European Council for Modelling and Simulation

www.scs-europe.net

ECMS Digital Library

http://www.scs-europe.net/dlib/dl-index.htm

ISBN 978-3-937436-72-2 (Print)
ISBN 978-3-937436-73-9 (CD-ROM)

ISSN 2522-2414 (Print)
ISSN 2522-2422 (Online)
ISSN 2522-2430 (CD-ROM)

Cover pictures front and and back side

© pictures: Kuwait College of Science &Technology

Printed by

Digitaldruck Pirrot GmbH
66125 Sbr.-Dudweiler
Germany

Communications of the ECMS
Volume 35, Issue 1, June 2021

Proceedings of the 35th ECMS International Conference on Modelling and Simulation ECMS 2021

May 31st – June 2nd, 2021
United Kingdom

Edited by:
Khalid Al-Begain
Mauro Iacono
Lelio Campanile
Andrzej Bargiela

Organized by:
ECMS - European Council for Modelling and Simulation

Web-hosted by:
Kuwait College of Science & Technology

Sponsored by:
Kuwait College of Science & Technology

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 2021 ORGANIZATION

Conference Chair

Khalid Al-Begain

Kuwait College of Science & Technology
Kuwait

Conference Co-Chair

Andrzej Bargiela

INFOHUB Ltd.
United Kingdom

Programme Chair

Mauro Iacono

Universita degli Studi della Campania
Luigi Vanvitelli,
Italy

Programme Co-Chair

Lelio Campanile

Universita degli Studi della Campania
Luigi Vanvitelli,
Italy

Editorial Board

INTERNATIONAL PROGRAMME COMMITTEE

Business Process Modelling and Simulation for Industrial Operations

Track Chair: **Romeo Bandinelli**
University of Florence, Italy

Co-Chairs:

Benoit Eynard
Technical University of Compiegne, France

Edward J. Williams
University of Michigan-Dearborn, USA

Finance, Economics and Social Science

Track Chair: **Kata Varadi**
Corvinus University of Budapest, Hungary

Co-Chair: **Agnes Vidovics-Dancs**
Corvinus University of Budapest, Hungary

Open and Collaborative Models and Simulation Methods

Track Chair: **Henrique M. Gaspar**
Norwegian University of Science and Technology, Norway

Co-Chair: **Ricardo da Silva Torres**
Norwegian University of Science and Technology, Norway

Simulation and Optimization

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

Co-Chairs:

Michael Manitz
University of Duisburg-Essen, Germany

Marco Trost
Technical University Dresden, Germany

Finite – Discrete – Element Simulation

Track Chair: **Peter T. Zwierczyk**
Budapest University of Technology and Economics, Hungary

Co-Chair: **Janos P. Radics**
Budapest University of Technology and Economics, Hungary

Machine Learning for Big Data

Track Chair: **Frederic T. Stahl**
DFKI German Research Center for Artificial Intelligence, Germany

Co-Chairs:
Mohamed Gaber
Birmingham City University, United Kingdom

Marwan Hassani
Eindhoven University of Technology, Netherlands

Modeling and Simulation for Performance Evaluation of Computer-based Systems

Track Chair: **Mauro Iacono**
Universita degli Studi della Campania
Luigi Vanvitelli, Italy

Co-Chairs:

Agnieszka Jakobik
Cracow University of Technology, Poland

Lelio Campanile
Universita degli Studi della Campania
Luigi Vanvitelli, Italy

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

IPC Members

Kolos Csaba Agoston, Corvinus University of Budapest, Hungary

Saleh Alaliyat, Norwegian University of Science and Technology, Norway

Donald Davendra, Central Washington University, USA

Julian Englberger, Boston Consulting Group, Germany

Nora Felfoeldi-Szucs, John von Neumann University, Hungary

Massimo Ficco, Universita degli Studi della Campania *Luigi Vanvitelli*, Italy

Icaro Aragao Fonseca, Norwegian University of Science and Technology, Norway

Ingo Frank, OTH Regensburg, Germany

Marton Groza, Karman Mechanics, Hungary

Ibrahim Hameed, Norwegian University of Science and Technology, Norway

Mahmood Hammoodi, University of Babylon, Iraq

Daniel Homolya, MOL Group, Hungary

Bogumil Kaminski, Warsaw School of Economics, Poland

Stelios Kapetanakis, University of Brighton, United Kingdom

Victor Korolev, Moscow State University, Russia

Andreasz Kosztopulosz, University of Szeged, Hungary

Frederick Lange, Maschinenfabrik Reinhausen GmbH, Regensburg, Germany

Giovanna Martinez-Arellano, University of Nottingham, United Kingdom

Michele Mastrioianni, Universita degli Studi della Campania *Luigi Vanvitelli*, Italy

Lusine Meykhanadzhyan, Financial University under the Government of the Russian Federation, Russia

Thiago Gabriel Monteiro, Norwegian University of Science and Technology, Norway

Frank Morelli, University of Applied Sciences in Pforzheim, Germany

Andras Oliver Nemeth, Corvinus University of Budapest, Hungary

Emilia Nemeth-Durko, Corvinus University of Budapest, Hungary

Laszlo Oroszvary, Knorr-Bremse Research, Hungary

Navya Prakash, DFKI GmbH Oldenburg, Germany

Peter Rausch, Nuremberg Institute of Technology, Germany

Simone Righi, University College London, United Kingdom

David Romero, Monterrey Institute of Technology, Mexico

Faruk Savasci, Krones AG, Germany

Oleg Shestakov, Moscow State University, Russia

Carlo Simon, HS Worms University of Applied Sciences, Germany

Janos Simonovics, Budapest University of Technology and Economics, Hungary

Janos Szaz, Corvinus University of Budapest, Hungary

Melinda Szodorai, Corvinus University of Budapest & Keler CCP, Hungary

Jacek Tchorzewski, Cracow University of Technology, Poland

Hajo Terbrack, Technical University Dresden, Germany

Nikolai Ushakov, Norwegian University of Science and Technology, Norway

Kaoly Varadi, Budapest University of Technology and Economics, Hungary

Erzsebet Terez Varga, Corvinus University of Budapest, Hungary

Agnes Vaskoevi, Corvinus University of Budapest, Hungary

Mattis Wolf, DFKI GmbH, Marine Perception, Germany

PREFACE

The 35th ECMS International Conference on Modelling and Simulation (ECMS 2021) comes during very disturbing circumstances due to the continuing pandemic caused by COVID-19. Holding this conference this year (even if virtually) represents both the resilience and insistence of the research community to carry on with the academic duties and the hope that the pandemic will soon be over and that we will meet again in person in the coming years.

ECMS 2021 is dedicated to all those who lost their lives due to Covid-19, to all those who suffered or still suffering due to catching the virus or losing loved ones, to all those heroes at the Frontlines fighting the virus and supporting the sick. A special dedication and appreciation to the research community in Europe and the world, who rose to the challenge and managed to develop the vaccines at an unprecedented speed. The whole world recognised and acknowledged the importance of research. In addition, the simulation and prediction models played a significant role in setting the preventive measures and setting governmental policies in the fight against Covid-19.

ECMS 2021 will also be unique as the virtual organisation allowed Kuwait College of Science and Technology (KCST) to host the conference. KCST is very proud to step in and support the European Council for Modelling and Simulation and ECMS 2021.

Big thank you to the loyal authors who submitted their papers to ECMS 2021. As usual, the Proceedings will be published as part of the Communications of the ECMS series, which enjoys now high recognition among the major indexing agencies.

The Editors

TABLE OF CONTENTS

Business Process Modelling and Simulation for Industrial Operations

Machine Learning for Big Data

Finance and Economics and Social Science

Open and Collaborative Models and Simulation Methods

Finite – Discrete - Element Simulation

Simulation and Optimization

Modeling and Simulation for Performance Evaluation of Computer-based Systems

ECMS 2021

SCIENTIFIC PROGRAM

Business Process Modelling and Simulation for Industrial Operations

APPLICATION OF MULTIAGENT SIMULATION MODELING TO FORECAST MILK RECEIVING PROCESS

Evgeny A. Nazoykin
Moscow State University of Food Production
Institute of Industrial Engineering, Information Technology and Mechatronics
Volokolamskoe highway 11, Moscow, Russia
E-mail: nazojjkinea@mgupp.ru

KEYWORDS

simulation modeling, multiagent modelling, food production facilities, milk receiving, production processes, AnyLogic.

ABSTRACT

This research paper discusses the application of multiagent simulation model of production processes by reference to milk receiving and storage. The paper describes basic parameters for general models, and presents the experiment results obtained during the model processing, as well as the approach towards implementation of the multiagent system employing AnyLogic simulation environment.
The introduction of general technique to plot multiagent simulation models allows to use digital tools to create a virtual copy of the true-life processes with the possibility to provide forecast and identification of the food production industries.
The employment of simulation modeling enables refinement of the production process under study, identification of its weaknesses, and provision of the expert opinion on improvement of production processes and as regards the results of virtual testing thereof.

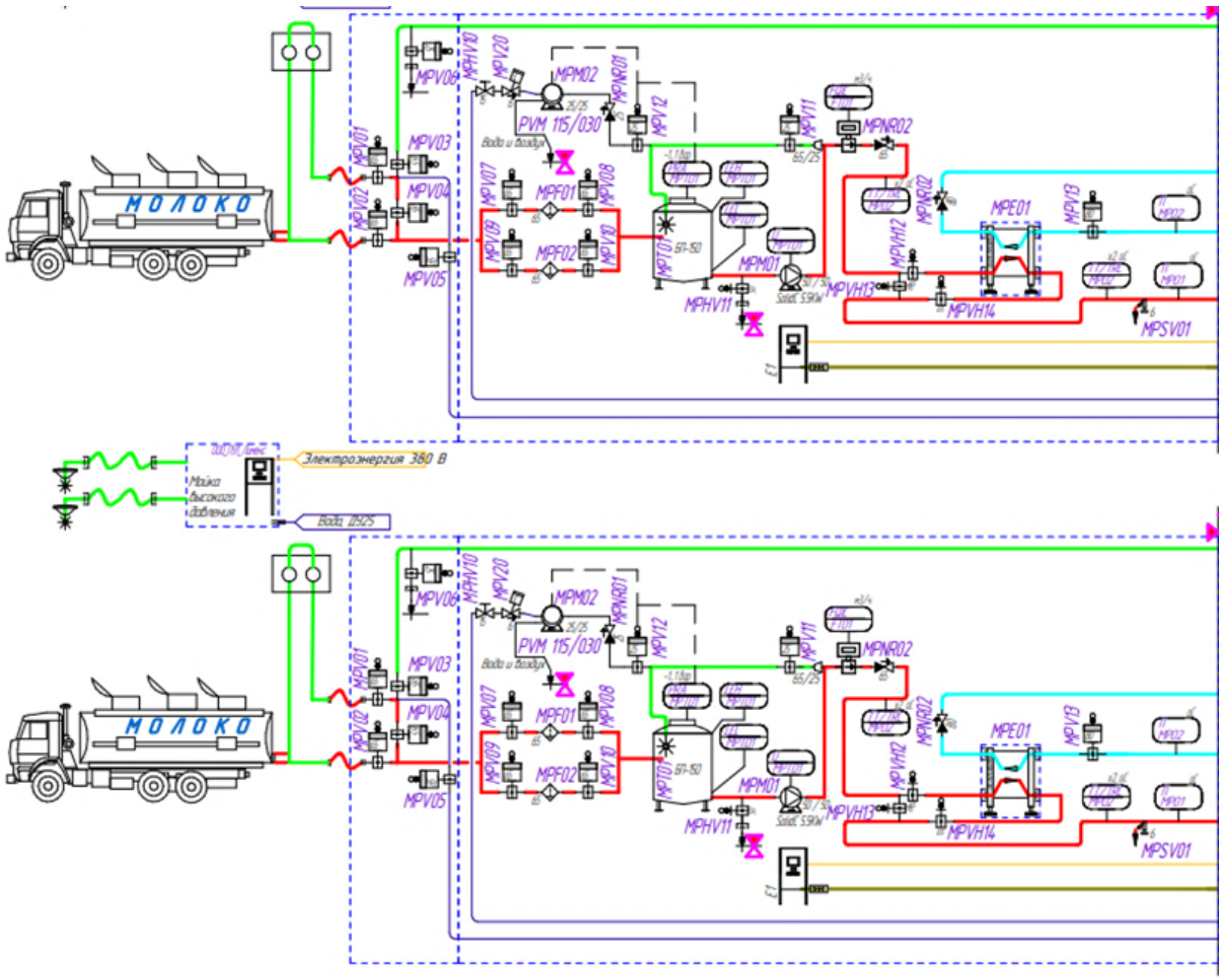
INTRODUCTION

The outdated production facilities in the Russian Federation often do not cope with the tasks the resolution of which is necessary in the context of contemporary production specifics (Sirota, 2006). Nowadays, the impact from external factors, and the demands from industrial customers result in the search for advanced means of production planning or its modernization (Sovetov, 2007).
The most relevant and feasible approach to the virtual representation of the real-life production processes is the employment of multiagent simulation modeling (Gabrin 2004; Karpov 2005; Blagoveschenskaya 2010). This type of modeling aims at carrying out experiments in order to determine the optimum parameters of production processes, to forecast and visualize the results in a user-friendly format, and to identify the weak points and find relevant managerial solutions for their elimination at the design stage without any costly procedures at real enterprises.
One of the production processes under study, for which the application of simulation modeling is relevant, is the milk receiving line (Lisin, 2009). The techniques currently available to analyse (Ponomarev et al., 2006) production processes do not always meet the requirements of both enterprises and business integrators. The multiagent simulation methods allow to evaluate the feasibility of the drafted production plan, calculate the process execution expectancy, predict the production output figures, and identify internal processes in the course of the system modeling, as well as providing guidelines on how to update the target parameters when it comes to the production process advancement.

THEORETICAL BACKGROUND

The automatization flowchart and production processes at the enterprise were reviewed and taken into account (Boev, 2011; Nazoykin, 2019) in the course of multiagent production model elaboration.
The research study is concerned with the milk processing line consisting of three sections: milk receiving area, milk storage, and milk processing area.
The subject of the study is the milk receiving area, as well as the milk tank vehicle supplier and certain varying parameters required to design the model and carry out further experiments. The following varying parameters were identified: number of incoming milk tank vehicles and their volume; time between arrivals of milk tank vehicles; pumps capacity; volume of the tank located at the milk receiving area.
The automatization flowchart and the afore mentioned parameters provide the basis for multiagent production model elaboration.
The principle of operation of the milk receiving module is described by means of the automatization flowchart.
Upon arrival, milk tank vehicles get connected to the milk receiving area by hose pipes (Ilyukhin et al., 2006). Then, raw milk from milk tank vehicles feeds tank *MRT01* located at the milk receiving area. After tank *MRT01*, milk enters cooler *MRE01*, where it gets cooled down to the required temperature for its further transportation to the storage area by means of the filling line. Milk enters the manifold valve through cooler *MRE01*. The manifold, in fact, changes the route, and fills or empties the tanks located in the milk storage area.
Having described the flowchart of the milk receiving area, it is possible to have it implemented in the simulation modeling environment of AnyLogic and employing the multiagent approach.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414



**Молоко – Milk*
Мойка высокого давления – Power-wash cleaning; Электроэнергия 360 В – Electricity 360 V; Вода Д925 – Water D925
Fig. 1. Flowchart of receiving area

RESEARCH STUDY

The purpose of this research is to create a multiagent milk receiving model and milk transportation by tank vehicles in line with the production process under study. In the context of the objectives of the study, it is compulsory to use dynamically varying parameters in order to conduct virtual experiments and to identify weak points in any particular data set. Consequently, this contributes to the rational use of production resources.

AnyLogic software and built-in libraries for discrete event simulation and flow modeling are used for the purposes of simulation model elaboration within the milk receiving module.

The following constituents of the milk receiving module are available for modeling, i. e. milk tank vehicle supplier, converter of milk tank vehicle agent into flow, tank *MRT01*, cooler *MRE01*. According to the process description (Nazoykin, 2018) of creating multiagent models for production processes, the milk tank vehicle supplier and the milk receiving module are different agents independent of each other. Thus, it is compulsory to ensure a more flexible configuration, as well as providing for the ability to incorporate the obtained agents into other projects.

Agent *MilkCarProvider* (Fig. 2) acts as a milk tank vehicle supplier imitating the arrival of a certain number of milk tank vehicles of a given volume at time intervals as stated by the input simulation parameters. The agent possesses 3 varying parameters, i. e. *carCount* (number of milk tank vehicles), *carRemainingInParking* (number of milk tank vehicles at stand-by in the parking lot, and *timeBetweenArrivals* (frequency of arrivals). Item *milkCarSource* is used to generate milk tank vehicles. Among its properties, parameter *timeBetweenArrivals* is assigned to this item so that the item is aware of the frequency of milk tank vehicles generation. Likewise, parameter *carCount* is used to generate items in a given number. When the milk tank vehicle agent leaves item *milkCarSource*, parameter *carRemainingInParking* is incremented; this means that the generated milk tank vehicle enters the parking lot and is waiting for connection to the milk receiving module. The generated items line up to wait and make *carQueue*. Then, the items enter area *nextFreeMilkCar*, which is accessed by the agent of receiving module, and the agent takes milk tank

vehicles for use. When the milk tank vehicle agent leaves item *nextFreeMilkCar*, parameter *carRemainingInParking* is decremented. This means that the generated milk tank vehicle leaves the parking lot. Further-on, parameter *carRemainingInParking* is visually displayed for the agent to show the number of remaining milk tank vehicles.

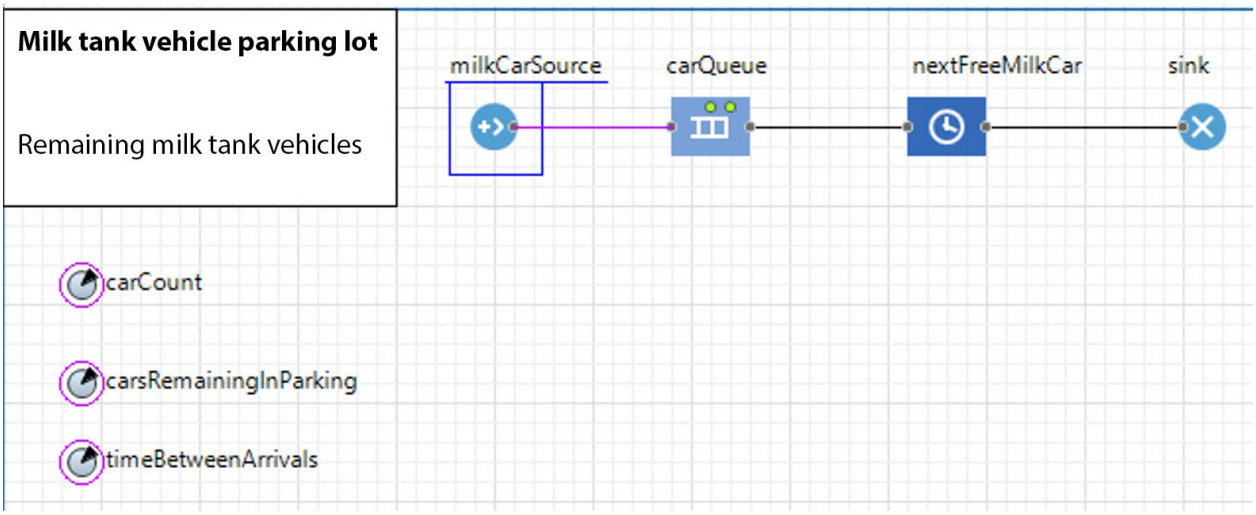


Fig. 2. Model of milk tank vehicle supplier and its visual representation

The task of *MilkReceivingStation* (milk receiving area) is to take item *milkCarSource* from agent *MilkCarProvider*, convert it into the volume of milk, select a free line and, using the selected line, fill the tank located in the milk storage area. Agent *MilkReceivingStation* possess 7 varying parameters as follows: *id* (identifier), *milkCarSize* (volume of one milk tank), *statusIndex* (current status), *statusStr* (line representation of status), *outputSpeed* (pump capacity), *tankSize* (volume of *MPT01* tank), *fillingLine* (reference to the filling line agent). Item *carSource* in this case does not generate the items of milk tank vehicle; instead, it gets filled in by calling command *inject()*. This command invokes the stage of the statechart, using which it is possible to transfer the logic of agent activities to AnyLogic software. The statechart is given in Fig. 3. Area *waitMilkCar* (waiting for a milk tank vehicle), Java code is exercised addressing to agent *MilkCarProvider* to request from item *nextFreeMilkCar* the next milk tank vehicle provided the value *carRemainingInParking* is more than 0. After the agent receives the next milk tank vehicle for unloading, the stage of the statechart called *connectingPipes* gets activated and lasts for 15 minutes simulating the connection of hose pipes to the milk tank vehicle. Item *agentToFluid* is used to convert the items of milk tank vehicle into a particular volume of liquid. This item obtains parameter *milkCarSize* to define the capacity of one milk tank vehicle. Using *Pipeline* facilities, milk is supplied first to *milkSource* (simulation of tank *MPT01*) and then to *freezingTank* (simulation of cooler *ME01*). The cooled raw milk enters area *FluidExit*. The case frames from flow simulation library *FluidExit* and *FluidEnter* are made to implement a dynamic network of flows. In this model, these case frames are needed to simulate the distribution manifold valve.

FluidExit includes method *connect*, to which item *FluidEnter* is connected. This method links *FluidExit* with *FluidEnter*. Then, the flow from *FluidExit* goes to *FluidEnter*; and this allows to dynamically convert the network of flows depending on the required configuration.

SIMULATION MODELING RESULTS

Item *Simulation* is used to do the initialization of input parameters. Command *getIntValue()* assists in choosing the integer values from the filled-up text boxes, command *getDoubleValue()* helps to choose the floating-point values. These values are applied to the previously discussed parameters in agent *Main*.

Besides, the simulation modeling environment of AnyLogic offers tools to plot different charts. In this case, a time-base chart from the available collection is selected, and the statechart for equipment is plotted. By doing so it is possible to analyse the states currently relevant for agents, and took managerial decisions. For instance, analysing the chart given in Fig. 4, it can be seen that the receiving module is in operation for only 1 hour, with the subsequent hours being idle. This means that under the set parameters the milk receiving module is able to process larger volumes of the supplied raw material.

Therefore, by carrying out the experiment with the model, it is possible to adjust the parameters of all production processes, and do the real-time tracking of the mode of operation for the equipment in order to prevent the equipment idle time, as well as identifying the system in its entirety.

