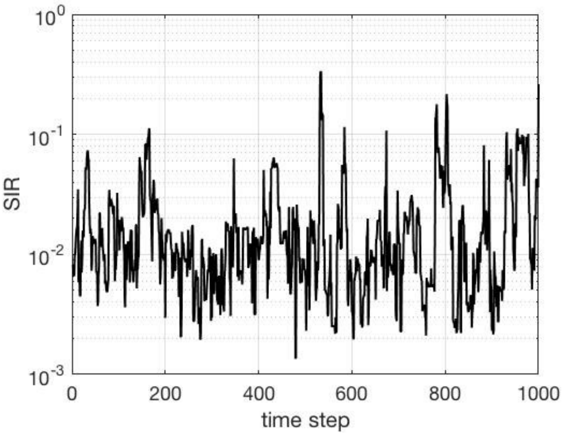


Fig. 7. Time series of SIR for Sierpinski triangle.

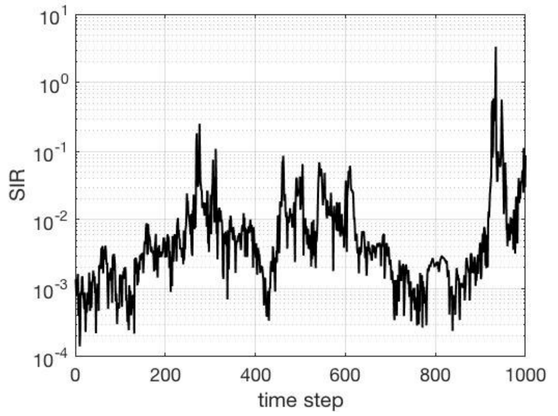


Fig. 8. Time series of SIR for Brownian motion.

Analyzing Fig. 9, one may observe that the geometry of the underlying set plays an important role affecting the structure of the pdf of SIR. Particularly, Sierpinski triangle set, whose sample paths are illustrated in Fig. 3, generates a distribution that significantly differs from other sets. The Hausdorff dimension of the underlying set also affects the statistical properties of SIR. In order to analyze this effect, we introduce the so-called *consistent stationary level* (Orlov 2016). According to its definition, this level is obtained as a solution to the following transcendent equation

$$1-\varepsilon = K\left(\sqrt{\frac{T}{2}}\varepsilon\right) \tag{13}$$

with respect to ε, where *T* is the overall number of observations and $K(z)$ is the Kolmogorov distribution function. The solution of (13) is equal to the limiting error rate of the statistical test under the null hypothesis. It consists in that the two samples of length *T* are drawn from the same distribution with the significance level of $\varepsilon(T)$, as the number of independent experiments tends to infinity.

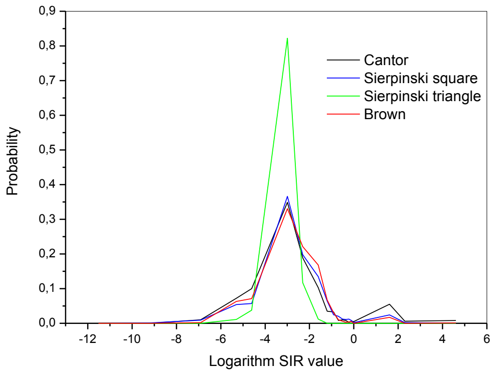


Fig. 9. Empirical SIR pdfs for different underlying sets

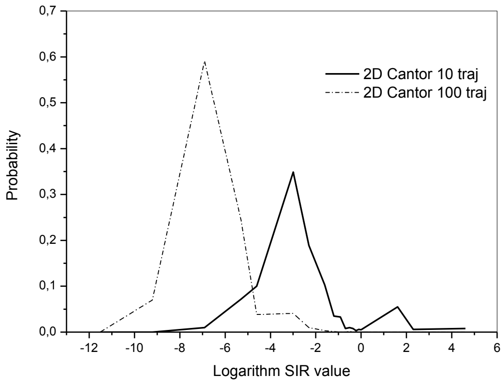


Fig. 10. Empirical SIR pdfs for Cantor square.