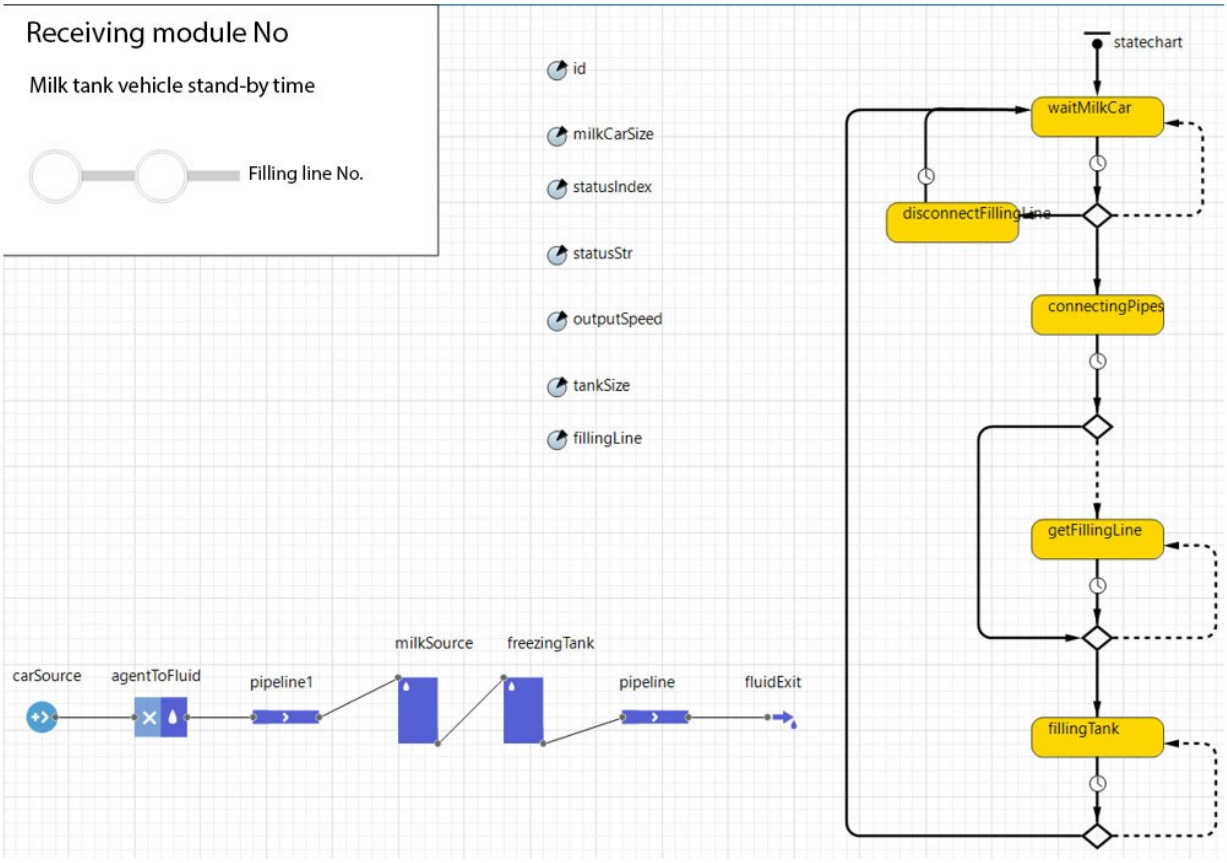


Fig. 3. Model of receiving module, its visual representation and statechart

Consequently, with the model undergoing the experiments, it is possible to select the appropriate equipment to achieve the target performance indicators.

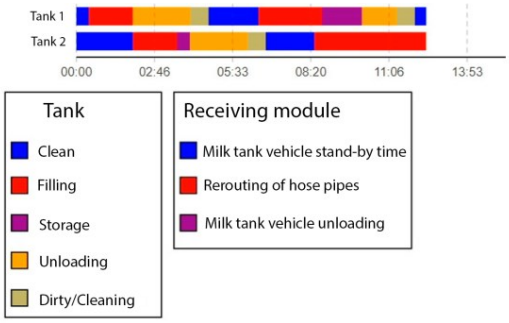


Fig. 4. Time-base chart

CONCLUSIONS

Based on the analysis of the production processes and using the equipment at the production facility, it is possible to elaborate the multiagent simulation model of the milk receiving area and the model of the milk tank vehicle supplier in the simulation modeling environment of AnyLogic.

Application of this agent-oriented model allows as follows:

- to identify the milk receiving processes at the production facility;
- to introduce the possibility of modifying the parameters for the purposes of conducting experiments;
- to employ the time-base chart to detect weak points of the production performance.

Using the designed simulation model makes it possible to test various equipment configurations by selecting relevant parameters.

REFERENCES

Sirota, A.A.. Computer modeling and evaluation of complex systems effectiveness: Coursebook / A.A. Sirota. – M.: Technosfera, 2006, – 280

B.Ya. Sovetov. Modeling of systems [Text]: University coursebook / B.Ya. Sovetov, S.A. Yakovlev. – M.: Vysshaya Shkola, 2007. – 343 p.

A. Fajar. Asynchronous agent-based simulation and optimization of parallel businesss / A. Fajar, R. Sarno // Telkomnika (Telecommunication Computing Electronics and Control) 17(4), 2019, pp. 1731-1739 (DOI: http://dx.doi.org/10.12928/telkomnika.v17i4.10846)

K.E. Gabrin. Fundamentals of simulation in economics and management: Coursebook / K.E. Gabrin, E.A. Kozlova. – Chelyabinsk: SUSU Publishing House, 2004. – 108 p.

Y.G. Karpov. Simulation of systems: introduction to modeling using AnyLogic 5 (+CD) / Y.G. Karpov. – SPb.: BHV-Petersburg, 2005. – 400 p.

M.M. Blagoveschenskaya, L.A. Zlobin. Information technology of production process control systems. M.: Vysshaya Shkola, 2010. 768 p.

V.D. Boev, D.I. Kirik, R.P. Sypchenko. Computer modeling: Manual for research and thesis project design. — SPb.: BAS, 2011. — 348 p.

E.A. Nazoykin. Multiagent models for forecasting and identifying production processes / E.A. Nazoykin, I.G. Blagoveshchensky // International Journal of Innovative Technology and Exploring Engineering - ISSN: 2278-3075, Volume-8 Issue-12, October 2019. pp. 3807-3809.

E.A. Nazoykin et al. Identification of processes for manufacturing marmalade produce using simulation methods / E.A. Nazoykin, I.G. Blagoveshchensky, M.M. Blagoveschenskaya, R.R. Naumov // Food production. 2019. No. 1 (39). pp. 40-41.

E.A. Nazoykin et al. Application of agent technologies into analysis of food production processes / E.A. Nazoykin, I.G. Blagoveshchensky, M.M. Blagoveschenskaya, R.R. Naumov // In compilation on Advanced food production technologies: development, trends, growth points. Proceedings of the Ist scientific and practical conference with international participation, November 29 - 30 2018. pp. 711-715.

V.V. Ilyukhin et al. Installation, start-up, diagnostics, repair and service of equipment for dairy production facilities / V.V. Ilyukhin, I.M. Tambovtsev, M.Y. Burlev // SPb.: GIORD, 2006. – 500 p.

V.B. Ponomarev et al. Mathematical modeling of production processes: Series of lectures / V.B. Ponomarev, A.B. Loshkarev. Yekaterinburg: GOU VPO USTU–UPI, 2006. 129 p.

P.L. Lisin et al. Contemporary process equipment for heat treatment of milk and dairy products: pasteurizing plants, heaters, coolers, starter tanks: Reference manual / P.L. Lisin, K.K. Polyansky, P.A. Miller. General ed. by Prof. K.K. Polyansky. - SPb.: GIORD, 2009. - 136 p.

AUTHOR BIOGRAPHIES

Evgeny A. Nazoykin was born in Moscow, Russia. In 2007 he entered Moscow State University of Applied Biotechnology where he studied information technology and automated systems. In 2011 he was awarded PhD. Currently Evgeny A. Nazoykin supervises the research group at Moscow State University of Food Production focusing on creating multiagent models for food production facilities. His e-mail address is nazojjkinea@mgupp.ru and ORCID profile is https://orcid.org/0000-0002-7859-1117.

DESIGNING AND OPTIMIZING PRODUCTION IN A HIGH VARIETY / LOW VOLUME ENVIRONMENT THROUGH DATA-DRIVEN SIMULATION

Virginia Fani, Bianca Bindi, Romeo Bandinelli
Department of Industrial Engineering
University of Florence
Florence, Viale Morgagni 40/44, 50134, ITALY
E-mail: virginia.fani@unifi.it

KEYWORDS
Data-driven simulation, Case study, Production, High Variety / Low Volume (HVLV).

ABSTRACT

HVLV environments are characterized by high product variety and small lot production, pushing companies to recursively design and optimize their production systems in a very short time to reach high-level performance. To increase their competitiveness, companies belonging to these industries, often SMEs working as third parties, ask for decision-making tools to support them in a quick and reactive reconfiguration of their production lines. Traditional discrete event simulation models, widely studied in the literature to solve production-related issues, do not allow real-time support to business decisions in dynamic contexts, due to the time-consuming activities needed to re-align parameters to changing environments. Data-driven approach overcomes these limitations, giving the possibility to easily update input and quickly rebuild the model itself without any changes in the modeling code. The proposed data-driven simulation model has also been interfaced with a commonly-used BI tool to support companies in the iterative comparison of different scenarios to define the optimal resource allocation for the requested production plan. The simulation model has been implemented into a SME operating in the footwear industry, showing how this approach can be used by companies to increase their performance even without a specific knowledge in building and validating simulation models.

INTRODUCTION

As suggested by the name, High Variety/Low Volume (HVLV) environments are manufacturing scenarios characterized by high product variety, frequent production order changes and small lot dimensions. As reported by White and Prybutok (2001), another possible definition of HVLV could be “non-repetitive companies”, where all the production stages operate on a non-repetitive base (Portioli-Staudacher and Tantardini, 2012). In this context, frequent changes of production mix have to be managed, often requiring the re-optimization or even re-design of production flows. HVLV represents a strategic choice for all the companies that aims to provide quick and reactive production, such as the ones working in dynamic and uncertain contexts like the fashion industry. For instance, a HVLV approach is frequently chosen by SMEs that, due to their size, have low volumes to produce and several clients to work with as third-party suppliers, facing with the trade-off between flexibility and high efficiency (Katic and Agarwal, 2018). Most of the manufacturing SMEs operates as job-shop, declared to be a HVLV manufacturing environment requiring skilled and flexible workforce to produce a wide range of products (Haider and Mirza, 2015; Huang and Irani, 2003). Each production unit produces a large variety of part types in small batches, characterized by their own routing and sequenced tasks (Slomp et al., 2009). The existing literature on HVLV is focused on the improvement of operational efficiency (Adrodegari et al., 2015; Cransberg et al., 2016; Hendry et al., 2013), even using approaches often adopted in high volume and low variety mass markets (Thomassen and Alfnes, 2017). For instance, even it is a common misunderstanding that lean is suitable for mass production only, it has been proposed to guarantee flexible productions in high variety environment (Haider and Mirza, 2015; Slomp et al., 2009). In lean paradigm, the elimination of non-value-added activities and wastes, such as overproduction and buffer, aims to reduce lead time, guaranteeing more responsiveness to customer demand (Haider and Mirza 2015). Other causes of waste are represented by long waiting and queue times that may occur due to the over-saturation of resources (Haider and Mirza 2015) or unbalanced scheduling plan (Fernandes et al, 2014; Fernandes et al, 2020), resulting in large work in process (WIP). The identification and monitoring of an appropriate set of indicators represents a key aspect especially within dynamic contexts, where changes in key performance indicators (KPIs) have to be immediately followed by the most appropriate reaction. As shown in literature (Haider and Mirza 2015; Slomp et al., 2009), main KPIs for production performance are WIP, lead time (LT), productivity, takt time (TT) and resource utilization. Despite the clear gainable benefits, KPIs monitoring and resource balancing are time-consuming activities, especially in HVLV contexts where they have to be often conducted due to the frequent change of production mix. In fact, each item has its own

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

production cycle in terms of tasks list, sequence and processing time, requiring production layout reconfiguration and re-assignment of tasks to resources (Haider and Mirza, 2015). Even if discrete-event simulation (DES) is widely used to optimize and predict the performance of job shops, frequent changes in production orders and unexpected events, typical of HVLV environments, ask for real-time models able to evaluate different scenarios in a very short time. Data-driven is an approach to simulation to overcome the long time needed to build and validate models in real environment, automatically re-building the model from data stored into structured dataset without any need to run programming code (Wang et al., 2011). According to this, they can be applied to both traditional and intelligence manufacturing systems (Zhang et al., 2019), interacting with real environments to update simulation models with on-field feedback (Goodall et al., 2019). In this paper, the data-driven approach has been used to give quick tips to final users to easily re-build the simulation model to optimize and balance resources' workload recursively. The proposed parametric data-driven model for HVLV scenarios has been applied in a footwear SME, representing the fashion industry one of the main dynamic sectors due to the high variants to be managed (d'Avolio et al., 2016), where simulation has already been successfully applied for optimizing production (Fani et al., 2017; Fani et al., 2018; Hassan et al., 2019). The work is structured as follows: in the first section, a clear overview of the purpose of the work is given; in the second section, the proposed data-driven model is described and the iterative procedure for its application summed up; the third section shows its implementation on a real scenario in the footwear industry; finally, main conclusions and further developments are shown.

PROBLEM STATEMENT

In HVLV environment, several KPIs have to be constantly monitored in operational dashboards. First, daily productivity (i.e. the number of units produced per day) represents a target value to be reached or, generally, to be maximised according to the resource availability. Frequently used within lean production systems, TT (i.e. the average time between the start of production of one unit and the next one) is a key indicator of the production rate, to be respected for matching the demand. If a process is unable to produce at takt time, in fact, additional resources or process re-engineering is needed to reach the productivity target. Besides TT, LT (i.e. the amount of time from the start of a process until its conclusion, including processing and waiting times) is another parameter to be measured and monitored to reach the productivity target. A shorter LT, in fact, results in a higher productivity. Because within most plants the largest contributor to LT was queue time (i.e. the amount of time a unit spends waiting before being processed), reducing queue time further reduces LT. The waiting time strictly depends on the queue length, a part of the work in process (WIP): queue size is the number of units waiting for being processed, while WIP is the overall number of items in a production system, including both waiting and processing items. From a lean perspective, the optimal WIP size should be equivalent to the number of workstations, having queue size equals to zero through the implementation of the one-piece-flow approach. Finally, resource utilization strongly influences WIP, because over-saturated resources represent bottlenecks in unbalanced production systems. According to this, the main key performance indicators monitored in the proposed data-driven simulation model are productivity, TT, LT, queue size and saturation. The related target values defined by companies can be reached changing variables that occur in production. For instance, additional capacity impacts on LT, reducing queue time and increasing productivity. Even the described KPIs reflect the critical success factors for companies working in several production contexts, main challenges for HVLV strategy are related to the frequent need of re-optimize or even re-design the production flows. According to this, the main challenge in HVLV strategies are not related to specific KPIs to be monitored but to identify the most suitable decision-support tool to make quick and reactive changes in production based on their value. Given certain production plan (i.e. Stock Keeping Units – SKUs - mix, delivery quantities and due dates) and production cycle per SKU (i.e. processing time per each task) as fixed input, capacity can be increased in many ways, such as enlarging the amount of working hours per day or adding more resources. Considering containers as handling units, related parameters have to be included in the analysis due to their impact on production performance. First, each containers can include a variable number of items, impacting on the processing time required per handling units and, consequently, on the queue over the system: higher container capacity is, less one-piece-flow approach is followed, increasing the WIP and slowing the overall production flow. Similarly, buffer capacity between workstations represents a variable that moves from 1, reflecting the one-piece-flow approach, to unlimited capacity, reducing the occurrence of waiting workers on the production line. Finally, restrictions on the number of containers to be daily moved over the production system can be included, especially when production is outsourced and target values are defined in supplier agreements.

MODEL DESCRIPTION

Starting from the problem statement, the proposed data-driven simulation model has been defined. The commercial simulator used is AnyLogic®, chosen for its interface with commercial databases, as well as for the easy importing procedure and its built-in database, adopted to store the input data needed to realize the data-driven model. In addition, the possibility to implement Java functions has been used to parametrize the processing times per SKU and the assignment of workers to workstations. The database structure has

been defined to make easier the import of a production plan, as well as a separate table to manage the production cycle of each SKU. For instance, in order to guarantee an easy management of changes in production mix, the database table related to production cycles has been structured including SKU, sequence, task, and task time as columns: a new SKU will only require to add rows related to its own tasks list and sequence. Moving to the parameters of each element, none of them has been included in the model as fixed value, but as a variable to be updated according to the dedicated field on the database uploaded at the model running. For instance, the assignment of each task per SKU to a workstation and a worker who processes it has been done directly on the database. Once the database structure for a parametric modeling of input and variables has been developed, the database views for collecting data to calculate the performance indicators have been realized. For instance, a datalog to track the queue size per workstation during the model running has been coded. The tracking frequency for queue size has been parametrically defined as model parameter to be easily changed before the model execution, in order to make the final user able to evaluate the trade-off between collecting more frequent information and increasing the execution speed. The model can be applied in real scenarios according to the iterative procedure shown in Figure 1.

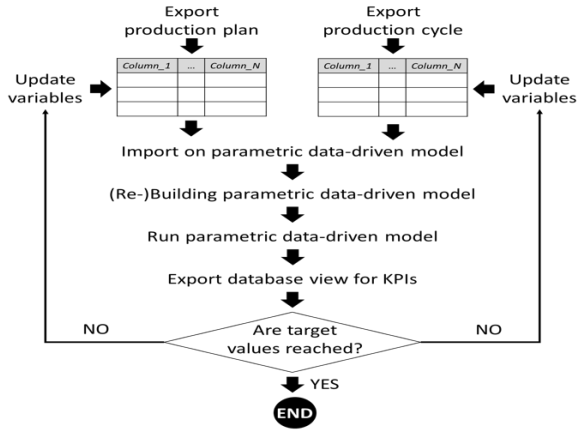


Figure 1: Proposed data-driven simulation model

Looking at Figure 1, the proposed procedure for the implementation of a parametric data-driven simulation model can be described as follows. First, the input data have to be exported from the company ERP and enriched filling values related to the variables included in the model. For instance, even the production cycle for the SKUs included in the production plan is given, the assignment of each task to a specific resource working on a certain workstation has to be done at this point. Once all the variables have been filled, the database structure for the model is ready and can be imported on the simulator database. Moreover, parameters such as containers and buffer capacities are set to be acquired by the model itself. The parametric data-driven model is then built according to the database values using the Java language available in AnyLogic®. In more detail, the generic layout of the realized discrete simulation model is composed of a parametric source and a generic "workstation" agent, as shown in the following paragraph. Moreover, the "worker" agent has been used together with dataset and schedule objects to dynamically define assignments and shifts respectively. At the model start, the assignment of workers to workstations is done using the Java language and processing time per item processed on each workstation is defined according to the value stored in the database table related to production cycles. Once the model has been run, the database views previously defined on the simulator to monitor the KPIs are exported and the values analysed. The comparison between the KPIs value coming from the simulator and the target values will determine if new iteration of the procedure is needed or not. New iterations mean changing the variables and parameters setting according to the results, in order to update the database and run again the re-built model. For instance, the productivity target could not be reached and resources will have to be re-assigned to better balance the production system, reducing queue and levelling workers' saturation.

CASE STUDY

The As-Is Scenario

The proposed simulation model has been applied into a footwear company to demonstrate its applicability in real scenarios. The footwear production cycle begins with the cutting process, followed by stiching, lasting and assembly and, finally, quality control and packing. The cutting department cuts all the parts needed for each shoe, then gathers the parts into kits (i.e. one kit includes all the parts for each pair of shoes). Cut kits then move to the stitching department for assembly. In the stitching department the operations are divided into simple steps and each worker is given few tasks, even only one. Generally, two stitching lines can support one assembly line. Once the stitching has been completed, the upper must be lasted before the outsole can be attached. Lasting is the operation that gives shoe its final shape. After the upper is heated and fitted around a plastic metal, or wood foot form called "last", the insole, midsole, and outsole are cemented to the upper. The last steps are quality control and, if shoes are compliant to the final check, their packing. Moving towards the case study, the simulation model has been applied to the stitching department of the company., composed by 4 production units organized as job shops: the first one is the preparation unit, where cut materials delivered in kit are re-organized in the stitching handling units, usually boxes, together with the other components needed (e.g. laces); the second and third units provide uppers and tongues respectively, assembled together in the last production unit. A simplified schema of the production units included in the case study is shown in Figure 2. Workstations (i.e.

"WSX" in the diagram) can be sewing machines or workbenches for manual activities, while workers (i.e. "wX" in the diagram) are resources that can be assigned to different type of tasks and machines.

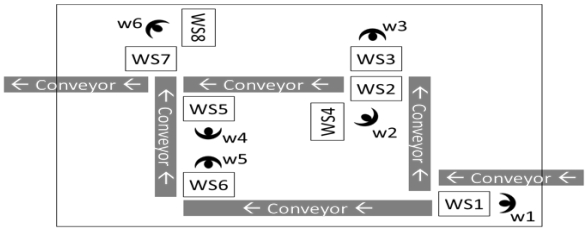


Figure 2: Resources in production units

Conveyors are used to speed the movement of handling units, but boxes can also be manually moved from one workstation to the next one according to the task sequence in the SKU production cycle. Moreover, a SKU can be worked by the same station more than once, as well as a single worker can be assigned to more than one workstation. Last, moving from one SKU type to another, different workstations can be used and different sequences can be followed, according to the SKU production cycle. For instance, considering a box filled with the generic item SKU1 entering the system shown in Figure 2, it will be processed according to the SKU production cycle, starting with tasks assigned to the worker w1 on the workstation WS1. Once w1 has completed the assigned tasks for all the SKUs included in the handled box, he will put the box back on the conveyor to move it from its workstation to the next one. Looking at the diagram, if the next task for the SKU has to be processed by w2 or w5, they can take the box directly from the conveyor; on the other hand, in case of task assigned to w4, no conveyor links the involved workstations and it is the worker himself who moves the box from the previous to his workstation. As shown in the diagram, w2 processes tasks on both WS2 and WS4, depending on the SKUs production cycle: for example, considering WS2 as sewing machine and WS4 as workbench for manual activities, SKU1 could require only sewing tasks while SKU2 also manual activities like the application of decorations or patches. Finally, non-sequential sewing activities on the same workstation could be included in the production cycle, requiring for example to process SKU2 on WS2, then on WS4 and then again on WS2. Boxes are usually mono-SKU, meaning that each box contains a certain number of the same SKU that requires the same tasks. In the case study, the capacity is equals to 2 pairs of shoes per box and 2 types of SKUs are included in the production mix.

The Application of Data-driven Simulation

Starting from the organization of the stitching department, the data-driven model for the case-study is composed by three building blocks, modelled as diagrams: one for the box and shoes generation, the second for the task processing and the third for the shoes sink and box recycling, as shown in Figure 3, Figure 4 and Figure 5 respectively. According to the resource assignments on the database, a specific number of workers, as agents, and workstations, as flowcharts, are generated at the model start.

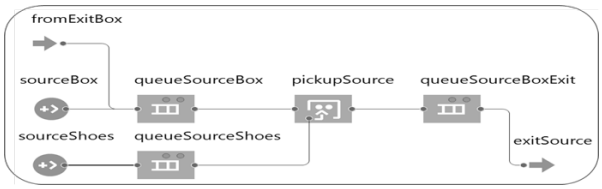


Figure 3: Source station workflow

The source diagram in Figure 3 generates both the boxes and shoes entering into the production system. According to the data stored in the database, the *sourceBox* element creates a fixed number of boxes (i.e. 300 in the case study), representing the maximum number of boxes allowed in the production system. The company has chosen to fix the number of allowed boxes to limit the WIP in the production system, but the simulation model can be run even setting an unlimited number of boxes. The *fromExitBox* element receives the empty boxes arriving from the *exitBox* station in Figure 5. The *sourceShoes* element generates shoes according to the production plan exported from the ERP and stored into the database, both in terms of pairs of shoes and scheduled date. The *pickupSource* element assigns shoes to boxes according to the box capacity: as shown in Figure 3, boxes and shoes are independent agents before the pickup element while filled boxes became the handling units after that. Filled boxes then move to the *queueSourceBoxExit* buffer, representing the company warehouse before the production area. The *exitSource* element dispatches each filled box to the right resource according to the workstation with sequence equals to "1" for the SKU contained in the box. That data is read on the database table related to the production cycle per SKU. More in details, the *exitSource* element in Figure 4 moves the filled boxes to the right *enterServiceXX* element in Figure 4.

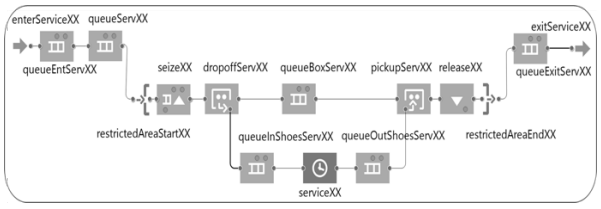


Figure 4: Generic workstation workflow

Figure 4 represents tasks processing on workstations, from the assignment of a filled box to its dispatching to the next workstation. Along the production system, boxes move from *ServiceXX* to *ServiceNN* until the final workstation listed on the SKU production cycle. Similarly to *exitSource* in Figure 3, the *exitServiceXX*

element in Figure 4 defines the criteria to move filled boxes to the next workstation, reading the sequence equals to "n+1" for the SKU contained in the box on the dedicated database table. The *restrictedAreaStartXX* and *restrictedAreaEndXX* elements are used in order to define the total number of boxes into a workstation. The size of *queueEntServXX* and *queueExitServXX* elements defines the capacity of the intermediate warehouses before and after the workstation respectively. The queuing discipline for bringing the right box to be processed from the conveyor (i.e. *queueEntServXX*) is priority-based. If a single box can be processed on the same workstation more than once, in fact, priority has to be given to boxes that have already been processed on the workstation. The *dropoffServXX* and *pickServXX* elements replicate the activities of unloading and loading of shoes done by the workers in each workstation. The *queueInShoesServXX* element defines the maximum number of shoes that can be unloaded from the boxes and release on the worker table. The *seizeXX* and *releaseXX* elements assign a specific worker to the workstation, choosing between the ones enabled from database to that workstation and according to their availability. Once the worker has been assigned, he will not be released until shoes in the *queueInShoesServXX* are completely worked, to manage workers assigned to more than one workstation. The *queueEntServXX* differs from the *queueInShoesServXX* because, while several boxes can be processed into the first buffer by different workers, once shoes have been removed from boxes and released on the worker table they will be processed by himself. To reflect the company's aim to implement a one-piece-flow strategy, the *queueEntServXX* buffer has been set to 1, while the *queueInShoesServXX* equals to the parameter related the number of shoes placed within a generic box.

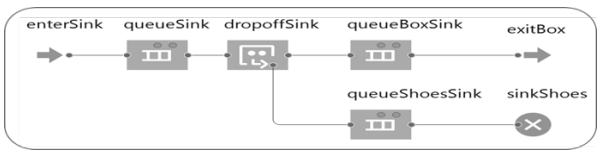


Figure 5: Sink and recycle workflow

Once the production cycle has ended, boxes enter the last building block (i.e. Figure 5), where shoes are unloaded from boxes and destroyed by the *sinkShoes* element. Boxes are moved to the *queueSourceBox* element through the *exitBox*, waiting to be filled with new shoes (i.e. Figure 3).

The Results

Starting from the described scenario, the database has been filled and imported on the simulator, according to the production plan given as input, as well as the assignment of resources to tasks hypothesised by the company that includes 36 workers and 55 workstations. A single run of 12 months with 12 replications has been carried out and the first 15 days represent the warm-up period. Microsoft PowerBI® has then been used to graphically report and navigate that results. For the analysed company, the main KPI to be monitored is productivity, with a target value to be reached of 165 pairs of shoes, mixed as 110 pairs of SKU type "1" and 55 pairs of SKU type "2". Reaching that target value has been the first objective for the company for implementing simulation, to both analyse if that productivity will be got and if possible bottlenecks could be identified in advance. The model running reached the daily target of 165 pairs of shoes, but 5 workstations showed an average queue size of more than 10 boxes, identified by the company as limit value. 2 workers operates on the critical workstations (i.e. the first worker on two workstations and the second on other three), showing each of them a saturation slightly less than 100%. According to this, the second scenario asked by the company aims to identify how many resources should be added to decrease the average queue size under the limit value. For example, new workers could be assigned to different tasks previously associated to other workers and even to different workstations. In the case study, the tasks associated to the almost saturated workers w1 and w2 have been partially re-assigned to a new resource (i.e. w37). Even if w37 did not reached a high saturation, the new configuration does not represent a suboptimal solution, because it better fits with the company need of resources able to absorb the frequent request of extra-capacity to match the high variable demand. Once the best balancing has been identified for the analysed production mix, the company asks for a quick re-building of the simulation model after changing the SKUs to be produced. In fact, the change of production mix for the analysed footwear company occurs every 4-5 weeks with a very short notice from the brand owners, requiring a re-balancing of the production line that should cover 3 days at most. The main issue the company has faced with is that, even re-balancing in advance, the first week of production for the new SKUs mix is usually spent to understand the reasons of disruptions physically detected on the production line. According to this, the expected results from using simulation have been the reduction of wrong re-balancing for changes in production mix and, consequently, a reactive re-assignment of resources to guarantee the productivity target. In the case study, instead of the two SKUs included in the first model runs, the change mix had replaced one of them with another SKU. The first run of simulation has been done re-building the data-driven model with the same number of workers and workstations, updating only database values related to SKU types and production cycles. The database views show a productivity of 157 pairs of shoes, 8 less than the target. Due to queue trends and saturation of two workers close to 100%, some of the processed tasks have been re-assigned to the under-saturated worker added in the last scenario. Once the data have been updated and the simulation model run, the productivity indicator reaches the target value. Finally, the last scenario analysed by the company

refers to how to readapt the production line to double the productivity. The approach followed has been, first, doubling the number of workers assigned to each task, processing each one the 50% of the production. According to the iterative procedure showed in Figure 1, once the data-driven model had been run with the updated production plan and resource assignment as input, the final user has analysed the results in terms of production performance. As expected, the productivity target has been reached doubling the involved resources. Many improvements could be introduced to optimize the resources balancing, due to the high undersaturation of many workers. The iterative tasks re-assignment and KPIs evaluation procedure has been conducted allocating tasks splitted to several resources to few ones, until the optimal configuration of resources to guarantee the productivity target has been identified. The application of that procedure reduces the number of workers. from 74 to 53.

CONCLUSION

The present work demonstrate the successful application of the proposed data-driven simulation model to a HVLV real context. The case study has demonstrated how an iterative approach to data-driven simulation can support companies in decision-making process towards production performance improvement. More in details, this work demonstrates how the limitations of traditional simulation modelling into a dynamic environment can be overcome, reducing the time needed to find the optimal solution in terms of association workstation-tasks, number of workstations and number of workers. Beside this, further developments can be identified starting from the main results listed above. On the one hand, a data-driven approach to the 2D and 3D modelling can be included in the proposed model to assess other KPIs, such as layout optimization to minimize workers' movement along the production line. On the other hand, manual updates on database parameters after each iteration represent a key value for reaching simulation benefits by low-tech users but could not fit more structured companies. According to this, the proposed data-driven model can be extended towards new trends, such as running optimization algorithms within the simulation model or introducing Industry 4.0 technologies like Artificial Intelligence to improve performance and reduce computational time.

REFERENCES

Adrodegari, F., Bacchetti A., Pinto R., Pirola F., Zanardini M. 2015. "Engineer-to-order (ETO) production planning and control: An empirical framework for machinery-building companies". *Production Planning and Control* 26(11), 910–932.

Cransberg, V., Land M., Hicks C., Stevenson M. 2016. "Handling the complexities of real-life job shops when implementing workload control: A decision framework and case study". *International Journal of Production Research* 54(4), 1094–1109.

d'Avolio, E., Bandinelli, R., Rinaldi, R. 2016. "How product development can be improved in fast fashion industry: An Italian case". *IFIP Advances in Information and Communication Technology* 467, 718–728.

Fani, V., Bandinelli, R., Rinaldi, R.. 2017. "Optimizing production allocation with simulation in the fashion industry: A multi-company case study". In *Proceedings of the 2017 Winter Simulation Conference (Las Vegas, Nevada)* 3917–3927.

Fani, V., Bindi, B., Bandinelli, R., Rinaldi, R.. 2018. "Optimizing production performances with simulation: A case study in the fashion industry". In *Proceedings of the XXIII Summer School Francesco Turco (Palermo, Italy)*. 200–205.

Fernandes, N.O., Land, M.J., Carmo-Silva, S. 2014. "Workload control in unbalanced job shops". *International Journal of Production Research* 52(3), 679–690.

Fernandes, N.O., Thürer, M., Pinho, T.M., Torres, P., Carmo-Silva, S. 2020. "Workload control and optimised order release: an assessment by simulation". *International Journal of Production Research* 58(10), 3180-3193.

Goodall, P., Sharpe, R., West, A. 2019. "A data-driven simulation to support remanufacturing operations". *Computers in Industry* 105, 48–60.

Haider, A. and Mirza, J. 2015. "An implementation of lean scheduling in a job shop environment". *Advances in Production Engineering & Management* 10(1), 5–17.

Hassan, M.M., Kalamraju, S.P., Dangeti, S., Pudipeddi, S., and Williams, E.J. 2019. "Simulation Improves Efficiency and Quality in Shoe Manufacturing". In *Proceedings of the 31st European Modelling and Simulation Symposium* 231–236.

Hendry, L., Huang, Y. and Stevenson, M. 2013. "Workload control: Successful implementation taking a contingency-based view of production planning and control". *International Journal of Operations and Production Management* 33(1), 69–103.

Huang, H. and Irani, S. 2003. "An enhanced systematic layout planning process for high-variety low-volume (HVLV) manufacturing facilities". In *Proceeding of the 17th International Conference on Production Research*.

Katic, M. and Agarwal, R. 2018. "The Flexibility Paradox: Achieving Ambidexterity in High-Variety, Low-Volume Manufacturing". *Global Journal of Flexible Systems Management* 19(s1), 69–86.

Portioli-Staudacher, A. and Tantardini, M. 2012. "Lean implementation in non-repetitive companies: a survey and analysis". *International Journal of Services and Operations Management* 11(4), 385–406.

Slomp, J., Bokhorst, J.A.C., and Germs, R. 2009. "A lean production control system for high-variety/low-volume environments: a case study implementation". *Production Planning and Control* 20(7), 586–595.

Thomassen, M. K. and Alfnes, E. 2017. "Managing Complexity". In *Proceedings of the 8th World Conference on Mass Customization, Personalization, and Co-Creation (Montreal, Canada)*.

Wang, J., Chang, Q., Xiao G., Wang N., Li S. 2011. "Data driven production modeling and simulation of complex automobile general assembly plant". *Computers in Industry* 62(7), 765–775.

White, R. E. and Prybutok, V. 2001. "The relationship between JIT practices and type of production system". *Omega* 29(2), 113–124.

Zhang, L., Zhou, L., Ren, L., Laili, Y. 2019. "Modeling and simulation in intelligent manufacturing". *Computers in Industry* 112, 103123.

Evaluation of Algorithm Performance for Simulated Square and Non-Square Logistic Assignment Problems

Maximilian Selmair
Sascha Hamzehi

BMW Group
Email: maximilian.selmair@bmw.de

Klaus-Jürgen Meier

University of Applied Sciences Munich
Email: klaus-juergen.meier@hm.edu

***Abstract*—The optimal allocation of transportation tasks to a fleet of vehicles, especially for large-scale systems of more than 20 Autonomous Mobile Robots (AMRs), remains a major challenge in logistics. Optimal in this context refers to two criteria: how close a result is to the best achievable objective value and the shortest possible computational time. Operations research has provided different methods that can be applied to solve this assignment problem. Our literature review has revealed six commonly applied methods to solve this problem. In this paper, we compared three optimal methods (Integer Linear Programming, Hungarian Method and the Jonker Volgenant Castanon algorithm) to three three heuristic methods (Greedy Search algorithm, Vogel's Approximation Method and Vogel's Approximation Method for non-quadratic Matrices). The latter group generally yield results faster, but were not developed to provide optimal results in terms of the optimal objective value. Every method was applied to 20.000 randomised samples of matrices, which differed in scale and configuration, in simulation experiments in order to determine the results' proximity to the optimal solution as well as their computational time. The simulation results demonstrate that all methods vary in their time needed to solve the assignment problem scenarios as well as in the respective quality of the solution. Based on these results yielded by computing quadratic and non-quadratic matrices of different scales, we have concluded that the Jonker Volgenant Castanon algorithm is deemed to be the best method for solving quadratic and non-quadratic assignment problems with optimal precision. However, if performance in terms of computational time is prioritised for large non-quadratic matrices (50×300 and larger), the Vogel's Approximation Method for non-quadratic Matrices generates faster approximated solutions.**

***Keywords*—Assignment Problem; Jonker Volgenant Castanon; Vogel's Approximation Method; Hungarian Method; Greedy Search; Autonomous Mobile Robots;**

I. Introduction

The transportation of goods and materials, which can take place within a plant as well as to and from it, is an integral part of every supply chain and contributes substantially to an enterprise's expenses. As such, every logistics manager who has to deal with matters of transport aims to find the most cost-efficient solution while meeting high customer service standards, i. e. meeting demands at the lowest possible expense. This is referred to as the Assignment Problem (AP) and has been the subject of extensive operational research (Díaz-Parra et al., 2014). Researchers, who attempt to solve this problem, also aim to minimise the total transportation costs whilst moving goods from several supply points (e. g. warehouses) to demand locations (e. g. customers). Theoretical and practical use cases are based on two conditions. Each transport origin features a fixed amount of goods that can be distributed. Correspondingly, every point of transport destination requires a certain number of goods (Shore, 1970).

Each point of origin can be connected to every designated destination, thereby creating several routes that are often associated with different costs. These varying costs can be visualised in matrices, where each cell represents a potential effort. Depending on the objective of the optimisation, these efforts can be a monetary value, the transport distance or the duration that an agent requires to fulfil a corresponding task. Matrices like this are common practice in studies dealing with the AP and are used to make allocation decisions.

The underlying use case, where tasks have to be assigned to vehicles of a fleet consisting of AMRs, differs from the classical AP. In our case, each AMR has a capacity restriction of 1, i. e. a maximum of one load carrier can be transported at a time. Furthermore, each task corresponds to a demand of 1. This means that every task can only be allocated to one single AMR. Additionally, the number of available AMRs rarely matches the number of to-be-assigned tasks in practice. Since the size of the matrices depends on this factor, non-quadratic matrices (e. g. 40×50) are common. There are different approaches to solve this particular assignment problem, e. g. Integer Linear Programming, machine-learning-based methods or the application of algorithms (see section II).

For the latter approach, algorithms, one needs to distinguish between those that solve every problem scenario optimally, here referred to as optimal methods, and heuristics, which in some cases fail to find the best solution and instead result in approximations. Furthermore, all methods vary greatly in terms of the computation time required to solve a given assignment problem.

This paper is structured as follows: the next section provides an overview of the different methods that can be applied to solving the AP. These being Integer Linear Programming (ILP), the Hungarian Method (HM), the Jonker Volgenant Castanon (JVC) algorithm, a variation of a Greedy Search (GS) algorithm and Vogel's Approximation Method (VAM) with its associated adapted version for non-quadratic matrices, VAM-nq. The third

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

section introduces the use case that gave rise to this study. The simulation study and the discussion of the results are presented in section IV, which is followed by the conclusion.

II. The History of Solving the Assignment Problem

The Generalised Assignment Problem (GAP) describes the generalisation of both the knapsack problem and the bin packing problem (Cattrysse and van Wassenhove, 1992). In its original characteristic, one is given m agents with a corresponding capacity B_j for each agent j, and n tasks such that each task i has size s_{ij} and yields profit p_{ij} when assigned to agent j. The objective is to find an optimal assignment of agents to tasks by maximising the total profit. Applications of GAP can be found in multiple areas including fixed charge location problems, grouping and loading for vehicle routing, flexible manufacturing systems, scheduling variable length commercials, allocating storage space, designing communication networks, scheduling payments on accounts, assigning software development tasks to programmers, scheduling projects, assigning jobs to computers in networks, and assigning ships to overhaul facilities (Krumke and Thielen, 2013; Cattrysse and van Wassenhove, 1992).

In this paper, we consider a variation of the problem in which all the agents' budgets and all tasks' costs are equal to 1 (Pentico, 2007). This section provides a description of the established solution methods for the AP. The first attempt at solving the AP dates back to 1955 with Kuhn's Hungarian Method. In the subsequent decades, numerous researchers proposed a multitude of methods and solutions as listed hereafter in their chronological order: (Kuhn, 1955; Munkres, 1957; Reinfeld and Vogel, 1958; Witzgall and Zahn, 1965; Edmonds, Johnson and Lockhart, 1969; Dinic and Kronrod, 1969; Hopcroft and Karp, 1973; Pulleyblank, 1973; Gabow, 1976; Lawler, 1976; Karzanov, 1976; Cunningham and Marsh, 1978; Micali and Vazirani, 1980; Burkard and Derigs, 1980; Kazakidis, 1980; Derigs, 1981; Shimshak, Kaslik and Barclay, 1981; Galil, Micali and Gabow, 1982; Minoux, 1982; Havel, Kuntz and Crippen, 1983; Gabow, 1985; Grötschel and Holland, 1985; Derigs, 1986; Derigs and Metz, 1986; Trick, 1987; Jonker and Volgenant, 1987; Derigs, 1988; Lessard, Rousseau and Minoux, 1989; Gabow, Galil and Spencer, 1989; Gabow, 1990; Gabow and Tarjan, 1991; Derigs and Metz, 1991; Gerngross, 1991; Applegate and Cook, 1993; Atamtürk, 1993; Miller and Pekny, 1995; Feder and Motwani, 1995; Cheriyan, Hagerup and Mehlhorn, 1996; Goldberg and Kennedy, 1997; Goldberg and Karzanov, 2004; Mucha and Sankowski, 2004; Harvey, 2006; Duan and Pettie, 2014; Cygan, Gabow and Sankowski, 2015; Gabow, 2017; Selmair, Meier and Wang, 2019)

The analysis of the relevant literature has yielded a total of four methods that are frequently applied in research (Paul, 2018; Dinagar and Keerthivasan, 2018; Nahar, Rusyaman and Putri, 2018; Ahmed et al., 2016; Korukoğlu and Ballı, 2011; Freling, Wagelmans and Paixão, 2001; Li, Mirchandani and Borenstein, 2007; Balakrishnan, 1990). These are Integer Linear Programming (ILP), the Hungarian Method (HM), Vogel's Approximation Method (VAM) and the Jonker Volgenant Castanon (JVC) algorithm. For the purpose of this paper, we chose the three optimal methods – ILP, HM and the JVC algorithm – and three heuristic approximation methods: Vogel's Approximation Method with its variation for non-quadratic matrices, VAM-nq, as well as a trivial Greedy Search algorithm for benchmark reasons. In the following, each method will be described in detail.

A. Integer Linear Programming

As stated previously, ILP is one if the approaches that is able to ascertain an optimal solution for different, even large-scale scenarios. Its application begins by formulating an objective function as well as applicable restrictions in order to receive correct results. In accordance with Osman (1995) and by adapting ILP to the use case at hand, the objective function reads as follows:

$$\min \sum_{t \in T} \sum_{a \in A} d_{ta} \cdot c_{ta} \tag{1}$$

$$\sum_{a \in A} d_{ta} = 1 \qquad \forall t \in T \tag{2}$$

$$\sum_{t \in T} d_{ta} \leq 1 \qquad \forall a \in A \tag{3}$$

$$d_{ta} \in \{0, 1\} \qquad \forall t \in T, \forall a \in A \tag{4}$$

The goal of the objective function 1 is to minimise the sum of all costs c_{ja} for all tasks $T = 1, \ldots, m$ and for all agents $A = 1, \ldots, n$, which is the result of multiplying the decision variable d_{ja} by the corresponding costs that arise when a task t is assigned to an agent a. For the underlying use case, an agent refers to an AMR. The first Constraint (Equation 2) ensures that every task is actually assigned to an agent whereas the second Constraint (Equation 3) makes sure that each agent's capacity of 1 is not exceeded, i. e. each agent can only perform one task at a time. The last Constraint (Equation 4), applies to both tasks and agents and restricts the decision variable d_{ja} to binary values.

B. Hungarian Method

The Hungarian Method was initially proposed by Kuhn (1955) to solve the AP. Similar to ILP and JVC, the HM is able to find an optimal solution to any given problem. The algorithm solves $n \times n$ matrices (e. g. 10×10) by carrying out the following steps until an optimum solution is found:

1) Identify the minimum value in each column and subtract this value from all other values in the corresponding column.
2) Identify the minimum value in each row and subtract this value from all other values in the corresponding row.
3) All zeros in the matrix must be covered by marking as few rows and/or columns as possible.
4) Check if the number of lines equals n. If it does, an optimal allocation of the zero-values is possible. If the number of lines is smaller than n, an optimal allocation is not yet feasible and Step 5 has to be carried out.

5) Find the smallest value which is not covered by a line and a) subtract this value from each uncovered row and b) add it to each covered column.
6) Return to Step 3.

It has to be noted that non-quadratic matrices, $n \times m$ matrices (e. g. 10×40), are converted to $n \times n$ matrices (e. g. 40×40), as the Hungarian Method can only be applied to quadratic matrices. The additional value cells are allocated with the highest value identified in the original matrix. This adaptation requires additional computing power since the algorithm has to consider, for instance, 1.600 cells of values (40×40) instead of 400 (10×40). It is evident that this is a drawback when non-quadratic matrices are to be solved, yet this is always the case when more tasks than agents have to be considered or vice versa.