For example, for $T = 1000$ the consistent stationary level is $\varepsilon_0 = 0.059$. We define the consistent stationary level for the distance statistics similarly, that is, it is a solution to the following equation

$$G_T(\rho) = 1-\rho \tag{14}$$

with respect to $\rho^*(T)$, where $G_T(\rho)$ is the cumulative distribution function of $\rho_T = \sup_x |F_{1,T}(x) - F_{2,T}(x)|$

that captures the distance between two samples in *C*-norm. Combining these two values, the non-stationary index can be defined as

$$J(T)=\frac{\rho^{*}(T)}{\varepsilon(T)}. \qquad (15)$$

When $J(T)\le 1$, the time series is considered stationary. Otherwise, it is non-stationary, as the null hypothesis of homogeneity at different time steps is rejected.

Our results show that the non-stationary indices for SIR samples on different sets with a square shape (2D Cantor set, Sierpinski square, Brownian motion case) are all greater than 1 and distinct. More importantly, even for the random walks on the same fractal set, the SIR samples differ more than those in the stationary case: for 2D Cantor set *J*(100)=2.5, for Sierpinski square *J*(100)=2.7, and for Sierpinski triangle *J*(100) = 3.1. As the number of observations grows, the index of non-stationarity decreases monotonically. The stationary statistics is achieved when *T* exceeds 10000. Finally, we note that SIR is inversely proportional to *N*. The results of our numerical simulation are in good agreement with this fact, as one may observe in Fig. 10.

CONCLUSIONS

In this paper, we analyzed the time-dependent behavior of SIR in D2D environments for different random walks on fractal sets. We demonstrated that the pdf of SIR depends on the fractal dimension and the spatial density of trajectories. We also showed that the pdfs of SIR in case of fractale random-walk models are indistinguishable from those for the non-stationary random-walk model.

Our proposed model can be used to investigate the time-dependent user-centric performance metrics including the link data rate and the outage time in complex indoor environments, such as shopping malls, where the movement of users is not purely stochastic but rather is modulated by dedicated attractor points.

ACKNOWLEDGEMENTS

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (the Agreement number 02.a03.21.0008) and RFBR (projects 16-07-00766, 17-07-00845).

REFERENCES

Hesham E.S., Ekram H., Mohamed-Slim A. 2014. Analytical Modeling of Mode Selection and Power Control for Underlay D2D Communication in Cellular Networks. *IEEE Transactions on Communications*, 62, No. 11, 4147-4161.

Hutchinson J.E. 1981. Fractals and self-similarity. *Indiana University Mathematics* Journal, 30, No. 5, 713-747.

Orlov Yu.N. 2014. Kineticheskie metody issledovanija nestacionarnykh vremennykh riadov (in Russian). – Moscow: MIPT.

Orlov Yu.N., Fedorov S.L. 2016. Distributions functionals modeling on the ensemble of non-stationary stochastic process trajectories (in Russian). *Preprints KIAM of RAS*, No. 101.

Orlov Yu.N., Fedorov S.L. 2016. Generation of non-stationary time-series trajectories on the basis of Fokker-Planck equation (in Russian). *Trudy MIPT*, 8, No. 2, 126-133.

Orlov Yu., Fedorov S., Samuylov A., Gaidamaka Yu., Molchanov D. 2016. Simulation of Devices Mobility to Estimate Wireless Channel Quality Metrics in 5G Network. *Proc. ICNAAM, September 19-25, Rhodes, Greece.*

Samuylov A., Ometov A., Begishev V., Kovalchukov R., Moltchanov D., Gaidamaka Yu., Samouylov K., Andreev S. Koucheryavy Y. 2015. “Analytical Performance Estimation of Network-Assisted D2D Communications in Urban Scenarios with Rectangular Cells,” in *Transactions on Emerging Telecommunications Technologies*. – John Wiley & Sons.

Zenyuk D.A., Mitin N.A., Orlov Yu.N. 2013. Modeling of random walk on Cantor set (in Russian). *Preprints KIAM of RAS*, No. 31, 18.

Zenyuk D.A., Orlov Yu.N. 2016. Fractional diffusion equation and non-stationary time-series modeling (in Russian). – Moscow: MIPT.

Asadi A., Wang Q., Mancuso V. 2014. A survey on device-to-device communication in cellular networks. IEEE Communications Surveys & Tutorials, 16(4), 1801-1819.

Proakis J. G., Salehi M., Zhou N., Li X. 1994. Communication systems engineering (Vol. 2). New Jersey: Prentice Hall.

Baccelli F., Błaszczyszyn B. 2010. Stochastic geometry and wireless networks: Volume I Applications. Foundations and Trends in Networking, 4(1–2), 1-312.