C. Jonker Volgenant Castanon Algorithm

Skiena (1990) defined the augmenting path as a simple path – thus, a path that does not contain cycles – through a network using only edges with positive capacity from the source to the sink. Compared to the Hungarian Method, which finds any augmenting paths out of all feasible ones, the JVC algorithm finds the shortest augmenting path among all options. Although the JVC algorithm is based on the implementation of the Hungarian Method, the following additional pre-processing steps are required:

1) *initialization step* with sub-steps such as a) column reduction, b) reduction transfer, c) reduction of unassigned rows, similar to auction algorithms,
2) *termination step*, if row assignment is complete,
3) *augmentation step*, where an auxiliary network is generated to determine an alternating path with minimal total objective cost, from unassigned row i to unassigned column j,
4) *update step of dual solution*, where the complementary slackness variables are updated, and finally a
5) *return step to the termination Step 2.*

where the algorithm output is an integer sequence that assigns each column to each corresponding minimal row element denoted by $v_j = \min_i(c_{ij})$. Note, that some rows may not be assigned, which is particularly important for the processing of non-quadratic matrices. Moreover, the JVC algorithm minimises the primary objective:

$$\min \sum_i \sum_j c_{ij} x_{ij}; \tag{5}$$

subject to the constraints:

$$\sum_i x_{ij} = 1, \ \sum_j x_{ij} = 1 \tag{6}$$

$$x_{ij} \leq 0, \ \forall (i,j) \in E \ , \tag{7}$$

and maximises the dual objective as follows:

$$\max\Big\{\sum_i u_i + \sum_j v_j\Big\} \text{ s.t. } c_{ij} - u_i - v_j \leq 0 \ . \tag{8}$$

where the reduction transfer is completed from unassigned to assigned rows where for each row i.

$$j_1 = x_i; \tag{9}$$

$$\mu = \min\{c_{ij} - v_j : j = 1, ..., n; \forall j \neq j - 1\} \tag{10}$$

$$v_{j1} = v_{j1} - (\mu - u_i) \, ; u_i = \mu \tag{11}$$

After augmentation Step 3, the partial solution assignments are updated, while the dual values are updated to restore the complementary slackness conditions in the following:

$$c_{ik} - u_i - v_k = 0, \text{if } x_i = k \ \forall \text{ assigned columns and} \tag{12}$$

$$k, i = 1, .., n \text{ and} \tag{13}$$

$$c_{ik} - u_i - v_k \leq 0 \tag{14}$$

Generally, the complexity of the first two initialisation procedures is reported to be $\mathcal{O}(n^2)$, whereas it has been shown that the augmenting reduction procedure has a complexity of $\mathcal{O}(R \cdot n^2)$. Here, R refers to the range of the cost coefficients (Jonker and Volgenant, 1987).

D. Greedy Search Heuristic

In order to compare the optimal method algorithms to a simple heuristic, we have used a Greedy Search method, which is described as follows:

1) Input the cost matrix C and determine the maximum cost value m_c,
2) while a minimum value can be found that is lower than m_c, locate the minimum value across all rows i and columns j,
3) update the objective value o_c^k, of current iteration k by adding its own value to the previous iterative objective value o_c^{k-1}.
4) for all columns j, locate the minimum value row and column position (i_{min}, j_{min}), assign the task to the assignment or matching list, and set the minimum value to maximum cost m_c for all residual row entries,
5) for all rows i, locate the minimum value, assign the task, and set the minimum value to the maximum cost m_c for all residual column entries,
6) until no minimum values are left that are smaller than the maximum value m_c.

E. Vogel's Approximation Method and VAM-nq

The following description of Vogel's Approximation Method is based on the original proposal by Reinfeld and Vogel (1958). VAM solves transport matrices by repeating the steps presented below until a feasible solution is found. The cells of the matrices are assigned with the costs c_{ij} associated with allocating a task to an agent. These costs occur when an agent transports goods from a point of origin i to a destination j. Depending on the objective of the optimisation, c_{ij} can be a monetary value, the transport distance or the duration that an agent requires to fulfil the corresponding task. Each source (origin) features a specific amount of goods that can be allocated (supply). Correspondingly, each sink (destination) usually requires a certain number of units (demand). The underlying case is special in regard to supply and demand. Each agent can only

provide a supply of 1. Correspondingly, each task equals a demand of 1. As mentioned earlier, this fact describes the special case of the Generalised Assignment Problem, refereed to the Assignment Problem. In order to carry out the allocation under these circumstances, the following steps are necessary:

1) Calculate the difference between the smallest and the second-smallest cell value for each row and each column.
2) Select the row or column which features the biggest difference. If the calculations yield the same value in two or more columns or rows, select the row or column containing the smallest cell value.
3) Choose the smallest cell value of the selected row or column and allocate the corresponding task to an agent.
4) Eliminate the row and column that has been used for the allocation.
5) Check if there are still agents and tasks left to allocate. If so, repeat Steps 1 - 4.

Different authors have previously attempted to improve the classic VAM in order to progress towards an optimal solution, which is generally achieved by applying either ILP or the HM. These are, for instance, (Paul, 2018; Dinagar and Keerthivasan, 2018; Nahar, Rusyaman and Putri, 2018; Ahmed et al., 2016; Korukoğlu and Ballı, 2011; Balakrishnan, 1990; Goyal, 1984; Shimshak, Kaslik and Barclay, 1981).

An example of the VAM can be found in Tables I a through c. Here, the agents V_i are assigned to the different rows and the tasks T_j to the columns. The individual cells show the costs c_{ij} for each possible task-agent combination. The row differences, which refer to those between the least and the second-least expensive option, can be found in the column Δi, whereas the column differences are shown in Δj. Table I a shows that the most substantial difference is associated with the third row, which features the lowest value in the third column (Table I b). Accordingly, Task 3 is assigned to Agent 3. After the allocation, the third row and column are eliminated (Table I c).

Selmair, Meier and Wang (2019) have shown that the original VAM is not always suitable when it comes to determining optimal or near-optimal results for non-quadratic matrices (see Figure 5 a). In fact, the calculated objective values are in some cases more than 100 % higher than the optimal objective value. For the purpose of comprehensibility, any matrix is described as having two dimensions, yet these are not identical in length for non-quadratic matrices. In line with this, it was proposed that these insufficient results arise if either the selected row or column, which features the maximum difference, is from the shorter dimension. This may mean that if a matrix contains more columns than rows, choosing a column with the maximum difference (which is achieved by subtracting cell values in the smaller dimension / rows) might result in values that deviate from the optimal objective value. The same also applies if there are more rows than columns. This can be explained by the fact that the dimension of rows features more values and the chance is therefore higher to find a smaller cell value within those. In order to mitigate the above stated disadvantage of VAM, an improved version of VAM was developed.

Figure 5 a shows that the results of the VAM start to deteriorate as soon as the matrix size is increased in only one dimension, i. e. a non-quadratic matrix is created. While VAM is able to generate optimal solutions in some cases, however, in those instances in which it fails, the calculated objective values are up to 5 times higher than the optimal objective value. The deviations from the optimal solution increase continuously as the difference between the number of rows and columns increases. This even applies to small non-quadratic instances with the dimensions of 4×5. For instance, in the case of 5×10 matrices, the calculated objective values can be three times as high as the actual optimal objective value.

Selmair, Meier and Wang (2019) provided an enhanced version of VAM, namely Vogel's Approximation Method for non-quadratic Matrices, which yields results that are much closer to the optimal objective value for said non-quadratic matrices. The description below is based on a scenario in which a matrix contains more columns than rows. If a matrix were to contain more rows than columns, these steps can be adapted accordingly by replacing "rows" with "columns" and vice versa. The Vogel's Approximation Method for non-quadratic Matrices (VAM-nq) solves allocation matrices featuring more columns than rows by carrying out the following steps:

1) Calculate the difference between the smallest and the second-smallest cell value for each row.
2) Select the row featuring the biggest difference. If there is a tie between rows, choose the row containing the smallest cell value.
3) Determine the smallest cell value for the selected row and allocate the corresponding task to an agent.
4) Eliminate the corresponding row and column that have been used for the allocation.
5) Check if there are still agents and tasks left to allocate, and repeat Steps 1 - 4 in case that there are.

On comparison of the original VAM and the enhanced VAM-nq, some aspects point towards a refinement of the original approach. For one, VAM-nq considers only the rows if there are more columns than rows (Step 1). Accordingly, only the largest differences within each of the rows and the corresponding smallest cell values are considered (Step 2 and 3). Applying this method to the other option, that is, to matrices that feature more rows than columns, results in the following approach. Steps 1 through 3 only apply to columns, their biggest differences within each column and smallest cell values. With Table II, the example of subsection II-E is solved by means of all three heuristics, illustrating that the proposed VAM-nq provides substantially better results even in small non-quadratic cases.

III. The Use Case at Hand

The use case presented in this study is part of a BMW project in the context of Industry 4.0, to which we are contributing. In order to automate its internal material flow, the BMW Group has developed its own AMR, the so-called Smart Transport Robot (STR) shown in Figure 1, which is designed to substitute

c_{ij}	T1	T2	T3	T4	Δi
V1	200	100	400	50	50
V2	60	80	30	350	30
V3	210	300	70	150	80
V4	120	510	340	80	40
V5	70	80	40	400	30
Δj	10	0	10	30	

(a) Initial Matrix to be solved by VAM

c_{ij}	T1	T2	T3	T4	Δi
V1	200	100	400	50	50
V2	60	80	30	350	30
V3	210	300	70	150	80
V4	120	510	340	80	40
V5	70	80	40	400	30
Δj	10	0	10	30	

(b) Matrix featuring the identified biggest difference (80)

c_{ij}	T1	T2	T3	T4	Δi
V1	200	100	400	50	50
V2	60	80	30	350	30
V3	210	300	70	150	80
V4	120	510	340	80	40
V5	70	80	40	400	10
Δj	10	0	10	30	

(c) Matrix after eliminating assigned row and column

Table I: Exemplary procedure of Vogel's Approximation Method

c_{ij}	T1	T2	T3	T4
V1	200	100	400	**50**
V2	60	80	**30**	350
V3	210	**300**	70	150
V4	120	510	340	80
V5	**70**	80	40	400

(a) Solution of Greedy Search with an objective value of 450

c_{ij}	T1	T2	T3	T4
V1	200	100	400	**50**
V2	60	80	30	350
V3	210	300	**70**	150
V4	**120**	510	340	80
V5	70	**80**	40	400

(b) Solution of the original VAM with an objective value of 320

c_{ij}	T1	T2	T3	T4
V1	200	100	400	**50**
V2	60	**80**	30	350
V3	210	300	70	150
V4	**120**	510	340	80
V5	70	80	**40**	400

(c) Solution of VAM-nq with an objective value of 290

Table II: Exemplary solutions of GS, VAM and VAM-nq with different resulting objective values

commonly used tucker trains. These play a central role in the automotive material handling processes. The industrial use-case of the BMW Group specifies the following requirements:

1) Every agent can carry one load carrier at a time.
2) All transportation requests are unknown prior to their receipt.
3) Agents without a task are required to either recharge or park.

The allocation of tasks to AMR is formulated as the AP. In this context, the costs are defined as the driving effort of each AMR to the source-point of a task. This study aims to identify the most efficient method to solve this assignment problem. Although it is highly likely that this specific problem results in non-quadratic matrices, we have also applied all methods to quadratic matrices.

Figure 1: BMW's STR in its natural habitat.

IV. Simulation Study

In order to evaluate the suitability of all previously described methods, simulations were performed to provide the necessary data. This section describes the methodology of the simulation study and presents its results.

A. Algorithm Performance Metric for Comparisons

We utilised the Mean Absolute Percentage Error (MAPE) metric, defined below, in order to measure the qualitative performances of the introduced methods.

$$E_{\text{MAPE}} = \frac{1}{n}\sum_{t=1}^{n}\frac{S_s - O_s}{O_s} \tag{15}$$

It indicates the percentage average and the relative deviation from an actual optimal objective value. Furthermore, in Equation 15, O_s denotes the optimal objective value of an optimal solution, whereas S_s is the objective value of an inferior solution. The optimal value therefore is based on the best solution achievable, whereas any deviation is considered to be sub-optimal.

B. Research Design

Simulation experiments based on different quadratic and non-quadratic matrices were carried out using an Intel Core i7-6820HQ 2.70 GHz featuring 32 GB RAM in order to compare the six methods discussed in this paper. The software *AnyLogic*

version 8.7.2 was utilised to generate and solve assignment problems.

For every experiment, a total of 20.000 assignment scenarios were randomly generated. In order to generate realistic sets of data, a map was randomly filled with nodes, which represent the AMRs and the tasks. Their distance from each other provided the cost values assigned to the cells in the matrices. The following overview shows which matrices were used to generate and evaluate the corresponding Key Performance Indicators (KPIs):

- Mean Computation Times for ILP, HM, JVC, GS, VAM and VAM-nq were computed for:
  a) 20 quadratic matrices starting at 10×10 and increasing to 200×200 in steps of 10
  b) 20 non-quadratic matrices starting at 50×50 and increasing to 50×525 in steps of 25
- Probability for the heuristic methods VAM, VAM-nq and Greedy Search to reach the actual optimal objective value was calculated for:
  a) 50 quadratic $n \times n$ matrices with $n = \{5, \ldots, 35\}$
  b) 50 non-quadratic $5 \times n$ matrices with $n = \{6, \ldots, 55\}$
- E_{MAPE}, Mean Absolute Percentage Error of VAM, VAM-nq and Greedy Search to map the deviation from the actual optimal solution was computed for 15 non-quadratic $5 \times n$ matrices with $n = \{6, \ldots, 20\}$
- E_{MAPE}, Mean Absolute Percentage Error of VAM, VAM-nq and Greedy Search to map the deviation from the optimal objective value were calculated for:
  a) 50 non-quadratic $50 \times n$ matrices with $n = \{51, \ldots, 100\}$
  b) 17 mixed matrices: 5×5, 5×50, 10×10, 10×20, 10×30, 10×40, 20×20, 10×60, 20×60, 30×30, 10×100, 40×40, 50×50, 50×100, 100×100, 100×200, 100×300

C. Discussion of Results

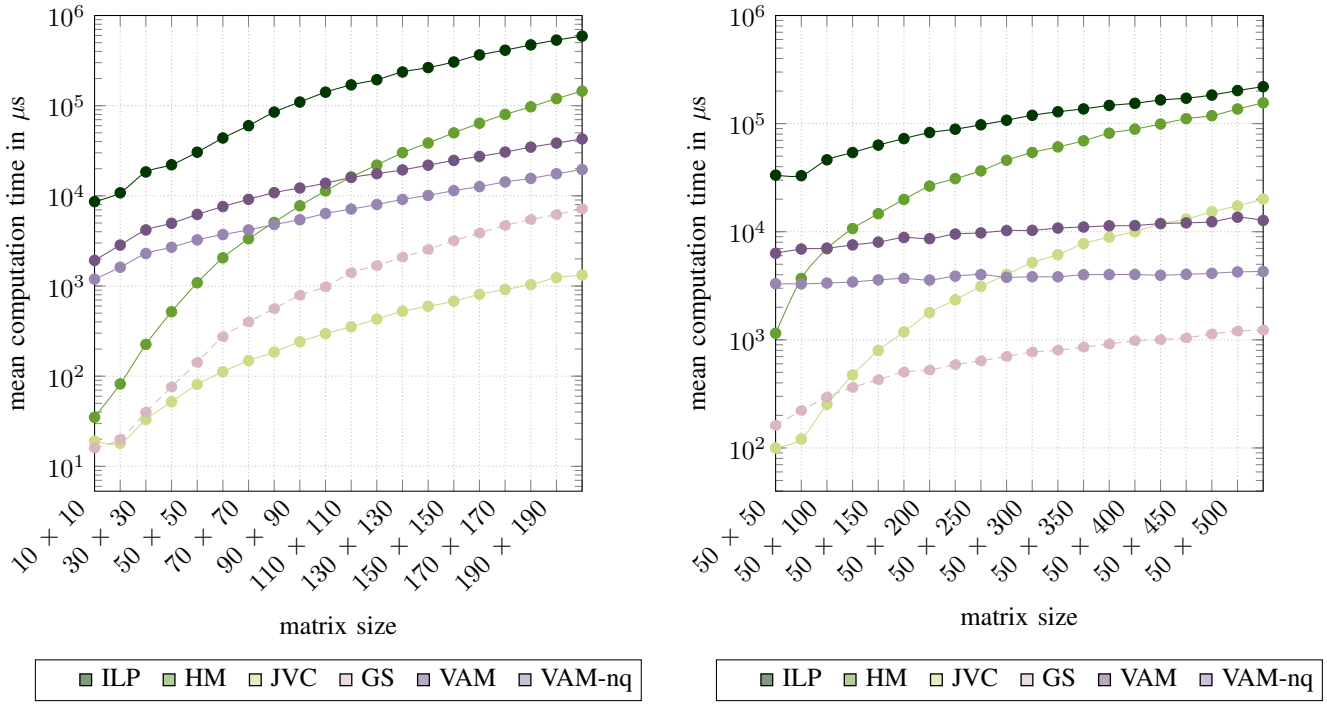
This section summarises the results of the simulation experiments. Figure 2 a shows clearly that ILP requires the most computation time for quadratic matrices of all sizes. For example, the time required to solve a 180×180 matrix is six times longer than that required by the Hungarian Method and even 1.000 times longer than for the JVC algorithm. JVC is by far the fastest method to solve quadratic matrices of all sizes. The results show JVC to be even faster than the Greedy Search heuristic even for large quadratic matrix sizes. For non-quadratic matrices, VAM-nq begins to outperform the JVC algorithm when matrices are equal to or larger than 50×300 as illustrated in Figure 2 b. More specifically, the computational time only increases slightly for VAM-nq as the matrix size increases even further, while also noting that the VAM-nq provides optimal objective values in most cases. In contrast, the computational time required by JVC, rises exponentially as the matrix size increases as presented in Figure 4.

HM, VAM and VAM-nq yielded computation times between 31 μs and 100.000 μs to solve matrices of all considered sizes. Even so, the results showed that the time required to compute results for matrices equal to or larger than 120×120 and 50×125, for non-quadratic matrices, was longer for the HM than both the VAM and VAM-nq. However, Figure 2 b shows that the computation time for the HM increases considerably when non-quadratic matrices are computed. This is most likely due to the fact that the HM has to generate additional rows or columns to produce quadratic matrices since it cannot, as previously mentioned, compute non-quadratic problems. VAM and VAM-nq, on the other hand, are able to compute quadratic and non-quadratic matrices regardless of their size in a relatively short amount of time (between 1.000 μs for small sizes and 30.000 μs for larger ones), which provides support for the superior scalability of VAM and VAM-nq.

As previously mentioned, the three heuristic methods GS, VAM and VAM-nq are not always able to produce an optimal solution. Figure 3 discusses the probability of these three methods to reach for the actual optimal objective value. For quadratic matrices, this probability approaches 0 % for matrix sizes of 20×20 and upwards. In contrast, for relatively small quadratic scenarios, such as 9×9, the chance for VAM and VAM-nq to calculate the best possible solution is only 24.7 %. For a 16×16 scenario, the chance is only 3.3 %. The examined GS method has an even lower probability than both VAM versions when applied to quadratic matrices. For the 9×9 scenario, the likelihood of attaining the optimal objective value lies at 4.8 %, for 16×16 matrices it is infinitesimal small at 0.2 %. For non-quadratic scenarios, VAM-nq and GS reveal a different picture: as the difference between dimensions increases, both methods approach a 100 % probability for reaching the optimal objective value, VAM-nq earlier than GS. For example, the probability of solving a 5×13 scenario optimally lies at 94.5 % for VAM-nq and at 68.5 % for the GS method. A 5×55 scenario is solved optimally 99.7 % of the time when applying the VAM-nq and 92.9 % of the time when the Greedy Search method is utilised.

Having presented the computational times and the probability of reaching the actual optimum for the examined combinations of variables, this section present the results associated with MAPE (see Figure 4). The mean deviation to the optimum of VAM increase continuously for non-quadratic matrices as the difference between the dimensions increases successively. That is, while the deviation from the optimal objective value of the original VAM continuous to grow as the matrix size increases, the deviation of the VAM-nq approaches 0 %. This means an optimal solution is generated more often as the difference between the dimensions increase. This clearly demonstrates that the VAM-nq is more suitable to compute non-quadratic instances than the original VAM.

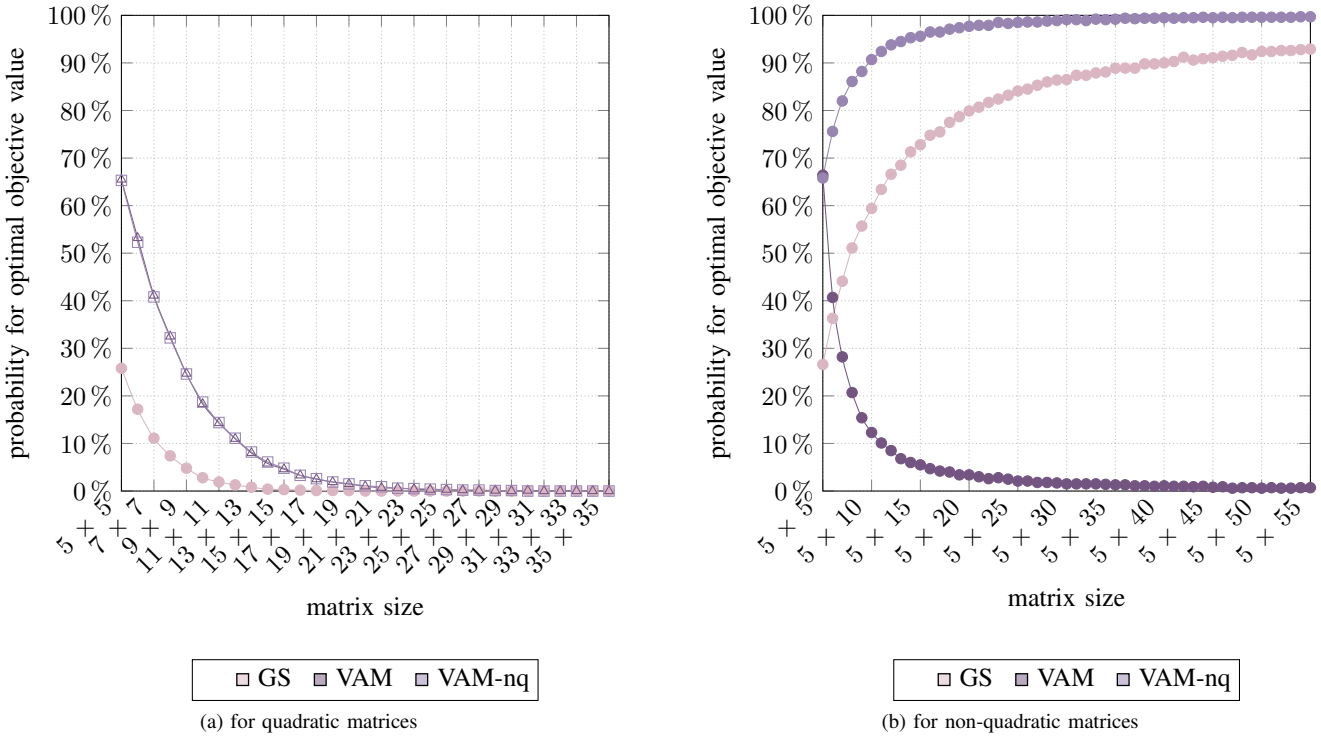
Figure 6 shows the results of simulation experiments performed by applying the original VAM, the VAM-nq and the GS algorithm to different assignment scenarios. For non-quadratic matrices, it is also evident that the original VAM produces objective values that are up to 300 % higher than the corresponding optimal solution. The results yielded by VAM-nq, on the other hand, display almost no deviation



(a) For quadratic matrices

(b) For non-quadratic matrices

Figure 2: Mean computation time for ILP (CPLEX-solver), HM, JVC, GS, VAM and VAM-nq for quadratic and non-quadratic matrices in microseconds (20.000 averaged samples for each test point)



(a) for quadratic matrices

(b) for non-quadratic matrices

Figure 3: Probability for reaching the optimal objective value for quadratic and non-quadratic matrices (20.000 averaged samples for each test point)

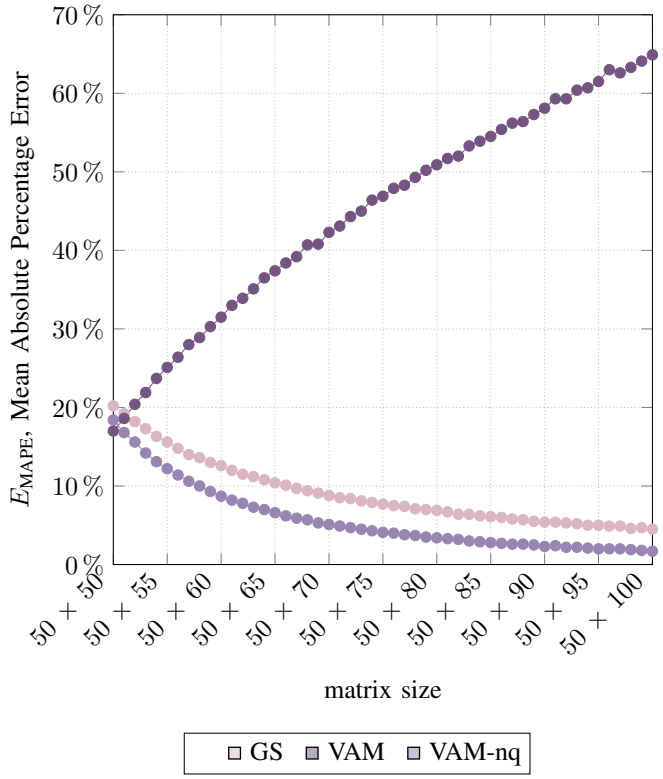


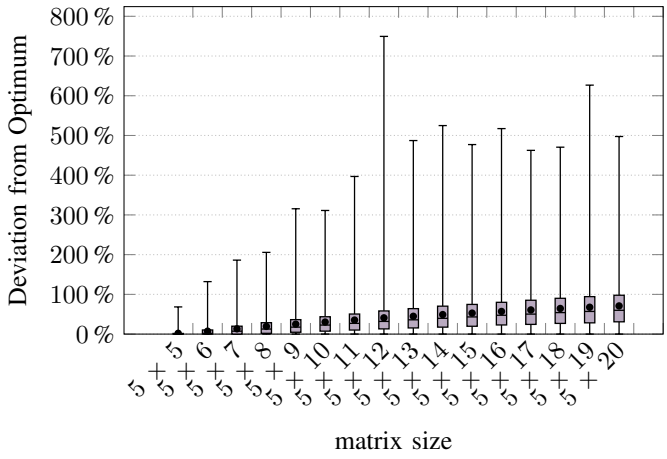
Figure 4: E_{MAPE}, Mean Absolute Percentage Error of the original VAM, VAM-nq and Greedy Search from the optimal objective value for non-quadratic matrices (20.000 averaged samples for each test point)

from the optimal objective value and this method, on average, manages to generate optimal solutions in a majority of non-quadratic cases. However, the results also show that the VAM produces slightly better results for quadratic matrices than the VAM-nq, but the differences in those cases are considered to be negligible.

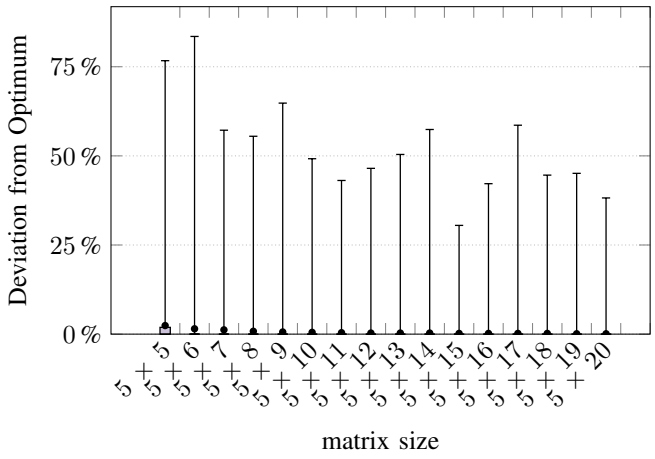
V. Conclusion

The results of our experiments have shown that JVC is superior to all other methods discussed in this paper. JVC is able to provide optimal solutions to each assignment scenario in a relatively short time. For large non-quadratic matrices, the approximation method VAM-nq yields better results than JVC in terms of computational time, even if some scenarios may not be solved optimally. As such, the selection of a method for practical application depends on the circumstances and prioritised objective.

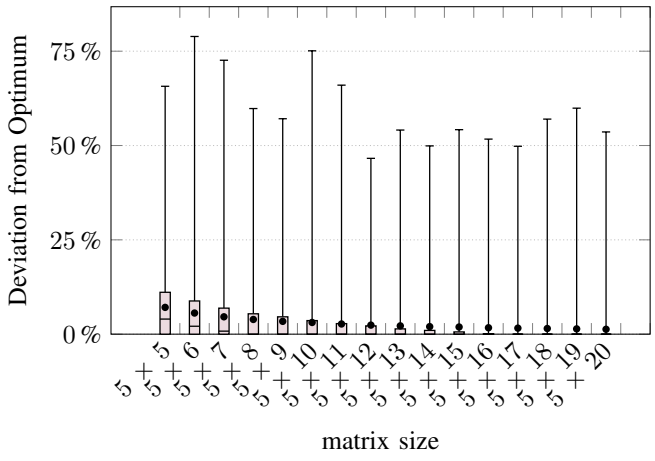
VAM is substantially faster at calculating results than the HM and ILP across all matrix sizes. However, the VAM results in a MAPE of 20 % from the optimal objective value for quadratic matrices (starting with approx. 15×15). In contrast, the MAPE increases to 65 % for larger non-quadratic matrices (starting at approx. 50×100). On the other side, VAM produces insufficient results for all non-quadratic matrice sizes and deviates considerably from the optimum. The adapted version



(a) VAM



(b) VAM-nq



(c) Greedy Search

Figure 5: Percentage Error of VAM, VAM-nq and Greedy Search from the optimal objective value with increasing matrix size (each box-plot represents a set of 20.000 samples)

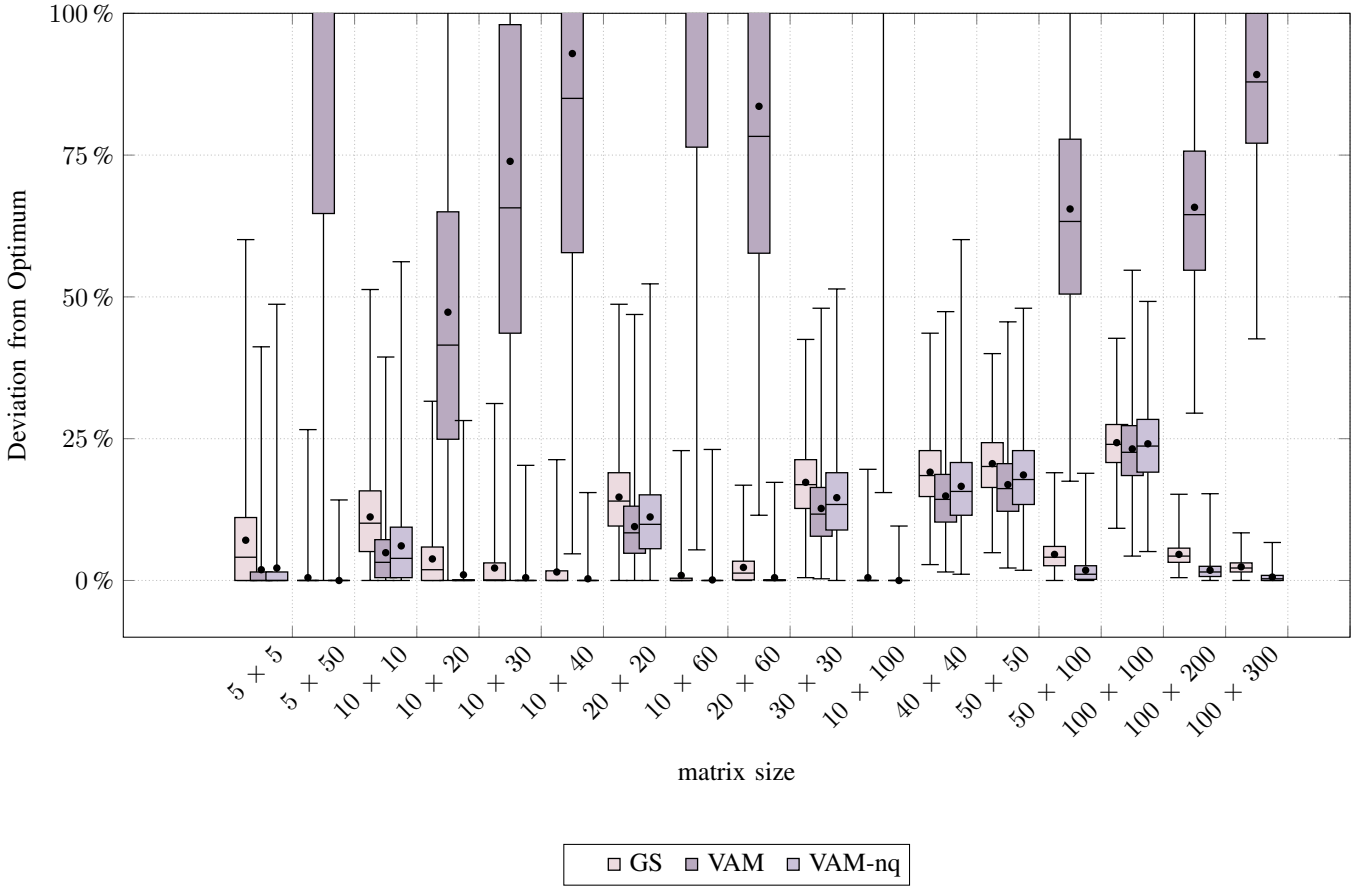


Figure 6: Deviation from optimum of Greedy Search the original VAM and VAM-nq from the optimal objective value for different matrix sizes (each box-plot represents a set of 20.000 samples)

of VAM, introduced as VAM-nq, yields objective values that deviate slightly more from the optimal objective value than the original VAM for quadratic instances, but provides much better results for non-quadratic scenarios, reaching an optimum solution in most of the cases. Moreover, the likelihood of solving a non-quadratic matrix optimally approaches 100 %, as the difference between the dimensions increases.

Based on these findings, we propose that the JVC algorithm proved to be most suitable for quadratic assignment problems. For non-quadratic scenarios, JVC is currently deemed to be the best method for solving quadratic and non-quadratic assignment problems with optimal precision. If performance in terms of computation time is more important than the quality of the solution, VAM-nq can be used for large non-quadratic matrices to generate a faster approximated solution. The probability for VAM-nq to calculate an optimal solution for non-quadratic problem instances larger than 5×25 approaches 100 % as the scale increases. Additionally, the time required by VAM-nq to calculate solutions is substantially faster than JVC for such scenarios and the computational time remains feasible even for large scale assignment problems. In summary, we propose that VAM-nq can be applied to streamline the computational effort and duration required to solve large non-quadratic scenarios at only minor to negligible detriment to the solution's quality.

References

Ahmed, M., Khan, A., Ahmed, F. and Uddin, M. (2016), 'Customized Vogel's Approximation Method (CVAM) for Solving Transportation Problems', *Buletinul Institutulu Politehnic* 62 (66), pp. 31–44.

Applegate, D. and Cook, W. (1993), 'Solving large-scale matching problems', *Network Flows and Matching: First DIMACS Implementation Challenge*, ed. by D. Johnson and C. McGeoch, vol. 12, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Providence, Rhode Island: American Mathematical Society, pp. 557–576, ISBN: 9780821865989.

Atamtürk, A. (1993), *Efficient algorithms for the minimum cost perfect matching problem on general graphs*, Bilkent University.

Balakrishnan, N. (1990), 'Modified Vogel's approximation method for the unbalanced transportation problem', *Applied Mathematics Letters* 3 (2), pp. 9–11, ISSN: 08939659.

Burkard, R.E. and Derigs, U. (1980), *Assignment and Matching Problems: Solution Methods with FORTRAN-Programs*, vol. 184, Lecture notes in economics and mathematical systems, Berlin, Heidelberg usw.: Springer, ISBN: 9783540102670.

Cattrysse, D.G. and van Wassenhove, L.N. (1992), 'A survey of algorithms for the generalized assignment problem', *European Journal of Operational Research* 60 (3), pp. 260–272, ISSN: 0377-2217.

Cheriyan, J., Hagerup, T. and Mehlhorn, K. (1996), 'An $o(n^3)$-Time Maximum-Flow Algorithm', *SIAM Journal on Computing* 25 (6), pp. 1144–1170.

Cunningham, W.H. and Marsh, A.B. (1978), 'A primal algorithm for optimum matching', *Polyhedral Combinatorics*, ed. by M.L. Balinski, E.M.L. Beale, G.B. Dantzig, L. Kantorovich, T.C. Koopmans, A.W. Tucker, P. Wolfe, V. Chvátal, R.W. Cottle, H.P. Crowder, J.E. Dennis, B.C. Eaves, R. Fletcher, M. Iri, E.L. Johnson, C. Lemarechal, C.E. Lemke, G.P. McCormick, G.L. Nemhauser, W. Oettli, M.W. Padberg, M.J.D. Powell, J.F. Shapiro, L.S. Shapley, K. Spielberg, H. Tuy, D.W. Walkup, R. Wets, C. Witzgall and A.J. Hoffman, vol. 8, Mathematical Programming Studies, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 50–72, ISBN: 978-3-642-00789-7.

Cygan, M., Gabow, H.N. and Sankowski, P. (2015), 'Algorithmic Applications of Baur-Strassen's Theorem: Shortest Cycles, Diameter and Matchings', *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, IEEE, pp. 531–540, ISBN: 978-0-7695-4874-6.

Derigs, U. (1981), 'A shortest augmenting path method for solving minimal perfect matching problems', *Networks* 11 (4), pp. 379–390, ISSN: 00283045.

Derigs, U. (1986), 'Solving large-scale matching problems efficiently: A new primal matching approach', *Networks* 16 (1), pp. 1–16, ISSN: 00283045.

Derigs, U. (1988), 'Solving non-bipartite matching problems via shortest path techniques', *Annals of Operations Research* 13 (1), pp. 225–261.

Derigs, U. and Metz, A. (1986), 'On the use of optimal fractional matchings for solving the (integer) matching problem', *Computing* 36 (3), pp. 263–270, ISSN: 0010-485X.

Derigs, U. and Metz, A. (1991), 'Solving (large scale) matching problems combinatorially', *Mathematical Programming* 50 (1-3), pp. 113–121, ISSN: 0025-5610.

Díaz-Parra, O., Ruiz-Vanoye, J.A., Bernábe Loranca, B., Fuentes-Penna, A. and Barrera-Cámara, R.A. (2014), 'A Survey of Transportation Problems', *Journal of Applied Mathematics* 2014 (3), pp. 1–17, ISSN: 1110-757X.

Dinagar, D.S. and Keerthivasan, R. (2018), 'Solving Fuzzy Transportation Problem Using Modified Best Candidate Method', *Journal of Computer and Mathematical Sciences* 9 (9), pp. 1179–1186.

Dinic, E.A. and Kronrod, M.A. (1969), 'An algorithm for the solution of the assignment problem', *Soviet Math. Dokl*, vol. 6, pp. 1324–1326.

Duan, R. and Pettie, S. (2014), 'Linear-Time Approximation for Maximum Weight Matching', *Journal of the ACM* 61 (1), pp. 1–23.

Edmonds, J., Johnson, E.L. and Lockhart, S.C. (1969), 'Blossom I: a computer code for the matching problem', *IBM TJ Watson Research Center*.

Feder, T. and Motwani, R. (1995), 'Clique Partitions, Graph Compression and Speeding-Up Algorithms', *Journal of Computer and System Sciences* 51 (2), pp. 261–272.

Freling, R., Wagelmans, A.P.M. and Paixão, J.M.P. (2001), 'Models and Algorithms for Single-Depot Vehicle Scheduling', *Transportation Science* 35 (2), pp. 165–180, ISSN: 0041-1655.

Gabow, H.N. (1976), 'An Efficient Implementation of Edmonds' Algorithm for Maximum Matching on Graphs', *Journal of the ACM* 23 (2), pp. 221–234.

Gabow, H.N. (1985), 'A scaling algorithm for weighted matching on general graphs', *26th Annual Symposium on Foundations of Computer Science (sfcs 1985)*, IEEE, pp. 90–100, ISBN: 0-8186-0644-4.

Gabow, H.N. (1990), 'Data structures for weighted matching and nearest common ancestors with linking', *Proceedings of the first annual ACM-SIAM symposium on Discrete algorithms*, pp. 434–443.

Gabow, H.N. (2017), 'A Data Structure for Nearest Common Ancestors with Linking', *ACM Transactions on Algorithms* 13 (4), pp. 1–28.

Gabow, H.N., Galil, Z. and Spencer, T.H. (1989), 'Efficient implementation of graph algorithms using contraction', *Journal of the ACM* 36 (3), pp. 540–572.

Gabow, H.N. and Tarjan, R.E. (1991), 'Faster scaling algorithms for general graph matching problems', *Journal of the ACM* 38 (4), pp. 815–853.

Galil, Z., Micali, S. and Gabow, H.N. (1982), 'Priority queues with variable priority and an O(EV log V) algorithm for finding a maximal weighted matching in general graphs', *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, IEEE, pp. 255–261.

Gerngross, P. (1991), *Zur Implementation von Edmonds' Matching Algorithmus: Datenstrukturen und verschiedene Varianten*, Universität Augsburg.

Goldberg, A.V. and Karzanov, A.V. (2004), 'Maximum skew-symmetric flows and matchings', *Mathematical Programming* 100 (3), ISSN: 0025-5610.

Goldberg, A.V. and Kennedy, R. (1997), 'Global Price Updates Help', *SIAM Journal on Discrete Mathematics* 10 (4), pp. 551–572.

Goyal, S.K. (1984), 'Improving VAM for Unbalanced Transportation Problems', *Journal of the Operational Research Society* 35 (12), pp. 1113–1114, ISSN: 1476-9360.

Grötschel, M. and Holland, O. (1985), 'Solving matching problems with linear programming', *Mathematical Programming* 33 (3), pp. 243–259, ISSN: 0025-5610.

Harvey, N.J.A. (2006), 'Algebraic Structures and Algorithms for Matching and Matroid Problems', *2006 47th Annual*

IEEE Symposium on Foundations of Computer Science (FOCS'06), IEEE, pp. 531–542, ISBN: 0-7695-2720-5.

Havel, T.F., Kuntz, I.D. and Crippen, G.M. (1983), 'The theory and practice of distance geometry', *Bulletin of Mathematical Biology* 45 (5), pp. 665–720, ISSN: 0092-8240.

Hopcroft, J.E. and Karp, R.M. (1973), 'An n5/2 Algorithm for Maximum Matchings in Bipartite Graphs', *SIAM Journal on Computing* 2 (4), pp. 225–231.

Jonker, R. and Volgenant, A. (1987), 'A shortest augmenting path algorithm for dense and sparse linear assignment problems', *Computing* 38 (4), pp. 325–340, ISSN: 0010-485X.

Karzanov, A.V. (1976), 'Efficient implementations of Edmonds' algorithms for finding matchings with maximum cardinality and maximum weight', *Studies in Discrete Optimization*, pp. 306–327.

Kazakidis, G. (1980), *Die Löung minimaler perfekter Matchingprobleme mittels kürzester erweiternder Pfade*.

Korukoğlu, S. and Ballı, S. (2011), 'A Improved Vogel's Approximatio Method for the Transportation Problem', *Mathematical and Computational Applications* 16 (2), pp. 370–381.

Krumke, S.O. and Thielen, C. (2013), 'The generalized assignment problem with minimum quantities', *European Journal of Operational Research* 228 (1), pp. 46–55, ISSN: 0377-2217.

Kuhn, H.W. (1955), 'The Hungarian method for the assignment problem', *Naval Research Logistics Quarterly* 2 (1-2), pp. 83–97, ISSN: 00281441.

Lawler, E.L. (1976), *Combinatorial optimization: Networks and matroids*, Unabridged reprint ... orig. publ. by Holt, Rinehart & Winston in 1976, Mineola, NY: Dover Publ, ISBN: 978-0486414539, URL: http://www.loc.gov/catdir/description/dover032/00060242.html.

Lessard, R., Rousseau, J.-M. and Minoux, M. (1989), 'A new algorithm for general matching problems using network flow subproblems', *Networks* 19 (4), pp. 459–479, ISSN: 00283045.

Li, J.-Q., Mirchandani, P.B. and Borenstein, D. (2007), 'The vehicle rescheduling problem: Model and algorithms', *Networks* 50 (3), pp. 211–229, ISSN: 00283045.

Micali, S. and Vazirani, V.V. (1980), 'An algorithm for finding maximum matching in general graphs', *21st Annual Symposium on Foundations of Computer Science (sfcs 1980)*, IEEE, pp. 17–27.

Miller, D.L. and Pekny, J.F. (1995), 'A Staged Primal-Dual Algorithm for Perfect b-Matching with Edge Capacities', *ORSA Journal on Computing* 7 (3), pp. 298–320, ISSN: 0899-1499.

Minoux, M. (1982), *A New Polynomial Cutting-plane Algorithm for Maximum Weight Matchings in General Graphs*.

Mucha, M. and Sankowski, P. (2004), 'Maximum Matchings via Gaussian Elimination', *45th Annual IEEE Symposium on Foundations of Computer Science*, IEEE, pp. 248–255, ISBN: 0-7695-2228-9.

Munkres, J. (1957), 'Algorithms for the Assignment and Transportation Problems', *Journal of the Society for Industrial and Applied Mathematics* 5 (1), pp. 32–38, ISSN: 0368-4245.

Nahar, J., Rusyaman, E. and Putri, S.D.V.E. (2018), 'Application of improved Vogel's approximation method in minimization of rice distribution costs of Perum BULOG', *IOP Conference Series: Materials Science and Engineering* 332.

Osman, I.H. (1995), 'Heuristics for the generalised assignment problem: simulated annealing and tabu search approaches', *OR Spectrum* 17 (4), pp. 211–225, ISSN: 0171-6468.

Paul, S. (2018), 'A novel initial basic feasible solution method for transportion problem', *International Journal of Advanced Research in Computer Science* 9 (1), pp. 472–474.

Pentico, D.W. (2007), 'Assignment problems: A golden anniversary survey', *European Journal of Operational Research* 176 (2), pp. 774–793, ISSN: 0377-2217.

Pulleyblank, W. (1973), *Faces of Matching Polyhedra*.

Reinfeld, N.V. and Vogel, W.R. (1958), *Mathematical Programming*, Prentice-Hall, Englewood Cliffs.

Selmair, M., Meier, K.-J. and Wang, Y. (2019), 'Solving non-quadratic Matrices in assignment Problems with an improved Version of Vogel's Approximation Method', *European Conference of Modelling and Simulation*.

Shimshak, D., Kaslik, A.J. and Barclay, T. (1981), 'A Modification Of Vogel'S Approximation Method Through The Use Of Heuristics', *INFOR: Information Systems and Operational Research* 19 (3), pp. 259–263, ISSN: 0315-5986.

Shore, H.H. (1970), 'The Transportation problem and the Vogel Approximation Method', *Decision Sciences* 1 (3-4), pp. 441–457, ISSN: 0011-7315.

Skiena, S. (1990), 'The Cycle Structure of Permutations', *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*, pp. 20–24.

Trick, M. (1987), *Networks with Additional Structured Constraints*, Georgia Institute of Technology.