Haenggi M. (2012). Stochastic geometry for wireless networks. Cambridge University Press.

Andrews J. G., Buzzi S., Choi W., Hanly S. V., Lozano A., Soong A. C., Zhang J. C. 2014. What will 5G be?. IEEE Journal on selected areas in communications, 32(6), 1065-1082.

AUTHOR BIOGRAPHIES

YURII N. ORLOV was born in Mytischy, Moscow region, Russia and at 1987 finished the Moscow Institute of Physics and Technology, where he studied statistical physics and kinetic theory. After post-graduation at the Keldysh Institute of Applied Mathematics he obtained the Cand.Sc. degree in 1992 and doctor degree on mathematical physics in 2007. His interests area belongs to non-stationary stochastic processes and time-series. His e-mail address is: ov3159f@yandex.ru.

DMITRY A. ZENYUK was born in Moscow. He received M.Sc. degree in applied math from Moscow State Technological University STANKIN in 2011. After graduation he attended postgraduate course in applied math and numerical simulation at Keldysh Institute of Applied Math, Russian Academy of Sciences. In his thesis he studied possible applications of fractional calculus to time series analysis and stochastic dynamics on fractal sets. His e-mail address is: eldrich@yandex.ru.

ANDREY SAMUYLOV received the M.Sc. in Applied Mathematics and Cand.Sc. in Physics and Mathematics from the RUDN University, Russia, in 2012 and 2015, respectively. Since 2015 he is working at Tampere University of Technology as a researcher, working on performance analysis of various 5G wireless networks technologies. His research interests include P2P networks performance analysis, performance evaluation of wireless networks with enabled D2D communications, and mmWave band communications. His e-mail address is: andrey.samuylov@tut.fi.

DMITRI MOLTCHANOV received the M.Sc. and Cand.Sc. degrees from the St. Petersburg State University of Telecommunications, Russia, in 2000 and 2002, respectively, and the Ph.D. degree from the Tampere University of Technology in 2006. He is a Senior Research Scientist with the Department of Electronics and Communications Engineering, Tampere University of Technology, Finland. He has authored over 100 publications. His research interests include performance evaluation and optimization issues of wired and wireless IP networks, Internet traffic dynamics, quality of user experience of real-time applications, and mmWave/terahertz communications systems. He serves as a TPC member in a number of international conferences. His e-mail address is: dmitri.moltchanov@tut.fi.

SERGEY ANDREEV is a senior research scientist in the Department of Electronics and Communications Engineering at Tampere University of Technology, Finland. He received the Specialist degree (2006) and the Cand.Sc. degree (2009) both from St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russia, as well as the Ph.D. degree (2012) from Tampere University of Technology. Sergey (co-)authored more than 100 published research works on wireless communications, energy efficiency, heterogeneous networking, cooperative communications, and machine-to-machine applications. His e-mail address is sergey.andreev@tut.fi.

OXANA ROMASHKOVA received her D. of S. in telecommunications in 2003. Since then, she has been a professor of Applied Informatics department of the Moscow City University. She is the author of more than 360 scientific and conference papers. Her current research focuses are telecommunications and information systems and networks for education and transport. Her e-mail is: ox-rom@yandex.ru.

YULIYA GAIDAMAKA received the Ph.D. in Mathematics from the Peoples' Friendship University of Russia in 2001. Since then, she has been an associate professor in the university's Department of Applied Probability and Informatics. She is the author of more than 50 scientific and conference papers. Her current research focuses on performance analysis of 4G/5G networks including M2M- and D2D communications, queuing theory, and mathematical modeling of communication networks. Her e-mail address is gaydamaka_yuv@rudn.university.

KONSTANTIN SAMOUYLOV received his Ph.D. degree from the Moscow State University and Doctor of Sciences degree from the Moscow Technical University of Communications and Informatics. In 1996 he became the head of the Telecommunication Systems Department of RUDN University, Russia, and later, in 2014, he became the head of Department of Applied Informatics and Probability Theory. During the last two decades, Konstantin Samouylov has been conducting research projects for the Helsinki and Lappeenranta Universities of Technology, Moscow Central Science Research Telecommunication Institute, several Institutes of Russian Academy of Sciences and a number of Russian network operators. His current research interests are performance analysis of 4G networks (LTE, WiMAX), signalling network (SIP) planning, and cloud computing. He has written more than 150 scientific and technical papers and 5 books. His e-mail address is: samuylov_ke@rudn.university.

AUTHOR INDEX