Witzgall, C. and Zahn, C.T. (1965), 'Modification of Edmonds' maximum matching algorithm', *J. Res. Nat. Bur. Standards Sect. B*.

Machine Learning
for
Big Data

On the effect of decomposition granularity on DeTraC for COVID-19 detection using chest X-ray images

Nicole Panashe Mugova[1], Mohammed M. Abdelsamea[1,2,*], Mohamed Medhat Gaber[1,3]
[1]School of Computing and Digital Technology, Birmingham City University, Birmingham, B4 7XG, UK
[2]Faculty of Computers and Information, Assiut University, Assiut, Egypt
[3]Faculty of Computer Science and Engineering, Galala University, Egypt
* Correspondence: mohammed.abdelsamea@bcu.ac.uk

Abstract

COVID-19 is a growing issue in society and there is a need for resources to manage the disease. This paper looks at studying the effect of class decomposition in our previously proposed deep convolutional neural network, called DeTraC (Decompose, Transfer and Compose). DeTraC can robustly detect and predict COVID-19 from chest X-ray images. The experimental results showed that changing the number of clusters (decomposition granularity) affected the performance of DeTraC and influenced the accuracy of the model. As the number of clusters increased, the accuracy decreased for the shallow tuning mode but increased for the deep tuning mode. This shows the importance of using suitable hyperparameter settings to get the best results from the DeTraC deep learning model. The highest accuracy obtained, in this study, was 98.33% from the deep tuning model.

INTRODUCTION

In late December 2019, an outbreak of a virus emerged from Wuhan, China (Wu, Chen, and Chan, 2020). The disease Sars-Cov-2, better known as COVID-19, spread at an alarming rate, affecting countries worldwide. COVID-19 primarily affects the respiratory system (similar to pneumonia) and chest X-rays play a crucial role in assessing the presence, severity, and progression of COVID-19. The virus has significantly affected the healthcare sector, resulting in shortages of staff and necessary protective equipment (Organization, 2020).

X-rays, also known as plain radiography (Johns et al. 1983) are a popular medical imaging technique. X-rays are significantly cheaper than CT scans and are painless, fast, and non-invasive. Understanding chest X-rays requires expert knowledge and experience but can be time-consuming. Using both chest X-rays and artificial intelligence will mean the extent of the disease can be determined. Long-term, diagnostic tools will benefit radiographers and other healthcare professionals involved in the diagnosis process.

Deep learning offers a reliable way to identify abnormalities in medical images, allowing for preventive screening and personalised patient data benefiting both the patient and doctor. Deep learning models have been applied to reduce pressure on healthcare professionals. It has been evident that pretraining deep neural networks on a large generic data set is an effective and efficient way to retrain such models on other tasks. The three types of Neural Networks that form the basis for most pre-trained models in deep learning are – artificial neural networks as multilayer perceptron, convolutional neural network (CNN) and recurrent neural networks (Wang et al., 2019).

CNN is a popular deep learning approach that has shown superior achievements in the medical imaging domain. The primary success of CNN is due to its ability to learn local features automatically from images. One of the most popular strategies for training CNN architecture is to transfer knowledge from a pre-trained network that fulfilled one generic task (e.g., large-scale image recognition) into a new domain-specific task (e.g., COVID-19 detection). Transfer learning (Abbas et al. 2018) is faster and easy to apply, especially without the need for a sufficient annotated dataset for training. Consequently, many researchers tend to utilise and adapt this strategy especially with medical imaging and COVID-19 detection (Abbas et al. 2020b). Transfer learning can be generally accomplished with two major scenarios: a) “shallow tuning”, which adapts the last few classification layers to deal with the new specific task and freezes the parameters of the remaining layers without training; and b) “deep tuning” which aims to retrain all the weights of the pre-trained network from end-to-end manner and requires a huge amount of data to overcome overfitting problem.

Using deep learning techniques to recognise the prominent features in chest X-rays allows for better diagnosis and in some cases is better than a radiologist (Hosny et al., 2018). Amongst the various deep learning techniques, DeTraC (decompose, transfer, and compose) has shown a high accuracy of 93.1% in diagnosing COVID-19. DeTraC is a deep CNN that can be trained using a small number of images (Abbas et al. 2020a), which is beneficial given the limited number of images currently available.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

One problem that can arise from using deep learning for medical diagnosis is the overfitting problem. This is where the algorithm performs well on the training and validation datasets but when presented with new testing data it does not perform well. This affects the reproducibility of the algorithm, so it is imperative to test different hyperparameters in the hope of finding the optimal hyperparameter settings.

In this paper, the behaviour of DeTraC is investigated when changing the hyperparameters – specifically the number of clusters in the class decomposition component using k-means clustering. A k-means clustering method is an example of an unsupervised learning technique. The main step involved in k-means clustering is to group different instances (examples) based on their similarities. One of the disadvantages of using the k-means clustering algorithm is that the number of clusters needs to be specified, if the number is unsuitable, the performance of DeTraC will be affected. So, changing the value of k will allow the effectiveness of the algorithm to be observed and to quantify its usefulness. Typically, the value of k is important as it determines the number of centroids going around the data (Pham et al. 2005). Although the value of k affects the performance of clustering directly, it affects the performance of DeTraC indirectly. Thus, using cluster performance metrics related to cohesion and separation may not be the best way to investigate the effect of k on DeTraC. In this paper, an experimental exploration on the effect of the number of clusters (i.e., decomposition granularity) on the performance of DeTraC is provided and discussed.

To achieve the main objective, the rest of the paper is structured as follows: The following section provides a brief overview of existing work on Deep Learning, details about the dataset, the workflow of DeTraC, and the experimental study. This is followed by the discussion of the results. Finally, the paper is concluded with a summary and pointers to possible future work.

RELATED WORK

Deep learning has "already left its mark" in healthcare and continues to provide new solutions to current problems in society (Esteva et al. 2021). Developing new advancements such as diagnostic tools benefits healthcare professionals and patients by providing quicker and more accurate diagnoses of certain diseases. The review of the literature has highlighted a gap in research and development. It shows the importance of why scientists and healthcare professionals need to work together, to minimise the effects of new diseases that arise by sharing data and research.

(Stephen et al., 2019) developed a CNN to detect Pneumonia from chest x-ray images. They noted how a CNN was the better choice for image segmentation because of its ability to extract abstract 2D features through learning. This is beneficial for illnesses such as Pneumonia, SARS, and COVID-19 – which all affect the lungs. (Abbas et al. 2020a) proposed a deep CNN called DeTraC (Decompose, Transfer and Compose) to detect COVID-19 from chest X-rays. One of the barriers researchers face when dealing with image classification is data irregularities and (Abbas et al. 2020a) suggests DeTraC can cope with this issue. (Rahaman et al., 2020; Ismael and Şengür, 2021) also use deep learning techniques to compare different pre-trained CNN models. Showing how CNNs are effective tools at image classification, however (Rahaman et al., 2020) note there are essential factors that influence the performance of the individual models, such as, image quantity, model complexity, and the distribution of the dataset. (Rahaman et al., 2020) used 860 images (300 healthy, 300 pneumonia, and 260 COVID-19). Data augmentation techniques were applied to the 1764 images and DeTraC achieved an accuracy of 97.35% (Abbas et al. 2020a). Even though (Rahaman et al., 2020) has a lower number of images, the CNN model performs well with an accuracy of 89.3% using a VGG19 model.

While the other sources use a pre-trained model and show the benefits of using a VGG19 model for image classification, (Stephen et al., 2019) does not rely on a pre-trained model but designs a CNN from scratch. The CNN designed can detect pneumonia from chest X-ray images with high accuracy (93.7%). Despite not relying on a pre-trained model the accuracy achieved is still higher than (Rahaman et al., 2020) but lower than (Abbas et al. 2021). (Stephen et al., 2019) suggests that the model they developed could alleviate some of the problems that researchers face such as reliability and interpretability. However, the method is not as detailed enough for it to be reproduced by someone else without seeing the source code which is not provided.

DeTraC

DeTraC consists of three phases – class decomposition, transfer learning, and class composition. Before the class decomposition phase, a pre-trained CNN model acts as the feature extractor, extracting deep features from the images inputted to build a deep feature space (Abbas et al. 2021). The deep feature space is very important and (Abbas et al. 2021) applied PCA to reduce the high dimensionality of the feature space. Applying PCA to the feature space results in a lower dimensionality and ignores any highly correlated features. Then the reduced feature space decomposes the original classes into decomposed classes (Abbas et al. 2021). Once the representation of each image is constructed based on the associated pre-trained model. In the second phase, DeTraC adds a novel class-decomposition layer to several pre-trained CNN models for the decomposition of classes, in an unsupervised way, and accomplish the training using sophisticated gradient descent optimisation method. Finally, we distinguish between normal and abnormal cases using an error-correction criterion. Class decomposition allows for easing the local

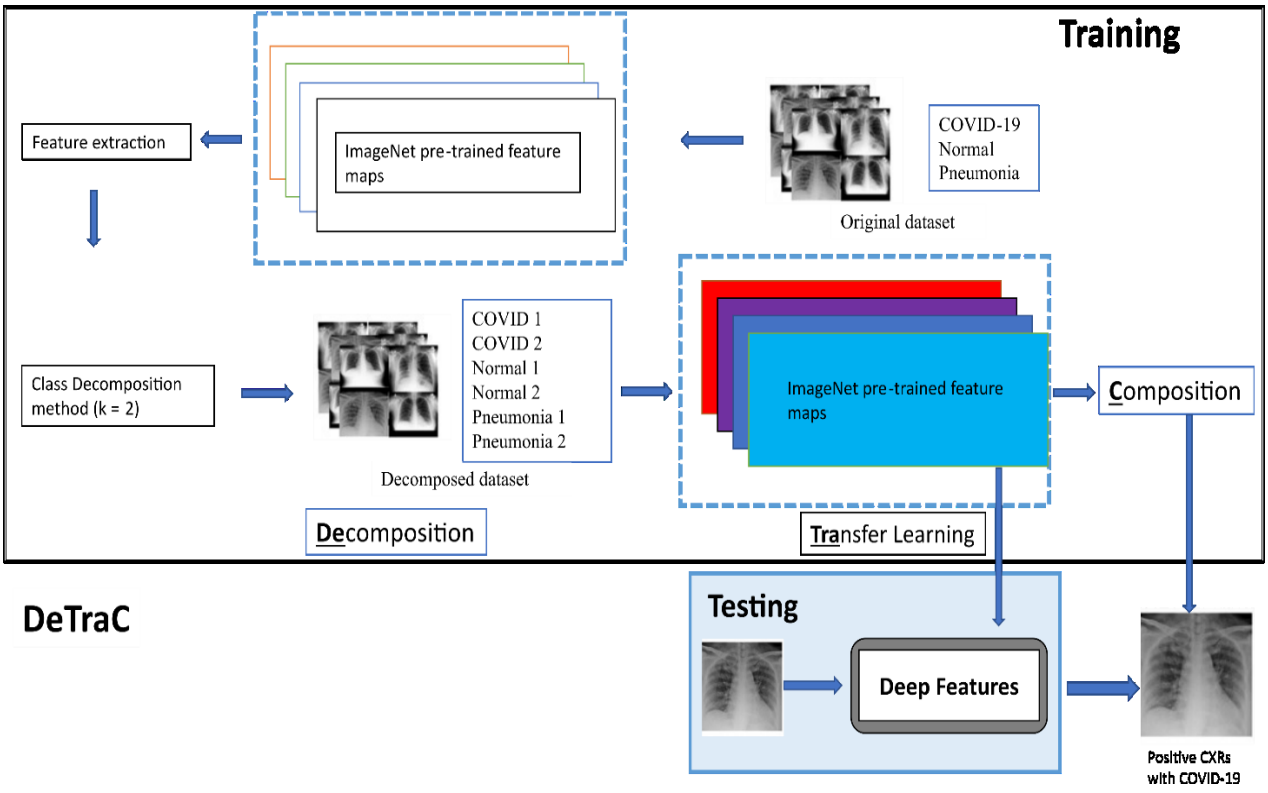


Figure 1: Flow chart outline of DeTraC method used to generate results.

structure of the dataset, and consequently enhancing the effectiveness of the model to deal with any irregularities (Abbas et al. 2020b) in the data distribution. This is achieved by making the complex problem easier to learn.

To illustrate the idea behind class decomposition, assume that the original dataset is denoted as $\boldsymbol{A}$, where $\boldsymbol{A}$ can be represented as:

$$\boldsymbol{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots \dots \dots, \boldsymbol{a}_n\},$$

where n number of images, and each image can be represented as a set of features as:

$$\boldsymbol{a_i} = (a_{i1}, a_{i2}, \dots, a_{in}).$$

Moreover, assume that $\boldsymbol{L}$ is a class category of dataset $\boldsymbol{A}$, then ($\boldsymbol{A}$, $\boldsymbol{L}$) can be rewritten as:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \boldsymbol{L} = \{l_1, l_2, \dots, l_k\},$$

where k is the number of classes while m is the number of features used for each image.

Class decomposition aims to divide each class in a dataset independently into subclasses. For example, if dataset $\boldsymbol{A}$ denoted to CXR images with 2 classes (i.e. normal and abnormal) then each class in $\boldsymbol{L}$ will be divided into two classes, resulting in a new dataset (denoted as dataset $\boldsymbol{B}$) with 4 sub-classes. Therefore, the relationship between dataset $\boldsymbol{A}$ and $\boldsymbol{B}$ can be mathematically described as:

$$\boldsymbol{A}=(\boldsymbol{A}|\boldsymbol{L}) \mapsto \boldsymbol{B}=(\boldsymbol{B}|\boldsymbol{C}),$$

where both $\boldsymbol{A}$ and $\boldsymbol{B}$ have an equal number of instances and $\boldsymbol{C}$ contains labels of the new subclasses, e.g.,

$$\boldsymbol{C} = \sum_{i=1}^{\kappa} \sum_{j=1}^{c} L_{ij} \quad , \quad c = 2$$

Accordingly, the feature space for dataset $\boldsymbol{A}$ and $\boldsymbol{B}$ can be illustrated as:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{11} & \dots & a_{1n} & \ell_1 \\ a_{21} & a_{22} & \dots & a_{2n} & \ell_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ell_2 \\ a_{m1} & a_{m2} & \dots & a_{mn} & \ell_2 \end{bmatrix},$$

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{11} & \dots & b_{1n} & \ell_{11} \\ b_{21} & b_{22} & \dots & b_{2n} & \ell_{1c} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ell_{21} \\ b_{m1} & b_{m2} & \dots & b_{mn} & \ell_{2c} \end{bmatrix}$$

Once the training is accomplished those subclasses are recombined back to the original problem, by relabelling patterns of new subclasses (Abbas, Abdelsamea and Gaber, 2020).

In this paper, we study the effect of this decomposition granularity on the performance of DeTraC (with different training modes) in detecting COVID-19 cases using chest X-ray images.

RESULTS

This section includes the dataset that we used and details our hyperparameters setting as well as discussing the experimental results.

The experiments were carried out in Google Colab on a laptop with the following configuration Intel(R) Core (TM) i7-8550U CPU with 8.00 GB RAM.

Dataset

The datasets used were collected from Kaggle (Rahman Tawsifur, 2021), and the owners of the dataset collected the images in collaboration with medical doctors.

The datasets used were composed of:

- 1200 samples of COVID-19 chest X-rays
- 1341 samples of Normal chest X-rays
- 1345 samples of Viral Pneumonia chest X-rays

DeTraC adaptation

In this paper, a shallow-tuning mode was used during the adaptation, weight initialisation, and training of the AlexNet pre-trained model. We used the off-the-shelf CNN local features (extracted from the last fully connected layer) of pre-trained models on the ImageNet dataset. However, due to the high dimensionality of the feature space associated with the images, we applied principal component analysis (PCA) to project the high-dimension feature space into a lower dimension (where the highly correlated features were ignored). This step was important for the class decomposition layer of DeTraC to produce more homogeneous classes and reduce the memory requirements.

In addition to the fact that DeTraC can cope with data irregularities using its class decomposition layer, DeTraC can also provide an efficient solution to overcome the limited availability of training images. This is by transferring knowledge from a generic object recognition task to our specific-domain tasks using ImageNet pre-trained model (e.g., ResNet) as adopted in the transfer learning component of DeTraC.

Hyperparameter settings

The datasets were split into training (80%) and testing (20%) sets. Then, to begin our investigation on the behaviour of the class decomposition component on DeTraC, the elbow method was adopted to change the value of k in the k-means algorithm. To begin with, the value of k was 2 and then was increased by +1 each time until k = 10. The k-means clustering was applied to all the classes (normal, COVID-19, and Pneumonia). Once the method of changing the value of k had been decided, the hyperparameter settings were set. The number of epochs was 5, the batch size was 64, the number of classes was 3, the number of k-folds was 5 and the learning rate for the feature extractor and feature composer was 0.001. They were kept the same for each value of k. Each value of k was run on both the shallow tuning mode and the deep tuning mode to see if there was a difference in the performance. Shallow tuning is where all the pre-trained layers have their weights frozen and the custom classification layer has its weights active. Deep tuning is where all the layers have their weights active (Abbas et al. 2020a). The pre-trained model has 37 layers. The fine-tuning mode allows the user to choose the number of layers to be frozen, but this option was not used for this experiment. The changing accuracy of the model was used as the evaluation metric, see Table 1.

Table 1: Summary table showing the results of changing the value of k on DeTraC.

Value of k	Shallow-Tuning Accuracy (%)	Deep-Tuning Accuracy (%)
2	94.64	97.49
3	94.44	97.66
4	93.98	97.82
5	92.23	97.41
6	91.81	98.01
7	88.91	98.33
8	94.82	98.15
9	92.61	97.78
10	92.24	97.59

DISCUSSION AND CONCLUSION

Class decomposition has been proposed to enhance low variance classifiers facilitating more flexibility to their decision boundaries. In (Abbas et al. 2021), we previously validated DeTraC for the detection of COVID-19 using chest X-ray images when data irregularities challenging problem is presented. This is by adding a class decomposition layer to the pre-trained models, which aims at partitioning each class within the image dataset into sub-classes and then assign new machine labels to the new set, where each subset is treated as an independent class, then those subsets are assembled back to produce the final predictions. Our previous experimental results showed the robustness of DeTraC in the detection of COVID-19 cases from a comprehensive image dataset. High accuracy of 95.12% with a sensitivity of 97.91%, a specificity of 91.87%, and a precision of 93.36%, was achieved in the detection of COVID-19 images from normal, and severe acute respiratory syndrome (SARS) cases.

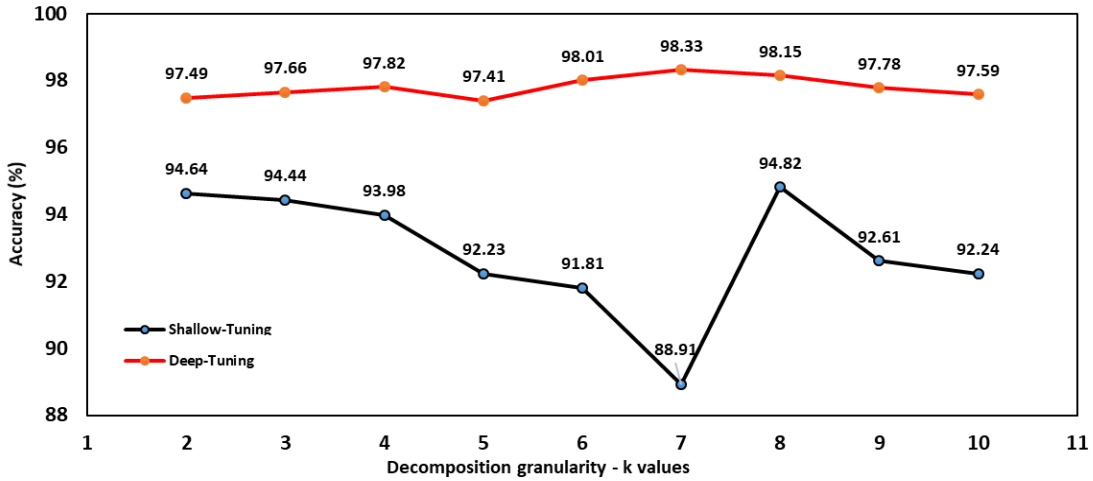


Figure 2: The effect of decomposition granularity on DeTraC.

In this paper, we study the effect of decomposition granularity on our previously proposed convolutional neural network architecture (DeTraC). From the results demonstrated in the previous section, there is a clear link between the value of k and the accuracy of DeTraC. Both the shallow and tuning modes behaved differently.

For the shallow tuning mode, when the value of k increased, the accuracy decreased. However, when k was equal to 8, the accuracy rose significantly (94.82%) before dropping again for k =9 (92.61%) and k = 10 (92.24%). This suggests for the shallow tuning mode, k = 8 is the optimal value to get the highest accuracy.

For the deep tuning mode, when the value of k increased, the accuracy increased also. When k was equal to 7, the accuracy was 98.33% – similar to shallow tuning and then dropped until k =10.

It is interesting, however, to observe that the decomposition granularity has an oppositive behaviour between the shallow and the deep tuning modes. In shallow tuning, the small values of k seem to have a good effect on the model's accuracy, and then higher values have a negative effect, before lifting up again at k=8, see Figure 2. This suggests that it may be the case that with shallow tuning, the learning of a higher number of classes can be more challenging. However, a more steading behaviour has been shown with deep tuning with a peak at k=7. This suggests that with the effectiveness of deep tuning, the decomposition granularity can be investigated carefully to gain the best possible outcome.

This study has shown that changing the parameters affects the behaviour of DeTraC. The value of k influences the overall accuracy of the algorithm in both the shallow and deep tuning modes. This is a positive development and allows for a better understanding of DeTraC and its use as a diagnostic tool.

In the future, it would be interesting to see how different clustering algorithms other than k-means affect DeTraC and if the results obtained would be similar. DeTraC has proved to be an effective tool in predicting COVID-19 and produces reliable results. The highest accuracy for the deep tuning mode was 98.33% which is similar to the 98.23% obtained by (Abbas et al. 2020a) which is a testament to the algorithm's reproducibility even with different images being used.

REFERENCES

Abbas, A., Abdelsamea, M.M. & Gaber, M.M. (2021) Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Appl Intell 51, 854–864. https://doi.org/10.1007/s10489-020-01829-7.

A. Abbas, M. M. Abdelsamea and M. M. Gaber (2020a) "DeTraC: Transfer Learning of Class Decomposed Medical Images in Convolutional Neural Networks," in IEEE Access, vol. 8, pp. 74901-74913.

Johns, Harold Elford, and John Robert Cunningham. "The physics of radiology." (1983).

Esteva, Andre, et al. "Deep learning-enabled medical computer vision." npj Digital Medicine 4.1 (2021): 1-9.

Hosny, A. et al. (2018) 'Artificial intelligence in radiology', Nature Reviews Cancer. Nature Publishing Group, pp. 500–510. doi: 10.1038/s41568-018-0016-5.

Ismael, A. M. and Şengür, A. (2021) 'Deep learning approaches for COVID-19 detection based on chest X-ray images', Expert Systems with Applications, 164, p. 114054. doi: 10.1016/j.eswa.2020.114054.

Pham, Duc Truong, Stefan S. Dimov, and Chi D. Nguyen. "Selection of K in K-means clustering." Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 219.1 (2005): 103-119. Organization, W. H. (2020) Coronavirus disease (COVID-19) (Accessed: 23 January 2021).

Parvathy, Velmurugan Subbiah, Sivakumar Pothiraj, and Jenyfal Sampson. "Hyperparameter Optimization of Deep Neural Network in Multimodality Fused Medical Image Classification for Medical and Industrial IoT." Smart Sensors for Industrial Internet of Things. Springer, Cham, 2021. 127-146.

Rahaman, M. et al. (2020) 'Identification of COVID-19 samples from chest X-Ray images using deep learning: A comparison of transfer learning approaches the Creative Commons Attribution Non-Commercial License (CC BY-NC 4.0)', Journal of X-Ray Science and Technology, 28, pp. 821–839. doi: 10.3233/XST-200715.

Rahman Tawsifur (2021) COVID-19 Radiography Database Kaggle, Available at: https://www.kaggle.com/tawsifurrahman/COVID19-radiography-database (Accessed: 10 February 2021).

Stephen, O. et al. (2019) 'An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare', Journal of Healthcare Engineering, 2019, pp. 4180949–4180949. doi: 10.1155/2019/4180949.

Wang, B. et al. (2019) 'Deep convolutional neural network with segmentation techniques for chest X-ray analysis', in Proceedings of the 14th IEEE Conference on Industrial Electronics and Applications, ICIEA 2019. Institute of Electrical and Electronics Engineers Inc., pp. 1212–1216. doi: 10.1109/ICIEA.2019.8834117.

Abbas, A., & Abdelsamea, M. M. (2018, December). Learning transformations for automated classification of manifestation of tuberculosis using convolutional neural network. In 2018 13th International Conference on Computer Engineering and Systems (ICCES) (pp. 122-126). IEEE.

Abbas, A., Abdelsamea, M. M., & Gaber, M. (2020b). 4S-DT: Self Supervised Super Sample Decomposition for Transfer learning with application to COVID-19 detection. arXiv preprint arXiv:2007.11450.

Wu, Y.-C., Chen, C.-S. and Chan, Y.-J. (2020) 'The outbreak of COVID-19', Journal of the Chinese Medical Association, 83(3), pp. 217–220. doi: 10.1097/JCMA.0000000000000270.

Towards Intrusion Detection of Previously Unknown Network Attacks

Saif Alzubi[1], Frederic Stahl[1,2], Mohamed Medhat Gaber[3,4]
[1]Department of Computer Science, University of Reading, Reading, UK
[2]Marine Perception Research Department,German Research Center for Artificial Intelligence (DFKI), Oldenburg, Germany
[3]School of Computing and Digital Technology, Birmingham City University, UK
[4]Faculty of Computer Science and Engineering, Galala University, Galala City 43511, Egypt

KEYWORDS

Anomaly Detection; Network Intrusion Detection; Unsupervised algorithms; Emsemble Learning

ABSTRACT

Advances in telecommunication network technologies have led to an ever more interconnected world. Accordingly, the types of threats and attacks to intrude or disable such networks or portions of it are continuing to develop likewise. Thus, there is a need to detect previously unknown attack types. Supervised techniques are not suitable to detect previously not encountered attack types. This paper presents a new ensemble-based Unknown Network Attack Detector (UNAD) system. UNAD proposes a training workflow composed of heterogeneous and unsupervised anomaly detection techniques, trains on attack-free data and can distinguish normal network flow from (previously unknown) attacks. This scenario is more realistic for detecting previously unknown attacks than supervised approaches and is evaluated on telecommunication network data with known ground truth. Empirical results reveal that UNAD can detect attacks on which the workflows have not been trained on with a precision of 75% and a recall of 80%. The benefit of UNAD with existing network attack detectors is, that it can detect completely new attack types that have never been encountered before.

INTRODUCTION

The Internet has become a part of our daily lives with billions of active users. New types of network attacks keep emerging, and there is a need to detect novel attacks without prior knowledge. Yet many Data Mining approaches to detect network attacks are supervised and are only suitable for detecting previously known attack types. There is a need for more exploration of unsupervised approaches as these approaches typically suffer from many false positives [1], a low precision or recall in detecting attacks, and some works are based on older attack types. Hence, the motivation of the paper is to fill this gap by developing the unsupervised ensemble-based Unknown Network Attack Detector (UNAD).

This paper first explores several unsupervised algorithms with respect to precision, recall, F1-Score for their suitability to be included as part of UNAD, namely the Local Outlier Factor (LOF) [2], Isolation Forest (iForest) [3] and Elliptic Envelope [4]. For this exploration, the CICIDS2017 [5] dataset is used. CICIDS2017 comprises 14 attack types, some of which emerged in recent years. Next, the paper proposes UNAD as a composition of some of the evaluated anomaly detecting methods as base learners (LOF and iForest). The reason for choosing an ensemble approach here is that ensemble approaches tend to improve the average accuracy over any member of the ensemble and reduce overfitting [6].

The contributions of the paper are (1) an experimental evaluation of the suitability of unsupervised anomaly detection methods for unknown attack detection; (2) a new heterogeneous unsupervised ensemble technique termed UNAD capable of detecting new previously unseen attack types; and (3) an experimental evaluation showing that UNAD it is capable of achieving high accuracy, precision and recall for detecting unknown attack types, and outperforms its standalone base learners.

Lastly the paper provides an outlook on ongoing and future research with respect to UNAD and also provides concluding remarks.

RELATED WORK

Supervised data mining approaches for Intrusion Detection Systems tend to achieve high accuracy, recall and precision such as [7], [8], [9]. However, they are not suitable for detecting unknown attacks types. Hence, unsupervised techniques have been explored, such as the Intrusion Detection System proposed in [10] based on One-class SVM. However, One-class SVMs tend to have a high computational overhead [11] and thus are not suitable for high-speed network traffic flow. The authors of [1] proposed an unsupervised ensemble model based Intrusion Detection System which achieved relatively high recall and precision. Yet both aforementioned unsupervised approaches have been trained and evaluated on relatively old datasets comprising none of the attack types that emerged over the last 10 years. More recently, iForest [3] was used in [12] to detect abnormal user behaviour on payroll access logs. Ensemble methods have also been used for insider threat detec-

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

tion such as in [13]. The authors of [14] used LOF to detect network attacks as anomalies. However, their study was conducted almost 10 years ago, and it is not clear if it still holds on recent attack types. Elliptic Envelopes [4], another unsupervised anomaly detection method, it has been used by the authors of [15] to detect Injection Attacks in Smart Grid Control Systems.

The research presented in this paper develops a new ensemble learner and workflow termed UNAD for unknown attack detection. Unlike previous ensemble learners for network intrusion/attack detection, UNAD integrates a heterogeneous set of standalone anomaly detection methods and improves upon their accuracy, precision and recall. Furthermore, UNAD is applied on recent data which contains more recent attack types. A problem with training models for unknown attacks is privacy issues. Therefore, publicly available synthetic benchmark datasets such as KDD Cup 99 [16], NSL-KDD[17], Kyoto 2006+[18], UNSW-NB15[19] and CICIDS2017 [5] can be used to mitigate privacy issues. The work presented in this paper uses CICIDS2017 as it contains more recent attack types.

UNAD BASE ANOMALY DETECTION METHOD SELECTION

In order to build the UNAD ensemble learner anomaly detection methods need to be selected as base learners. In total, 4 different kinds of anomaly detection methods that have previously been applied for similar applications to network attack detection (see RELATED WORK section) were considered. The considered techniques are One-Class SVM [20], iForest [3], LOF [2] and Elliptic Envelope [4]. The One-Class SVM method was ruled out early in the selection process since it is unsuitable for fast network flows due to its high computational demand [11]. The remaining three algorithms were experimentally optimised on the CICIDS2017 dataset and subsequently evaluated for their inclusion in the UNAD ensemble.

Experimental Setup

Evaluation Metrics

The metrics used to evaluate base learners are precision, recall and F1-Score. In UNAD, precision is equivalent to the portion of true positive attacks of all detections and recall is equivalent to the portion of attacks detected from all attacks present in the network flow. A high precision is equally important as detecting the majority of attacks. This is because false positive attack detections may trigger expensive actions to counter a non-existing threat. Since precision and recall are both equally important in this application, the base learners have been selected based on the F1-Score, which is the harmonic mean between precision and recall. An alternative measure to use instead of F1-Score could have been ROC_AUC; however, ROC_AUC measure is more reliable on balanced data and the ratio of benign data to data comprising attacks is 3:1 in the test set and validation set.

Dataset and Pre-Processing

For evaluating the algorithms, CICIDS2017 [5] dataset is used. CICIDS2017 is a publicly available benchmark dataset generated by the Canadian Institute for Cybersecurity, it covers five day, consists of 84 features, about 3 million data instances and covers 14 attack types including newer types that emerged in recent years. For generating the dataset [5] a complete network topology was created including Modem, Firewall, Switches, Routers. In addition, nodes in the network comprised various operating systems such as Windows, Ubuntu and Mac, all using commonly available protocols such as HTTP, HTTPS, FTP, SSH and email protocols. Table I summarises the number attacks per type and benign data instances in CICIDS2017.

TABLE I: CICIDS2017 overall traffic type distribution

Traffic Type	Count
Benign	2,358,036
DoS Hulk	231,073
Port Scan	158,930
DDoS	41,835
DoS GoldenEye	10,293
FTP Patator	7,938
SSH Patator	5,897
DoS SlowLoris	5,796
DoS SlowHTTPTest	5,499
Botnet	1,966
Web Attack: Brute Force	1,507
Web Attack: XSS	625
Infiltration	36
Web Attack: SQL Injection	21
HeartBleed	11
Total	**2,829,463**

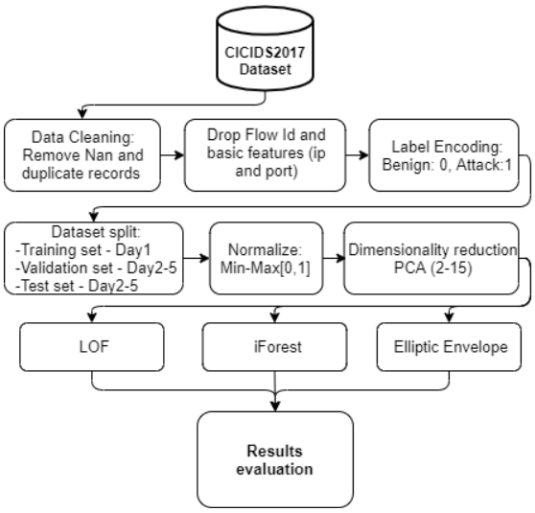


Fig. 1. Experimental workflow

All experiments were implemented in Python 3.6 using Google Colaboratory. The dataset was pre-processed before application of the anomaly detection algorithms. The pre-processing workflow is depicted in Figure 1. The first step was data cleaning which comprises removing missing and NaN values, as the used algorithms are designed for numerical data only. Moreover, duplicated records were removed to maintain data quality and avoid biased results. This step is

followed by dropping out some features that could affect the model's performance; for instance, ID features were removed as they do not have discriminatory value with respect to attacks. Next features containing IP addresses were also removed as attackers often spoof their email addresses to avoid IP filtering systems [8]. Finally, features representing port information were removed as they cause models to overfit towards socket information [21]. Next, the categorical text in the Label feature was converted to numeric form. Hence, the label for all attack types were converted to 1 and for benign instances to 0. This is because anomaly detection methods are essentially binary classification methods since they distinguish normal data (i.e. benign) versus anomalies (i.e. attacks). After pre-processing, the dataset was split into training, validation, and test sets. Assuming that there is no prior knowledge about the network attacks, the models will be trained only on benign flow. Thus, data from the first day was used to train the model which comprises 529,445 normal data instances (about 19% of the entire dataset). The remaining four-day dataset, which contains 2,298,225 data instances of both attacks and benign flow, were split for validation and testing (50% each). All data instances were normalised between 0-1 using min-max scaling to reduce inductive bias while keeping the shape of the original data distribution. For the experiments, Principal Component Analysis (PCA) is used. PCA has been applied in the Intrusion Detection area since it only requires a few parameters of the principal components to be managed for future detections and most importantly, the statistics can be estimated in a short amount of time during the detection stage, which enables real-time usage of PCA [22], [23].

Evaluation of Anomaly Detection Algorithms as Base Learners for UNAD

The anomaly detection models were learned from the training data (comprising only benign network flow), and the validation data (including all types of attacks) was used to find the best combinations of hyperparameter to maximise F1-Score. For hyperparameter tuning, various Principal Components (PCs) were considered (2-15 PCs) to reduce the data's dimensionality.

Local Outlier Factor (LOF)

LOF detects local outliers by comparing the local density of an object to its adjacent neighbours. LOF considers an object as an outlier if the average of the local reachability density of that object is lower than the local reachability density of its adjacent neighbours [2]. LOF's main advantage is detecting local and neighbouring outliers to data instances in very large datasets with heterogeneous densities [24], [25]. Therefore, for massive network traffic, LOF is expected to play a significant role in detecting attacks. Accordingly, LOF is evaluated here as a potential part of the proposed ensemble-based UNAD. The LOF module from scikit-learn [26] was used. The hyperparameters are contamination and n_neighbours. Contamination is the proportion of the outliers expected in the dataset ranging from 0 to 0.5. We assumed no knowledge about the proportion of outliers constituting non-attacks in the training data. The hyperparameters were tuned using various combinations of values for the contamination value. It was tuned from 0.01 to 0.5 in steps of 0.01. The value of n_neighbours was selected within a range of 5 to 50 in steps of 5. Once the hyperparameters were optimised and the best combination was determined, they were applied to the test set for every number of PCs ranging between 2-15.

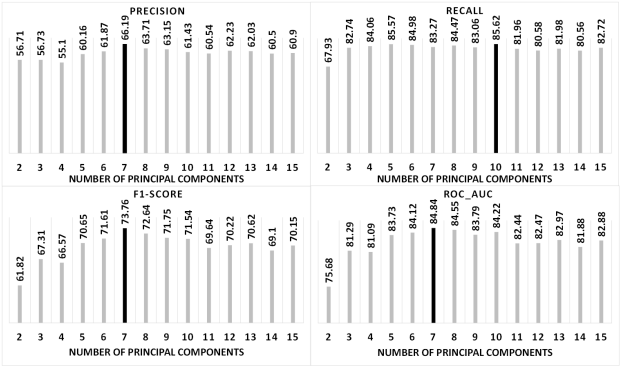


Fig. 2. Summary of Experimental Results expressed in percentages for the LOF based workflow

Fig. 2 shows the best results for each number of PCs used. For each number of PCs, always the best setting of contamination and the number of nearest neighbours is displayed. The figure shows that the highest precision and F1-Score was achieved for 7 PCs (with contamination parameter 0.07 and 30 neighbours), while the highest recall was observed for 10 PCs (with contamination parameter 0.08 and 35 neighbours). Based on the F1-Score, the optimal number of PCs for LOF is 7.

Isolation Forest

iForest consists of a random trees forest that keeps portioning all instances until they are fully separated. Moreover, iForest assumes that anomalies are expected to be split in early partitioning; therefore, instances with shorter path lengths are very likely to be anomalies [27]. iForest provides low linear time-complexity with a low memory requirement, making it ideal for detecting network attacks in a fast and timely manner. Furthermore, iForest can deal with high dimensional data with unrelated attributes [27]. Hence, making it perfect to be integrated in the proposed UNAD ensemble. iForest implementation from scikit-learn [26] was used. The hyperparameters here are contamination factor, n_estimators (number of trees) and max_samples. The contamination parameter is the same as for LOF. We assumed no knowledge about the proportion of outliers constituting non-attacks in the training data. The hyperparameters were optimised using various combinations of values for the contamination value. It was tuned from 0.01 to 0.5 in steps of 0.01. The number of n_estimators was selected from 50 to 450 in steps of 50.

Regarding the max_samples parameter, which selects the portion of the training data for each base estimator, a proportion settings of 25%, 50%, 75% and 100% were used in addition to the default setting of 256 samples. Concerning other parameters, max_features parameter which controls the number of features to be extracted from the dataset to train each estimator [26], it was set to its default values (1.0) to draw all features to train the estimators, and the random_state parameter was set to a fixed number (42) for results reproducibility. Once the hyperparameters were optimised and the best combination was determined, they were applied to the test set for every number of PCs ranging between 2-15.

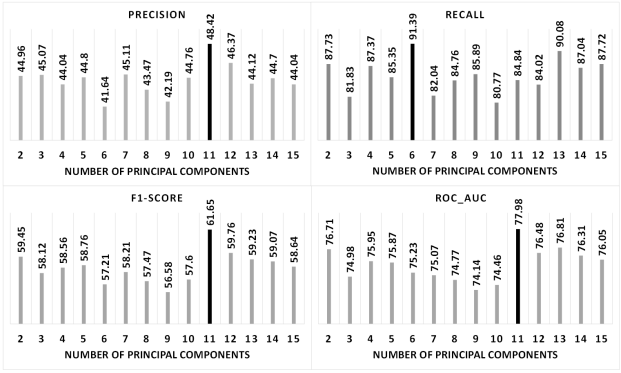


Fig. 3. Summary of Experimental Results expressed in percentages for the iForest based workflow

Fig. 3 shows the best results of iForest for each number of PCs used. For each number of PCs, always the best stetting of contamination and the number of nearest neighbours are displayed. The figure shows that the highest precision and F1-Score was obtained using 11 PCs (with contamination parameter of 0.24, 400 estimators and 25% max_samples), while the highest recall was observed using 6 PCs (with contamination parameter of 0.43, 200 estimators and default setting (256) max_samples). Based on the F1-Score the optimal number of PCs for iForest was at 11.

Elliptic Envelope

Elliptic Envelope detects outliers on multivariate Gaussian distributed datasets. Elliptic Envelope creates and fits an ellipse around the centre of a group of data instances using the Minimum Covariance Determinant. Hence, any data instance that is outside the ellipse is considered an outlier [4]. As the method was developed for Gaussian distributed datasets, it may not perform well on data streams, because the distribution a data stream can change over time due to concept drift. However, since the method has a low computational complexity, and is readily available in scikit-learn [26] it has been evaluated as a potential base learner for UNAD. The Elliptic Envelope contamination hyperparameter is the same as for LOF and iForest. Again, we assumed no knowledge about the proportion of outliers constituting non-attacks in the training data. The contamination parameter value was set ranging from 0.01 to 0.5 in steps of 0.01. Once the contamination parameter was optimised and its best value determined, it was applied to the test set for every number of PCs ranging between 2-15.

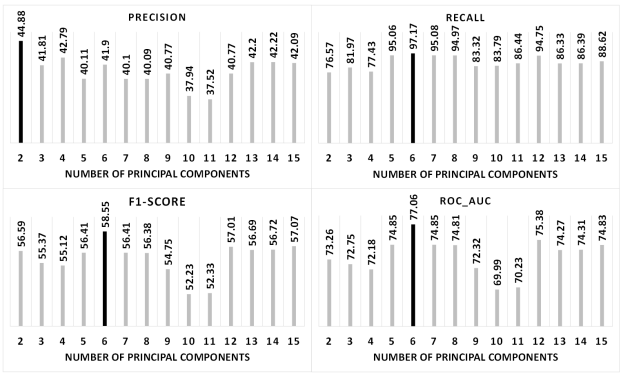


Fig. 4. Summary of Experimental Results expressed in percentages for the Elliptic Envelope based workflow

Fig. 4 depicts the Elliptic Envelope results for each number of PCs used. The figure shows that the highest recall and F1-Score was seen for 6 PCs (with contamination parameter of 0.44), while the highest precision was observed also for 6 PCs (with contamination parameter of 0.29). Based on the F1-Score and assuming equal importance of precision and recall, Elliptic Envelope's best setting was thus at 6 PCs, with contamination parameter of 0.44.

Conjectures and Selection of UNAD Base Learner Types

Although the anomaly detector candidates have been optimised with F1-Score as a target, ROC_AUC was included in the evaluation metrics as well since it is frequently used in anomaly detection literature. Interestingly in all cases using ROC_AUC rather than F1-Score would have lead the same outcome.

With respect to base anomaly detector selection for UNAD, LOF and iForest have been chosen. The reason for choosing LOF is that it achieves a relatively good F1-Score at 7 PCs on its own with 74% and a relatively high recall with 83% for 7 PCS. The precision of LOF is moderate with 66% at 7 PCs. The metrics for iForest are similar, but a bit more extreme. F1-Score is moderate at 61% for 11 PCs. the recall is high at 85% using 11 PCs, yet precision is relatively low with 48%, meaning that about half the anomaly detections are false alarms. Elliptic Envelope achieves the lowest F1-Score of all anomaly detectors with 59% at 6 PCs and even lower precision than iForest at 6 PCs with 42%. It has the highest recall though, with 97%. Since Elliptic Envelope achieves a very low precision, the technique is likely to be counterproductive in the UNAD ensemble and hence is excluded.

UNSUPERVISED ENSEMBLE LEARNER ARCHITECTURE FOR UNKNOWN ATTACK DETECTION

Based on the preliminary results discussed in the previous section, the UNAD ensemble was developed using of iForest and LOF as base learners, excluding Elliptic Envelop. The UNAD ensemble is depicted in Figure 5.

The dataset is cleaned the same way as described in the previous section, normalised and then two versions of the dataset are produced each projected on the best number of PCs for LOF (7 PCs) and iForest (11 PCs) respectively as determined in the experiments outlined in the previous section. Diversity among each type of base learner is created through bagging. For each base learner, bagging is applied on the 529,445 benign data instances of day one. In order to improve the stability and predictive performance of a composite learner [28], bagging was first introduced by Breiman [29]. It involves random sampling of the data instances with replacement. Each data instance is randomly selected whether to be in the sample or not. The size of the sample is equal to that of the original number of data instances. This suggests that some training instances may appear more than once in the same sample set, and some may not be included at all.

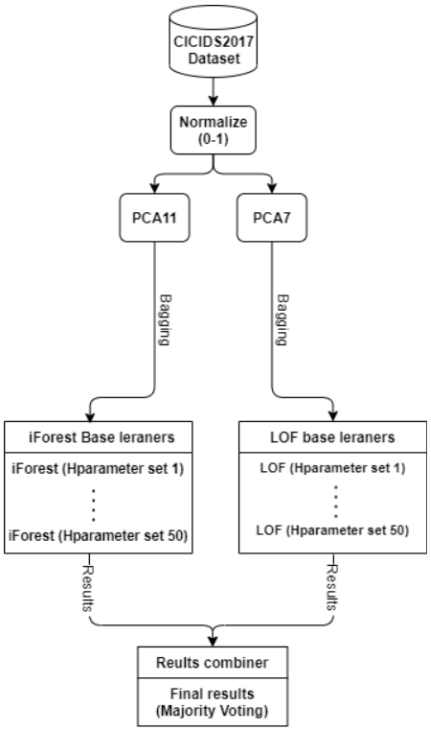


Fig. 5. Proposed UNAD workflow

UNAD induces 50 LOF and 50 iForest base learners, all of which are likely to be different since each has been induced on a separate bootstrap sample. UNAD further diversifies the base learners by randomly choosing a set of parameter values for the number of Nearest Neighbours and a contamination factor. Here parameter values from the top 3 best performing instances of LOF for 7 PCs were considered. These are:

- **Nearest Neighbours**: 25, 30 and 40
- **Contamination**: 0.06, 0.07 and 0.08

Concerning the iForest base learners, similar to the LOF base learners, UNAD chooses randomly the best parameter values from the iForest preliminary experiments with 11 PCs. These are in particular:

- **Number of estimators**: 150, 350 and 400.

The contamination parameter and the max samples parameters for iForest were identical for all top three instances in the preliminary experiments with 11 PCs. Hence, UNAD uses contamination 0.24 and 25% max samples for all iForest instances.

To detect anomalies, UNAD uses a majority voting scheme of all 100 base anomaly detector instances. There is an equal vote per base learner instance and per classification (benign or attack). Please note that due to equal voting and even number of base learners tie breaks are possible. The same number of base learners for LOF and iForest has been chosen to avoid bias towards one type of base learner, hence there is an even number of base learners. Currently, tie breaks are analysed by a human analyst, since they represent an uncertainty of the system. In the *ONGOING and FUTURE WORK* Section we consider reducing the number of tie breaks through a weighted voting mechanism.

EXPERIMENTAL EVALUATION

The UNAD learner was applied on the CICIDS2017 dataset as a case study. The data is pre-processed as already described in Section *UNAD BASE ANOMALY DETECTION METHOD SELECTION*. The dataset attack types and benign distribution is highlighted in Table I. Data from day one of the network flow was used as training data. These 529,445 instances comprised only benign cases. The test set comprising 1,149,112 instances was used to evaluate UNAD. The test set comprised instances with all attack types and also benign data instances.

TABLE II: LOF, iForest and UNAD results comparison in %

Measure(%) \ Method	**LOF**	**iForest**	**UNAD**
Accuracy	86	74	87
Precision	66	48	71
Recall	83	85	80
F1-Score	74	61	75
ROC_AUC	85	78	85

Table II depicts the EXPERIMENTAL results of UNAD compared with the standalone LOF and iForest results. Although the UNAD's recall was slightly lower than in standalone LOF and iForest algorithms, the precision was considerably improved and also there is some improvement of the F1-Score. The ROC_AUC results are the same compared with the best standalone competitor, and accuracy has slightly improved. However, accuracy and ROC_AUC are not considered a good evaluation metric since the data is imbalanced 3:1 in favour of benign data. Thus, F1-Score is considered a suitable metric, since it describes how well attacks have been detected in terms of true positive detections

and portion of overall attacks being detected. It can be seen that F1-Score has improved due to a considerable improvement of precision.

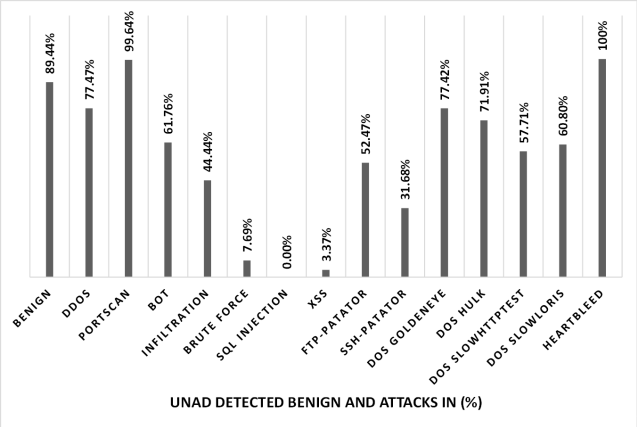


Fig. 6. Summary of UNAD detected benign and attacks

Figure 6 illustrates the percentage of identified benign cases and detected attacks using the UNAD system. UNAD was able to detect all the heartbleed attacks and almost all the portscan attacks (99.64%). UNAD was also able to identify 89.44% of the benign flow. DDoS and DoS detection rate were between 77.47% and 71.91% expect for DoS Slowloris and DoS Slowhttptest, which were 60.8% and 57.71% respectively. Attacks under the Web Attack category were the least well detected attacks with 7.69% for Brute Force, 3.37% for XSS and none of the SQL Injection attacks were detected. One can see that although there is overall a high recall / F1-Score the proportion of identified attacks varies by a large amount from attack type to attack type, yet all attacks, except Injections, are represented in the positive attack detections. However, considering the precondition that UNAD has never seen any of the attack types in the test set, it performs relatively well finding most attacks with a high precision and also almost all attack types are represented. In the *ONGOING WORK* section an approach to improve upon the recall of some attack types is briefly discussed.

TABLE III: Traffic type instances abstained from detection

Type	Count	Percentage (%)
BENIGN	151663	17.4
DDoS	3272	5.1
FTP-Patator	1546	38.9
DoS Slowhttptest	1160	42.1
DoS slowloris	868	29.9
DoS Hulk	824	0.7
DoS GoldenEye	789	15.3
Web Attack: Brute Force	386	51.2
SSH-Patator	671	22.8
PortScan	252	0.3
Web Attack: XSS	143	43.9
Bot	12	1.2
Infiltration	8	44.4
Web Attack: Sql Injection	3	30
Total	**164597**	**14.3**

Since UNAD abstains from detection when uncertain, there are also 161,597 test instances for which UNAD flagged that it was uncertain. How these traffic instances are composed is depicted in Table III. In the next section an approach to mitigate abstaining is briefly discussed.

ONGOING WORK

It was observed in the experimental evaluation that although UNAD did perform with a high F1-Score and precision on detected attacks, two limitations surfaced: (1) a low proportion of some attack types were detected and (2) abstaining from classifying some test instances.

With respect to limitation (1), a supervised adaptive component alongside UNAD is currently being developed that augments UNAD by training on new attack types previously detected by UNAD. Thus, once a new threat has been identified UNAD actively tries to further improve the detection of this particular type of attack. With respect to (2), a weighted majority voting is being considered since it is expected to improve UNAD's performance in general and lower the risk of tie breaks. For this a detection score is currently being developed (calculated on the *out of bag* sample from the bagging procedure). This detection score will be used to weight votes of UNAD base learners and thus reduce the possibility of tie breaks and further improve F1-Score. Tie-breaks resulting in abstaining from detection attempts may still occur; however, a lower number of abstained instances is expected to be more feasible to be examined manually by human analysts.

In addition, ongoing work also includes investigating why SQL injection are not well detected and which type of attacks are causing the ensemble's lower recall.

CONCLUSIONS

The paper discussed the need for unsupervised machine learning techniques to detect network attacks, because new types of network attacks constantly emerge. However, if an attack-type is and previously unknown, a supervised model is generally not capable of detecting such an attack sufficiently. Hence, this paper explores experimentally various anomaly detection methods for their detection capabilities of recent unknown network attacks. Based on this experimental evaluation the authors proposed a heterogeneous ensemble-based Unknown Network Attack Detection (UNAD) system which is composed of some of the evaluated anomaly detection methods, in order to improve precision and recall of unknown attack detection compared with standalone anomaly detection methods. UNAD is evaluated on the CICIDS2017 dataset, which does not pose any privacy issues and comprises recent attack types. The ensemble achieved a high precision and F1-Score and generally outperformed its standalone base anomaly detectors. Ongoing and future work comprise an improved voting strategy for base learners to further improve UNAD's performance and reduce tie breaks. Also an augmentation of UNAD is considered which provides a simultaneously running adaptive parallel supervised learner, which trained / adapted if a new, previously unknown attack, has been identified by UNAD.

References

[1] W. Chen, F. Kong, F. Mei, G. Yuan, and B. Li, "A novel unsupervised anomaly detection approach for intrusion detection system," in *ieee 3rd international conference on big data security on cloud (bigdatasecurity), ieee international conference on high performance and smart computing (hpsc), and ieee international conference on intelligent data and security (ids)*, pp. 69–73, 2017.

[2] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "LOF: Identifying density-based local outliers," in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, SIGMOD '00, (New York, NY, USA), p. 93–104, ACM, 2000.

[3] F. T. Liu, K. M. Ting, and Z. Zhou, "Isolation forest," in *2008 Eighth IEEE International Conference on Data Mining*, pp. 413–422, 2008.

[4] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, p. 212, Aug 1999. 3.

[5] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization.," in *4th International Conference on Information Systems Security and Privacy (ICISSP)*, pp. 108–116, 2018.

[6] O. Sagi and L. Rokach, "Ensemble learning: A survey," *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018.

[7] H. Zhang, S. Dai, Y. Li, and W. Zhang, "Real-time distributed-random-forest-based network intrusion detection system using apache spark," in *2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC)*, pp. 1–7, 2018.

[8] C. B. Freas, R. W. Harrison, and Y. Long, "High performance attack estimation in large-scale network flows," in *2018 IEEE International Conference on Big Data (Big Data)*, pp. 5014–5020, 2018.

[9] Y. Zhou, G. Cheng, S. Jiang, and M. Dai, "Building an efficient intrusion detection system based on feature selection and ensemble classifier," *Computer Networks*, vol. 174, p. 107247, 2020.

[10] M. Zhang, B. Xu, and J. Gong, "An anomaly detection model based on one-class svm to detect network intrusions," in *2015 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, pp. 102–107, 2015.

[11] I. Razzak, K. Zafar, M. Imran, and G. Xu, "Randomized nonlinear one-class support vector machines with bounded loss function to detect of outliers for large scale iot data," *Future Generation Computer Systems*, vol. 112, pp. 715 – 723, 2020.

[12] L. Sun, S. Versteeg, S. Boztas, and A. Rao, "Detecting anomalous user behavior using an extended isolation forest algorithm: An enterprise case study," 2016.

[13] D. Haidar and M. M. Gaber, "Adaptive one-class ensemble-based anomaly detection: an application to insider threats," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–9, IEEE, 2018.

[14] A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava, "A comparative study of anomaly detection schemes in network intrusion detection," in *Proceedings of the 2003 SIAM international conference on data mining*, pp. 25–36, SIAM, 2003.

[15] M. Ashrafuzzaman, S. Das, A. A. Jillepalli, Y. Chakhchoukh, and F. T. Sheldon, "Elliptic envelope based detection of stealthy false data injection attacks in smart grid control systems," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1131–1137, 2020.

[16] KDD99, "The UCI KDD Archive." http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. Accessed: 18.06.2020.

[17] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the kdd cup 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pp. 1–6, 2009.

[18] J. Song, H. Takakura, Y. Okabe, M. Eto, D. Inoue, and K. Nakao, "Statistical analysis of honeypot data and building of kyoto 2006+ dataset for nids evaluation," in *Proceedings of the First Workshop on Building Analysis Datasets and Gathering Experience Returns for Security*, BADGERS '11, (New York, NY, USA), p. 29–36, ACM, 2011.

[19] N. Moustafa and J. Slay, "Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set)," in *2015 Military Communications and Information Systems Conference (MilCIS)*, pp. 1–6, 2015.

[20] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Platt, "Support vector method for novelty detection," in *Proceedings of the 12th International Conference on Neural Information Processing Systems*, NIPS'99, (Cambridge, MA, USA), p. 582–588, MIT Press, 1999.

[21] O. Faker and E. Dogdu, "Intrusion detection using big data and deep learning techniques," in *Proceedings of the 2019 ACM Southeast Conference*, pp. 86–93, 2019.

[22] K. S. M. Shyu, S. Chen and L. Chang, "A novel anomaly detection scheme based on principal component classifier," in *Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop, in conjunction with the Third IEEE International Conference on Data Mining (ICDM03)*, p. 172–179, 2003.

[23] S. Almotairi, A. Clark, G. Mohay, and J. Zimmermann, "A technique for detecting new attacks in low-interaction honeypot traffic," in *2009 Fourth International Conference on Internet Monitoring and Protection*, pp. 7–13, 2009.

[24] D. Pokrajac, A. Lazarevic, and L. J. Latecki, "Incremental local outlier detection for data streams," in *2007 IEEE Symposium on Computational Intelligence and Data Mining*, pp. 504–515, 2007.

[25] M. Salehi, C. Leckie, J. C. Bezdek, T. Vaithianathan, and X. Zhang, "Fast memory efficient local outlier detection in data streams," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 12, pp. 3246–3260, 2016.

[26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[27] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation-based anomaly detection," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 6, no. 1, pp. 1–39, 2012.

[28] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1, pp. 1–39, 2010.

[29] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, p. 5–32, Oct. 2001.

SAIF ALZUBI is a PhD student in Computer Science at the University of Reading, UK. His main research interests are in Machine Learning and Data Mining. Mr Alzubi worked as an Information System Developer at the e-learning centre, and as a senior database programmer at the IT centre, University of Bahrain, Bahrain. He obtained his MSc degree in Computer Science in 2018 from Coventry University, UK and his BSc degree in Computer Science in 2009 from University of Jordan, Jordan.

FREDERIC STAHL is Senior Researcher at the German Research Center for Artificial Intelligence (DFKI). He has been working in the field of Data Mining for more than 15 years. His particular research interests are in (i) developing scalable algorithms for building adaptive models for real-time streaming data and (ii) developing scalable parallel Data Mining algorithms and workflows for Big Data applications. In previous appointments Frederic worked as Associate Professor at the University of Reading, UK, as Lecturer at Bournemouth University, UK and as Senior Research Associate at the University of Portsmouth, UK. He obtained his PhD in 2010 from the University of Portsmouth, UK and has published over 65 articles in peer-reviewed conferences and journals.

MOHAMED MEDHAT GABER Mohamed Gaber is a Professor in Data Analytics at Birmingham City University, and currently seconded as the Dean of the Faculty of Computer Science and Engineering at Galala University. He has published over 200 papers, co-authored 3 monograph-style books, and edited/co-edited 7 books on Artificial Intelligence. His work has attracted nearly six thousand citations, with an h-index of 40. According to the latest study conducted by Stanford University and Elsevier, and released in 2020, Mohamed is among the top 2% of the most cited scientists worldwide. Mohamed's research interests span many areas of Artificial Intelligence including, but not limited to: (1) ensemble learning, (2) learning from data streams, (3) medical image analysis, (4) natural language processing, (5) time series classification, and (6) deep learning.

DATA STREAM HARMONIZATION FOR HETEROGENEOUS WORKFLOWS

Eleftherios Bandis
Nikolaos Polatidis
Maria Diapouli
Stelios Kapetanakis
University of Brighton
Moulsecoomb Campus, Brighton NB2 4GJ, UK
{e.bandis, n.polatidis, m.diapouli, s.kapetanakis}@brighton.ac.uk

KEYWORDS
Data stream workflows, Graph Reasoning, Monitoring

ABSTRACT

Transport infrastructure relies heavily on extended multi sensor networks and data streams to support its advanced real time monitoring and decision making. All relevant stakeholders are highly concerned on how travel patterns, infrastructure capacity and other internal / external factors (such as weather) affect, deteriorate or improve performance. Usually new network infrastructure can be remarkably expensive to build thus the focus is constantly in improving existing workflows, reduce overheads and enforce lean processes. We propose suitable graph-based workflow monitoring methods for developing efficient performance measures for the rail industry using extensive business process workflow pattern analysis based on Case-based Reasoning (CBR) combined with standard Data Mining methods. The approach focuses on both data preparation, cleaning and workflow integration of real network data. Preliminary results of this work are promising since workflow integration seems efficient against data complexity and domain peculiarities as well as scale on demand whilst demonstrating efficient accuracy. A number of modelling experiments are presented, that show that the approach proposed here can provide a sound basis for the effective and useful analysis of operational sensor data from train Journeys.

INTRODUCTION

The modernisation of Rail industry has led to increasing usage of computer systems for logistics, tactical, planning, performance and maintenance reasons. Rail industry has experienced substantial growth over the last decade in terms of operational method advancement (wayside detectors, wheel profile monitors, extended sensor network), processes, software and hardware equipment (Rail Defect Test Facility, Asset Health Strategic Initiative, and others). These systems generate millions of records per day that are constantly monitored, enhanced and analysed with the aim to improve industry capability, reduce cost and ultimately increase customer satisfaction.

Most rail operations, such as scheduled train services can be treated as business workflows, since they comprise event trails of spatio-temporal data. Techniques developed and tested for monitoring workflow operations can also be used in the context of live train journey auditing and performance measurement.

An example of such systems that fit well workflow orchestration and choreography is Remote Condition Monitoring (RCM) systems. RCM comprise multi-sensor systems per any running vehicle that can offer the full picture of a how a locomotive performs within a pre-determined time span (minute, hour, day, etc.). Its captured information is very low level and can reproduce a train journey with all relevant mechanical data. RCM is primarily used for technical -incident- monitoring, however it has also been observed as an accurate indicator of performance malfunctioning over a period of time.

Rail networks are prone to delays since order has to be maintened with emphasis to driver and passenger safety, cost and performance. Workflow techniques based on data streams and process mining can be increbibly valuable to Train Operator Companies (TOCs) to understand bottlenecks, increase capacity and minimize cost throughout the networks. This paper presents a data harmonization approach for spatio-temporal data using graph representation and general time theory (Ma, 1994) which enable data harmonization across multi-provenance sensor streams. This work, although quite recent in inception, has been proven reliable for heavy volume data (Agorgianitis, 2016)systems and effective in real time TOC data. This paper is structured as follows: Literature section witll refer to state of the art work in the field, Methodology will present the rationale and foundation principles of this work, Evaluation will presents real life data integrations with TOC Data. Finally, Conclusion will describe results as well as next steps for this work.

LITERATURE

Modern organisations use Business Process Workflows (BPW) to coordinate their processes, tasks, roles and manage resources with the aim to improve efficiency, efficacy and profitability. Workflows can automate processes, make them more agile and increase monitoring for obscure, erroneous or complex events to

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

company managers to increase productivity (Workflow Management Coalition, 2021; BPMI, 2021). BPW management differs across organisations. The size, sector and strategic orientation of an organization plays a key role on how they adopt, analyse and practice BPWs (Van der Aalst, 2003). A common taxonomy includes the phases of: Design, Implementation, Enactment, Monitoring and Evaluation as the workflow life cycle in BPW management (Muehlen, 2004). Among those the Monitoring phase enables the supervising of business processes in terms of management (e.g. performance, accuracy) and organization (e.g. utilization of resources, length of activities etc.) (Reijers, 2003). Monitoring is key operation informing process managers and workflow designers necessary adjustments to improve their processes.

In the case of using Business process Modelling techniques to monitor train jour-ney operation there is a need to integrate various data from different rail systems, as well as the timetable to provide a detailed insight into real train journeys. RCM data are key to provide the basis of this analysis, but there is a considerable challenge to associate, workflow execution trails with the expected business process instances (i.e. timetable). This has proven to be a complicated task as several problems exist within the Railway data collection systems. For example:

- RCM systems are independent enough, installed on several trains at contrasting times. They generate data that denote a workflow process execution, however, there is no available information (linkage) between monitored workflow traces and their corresponding workflow on a seasonal timetable.
- Data monitoring has several phases. Firstly, telemetric sensors are used to gather data as "low level events". Then data is filtered by a processing system to produce workflow processes. Finally, the extracted workflows are stored on persistence lay-ers of variant formats. Each phase represents a single entity since it is created at various times and by different architectures. Consequently, the data transformation along each phase allow margin for error which leads to partially inconsistent, in-complete and ultimately faulty data. Through data analysis which has been con-ducted on real RCM datasets we found that such percentage can vary but it ultimately can affect crucial attributes making workflow generation and workflow alignment to business process extremely difficult.
- Transport industry has many similar processes. For instance, the same route might run multiple times within a few minutes interval. It is difficult to distinguish identical processes since most of their attributes having significant similarity.
- RCM data can contain missing and erroneous values -due to different clocks, analogue sensors and error-prone data transmission systems and areas (such as tunnels)-.
- TOCs have several fleets of similar trains that may employ several dif-ferent RCM systems. Several processes can be stored in different da-tasets which make workflow operations substantially complex.
- Data format can follow several popular or bespoke formats, hardening a universal workflow monitoring approach.

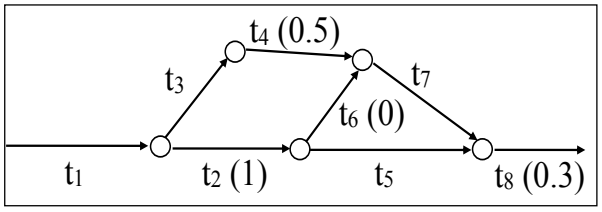
Workflow experts can use various methods to evaluate their processes, however, large or extended volumes of data can make the analysis of event logs extremely difficult. Process Mining (PM) is the technique used to extract knowledge and insights by discovering and analysing processes from event logs (Van der Aalst, 2011). By applying process mining, domain experts can use the derived information as feedback to design new processes or revise and enact predefined ones. In the literature, several algorithmic techniques have been introduced to solve the process mining problem. Algorithms like Alpha miner and alpha+ have been used extensively but other heuristics, genetic and fuzzy algorithms have also been applied (Tiwari, 2008). Each algorithm has its limitations on a different aspect of the process discovery such as fitness, simplicity and precision, and they may be unfit to areas where uncertainty, inconsistency and fuzziness is present. In such cases a CBR approach (Alshammari, 2017) may be more appropriate. CBR has been proven effective in monitoring business process workflow instances under uncertainty (Kapetanakis, 2009; 2010; 2011; 2012; 2013; 2014) in different interdisciplinary domains (Adedoyin, 2017), (Al Murayziq 2015, 2017), (Amin, 2019, 2020), (Ekpenyong, 2019), (Lansley,2019), (O'Connor, 2018) by retrieving similar solutions for similar problems.

RESEARCH METHODOLOGY

Our workflow data follow a sequential temporal and spatial pattern since they represent a variety of activities over time. Information about workflows can be encoded as events (points in time) or states (time intervals). In order to combine the two representation primitives and retain the full information and its provenance, there is a need for a formal underlying theory and representation that captures both temporal information and temporal relations (order, concurrency etc). To represent effectively workflows and their sequence and relationships in a formal way we use the General Time Theory (GTT) (Bandis, 2017; 2018), (Petridis, 2014). The general time theory takes both points and intervals as primitive. It consists of a triad (T, Meets, Dur), where:

- T is a non-empty set of time elements;
- Meets is a binary order relation over T;
- Dur is a function from T to R0+, the set of non-negative real numbers.

A time element t is called an interval if Dur(t) > 0; otherwise, t is called a point



Graphical representation of a log temporal inference using the GTT

In a graph representation each node represents a station whereas any edge represents the duration from station A to station B. A GTT workflow representation allows for a unified log interpretation which in conjunction with the multi-level similarity representation presents a foundation for adequate CBR workflow cases (Kapetanakis, 2014).

REPRESENTATION

A workflow process consists of multiple activities. Activities involve tasks such as "start of a journey", "departure from a station", "arrive on a station" or "end of a journey". The tasks contain multi-perspective information such as:

1. Time-related information: The start and the end of each activity is marked with a timestamp. The duration of an activity is also given.
2. Location: The station of which the activity takes place
3. Relationships: One activity holds which activity follows as well as the time duration between them

General information about the workflow is also available:

1. The total duration of all activities
2. The train unit responsible to undertake all the workflow activities
3. The day of the week the workflow took place
4. The workflow start and end time

Workflows are represented as GTT event-duration graphs with spatial information as node-specific tags. Every node can be represented as: {StationName$_q$, StopDuration$_q$, NextStation$_q$, TimeUntilNextStation$_q$}

Similarity among graphs is represented using multi-level representation based on the workflow structure. This can be annotated as:

Level 1: Relevant timestamps from workflow data. For example, Let case 1, C_1 and case 2, C_2 as workflow representations and C_{1L}, $C_{2L'}$ their respective list of stations. For C_1 and C_2 if Start date is the same (Binary equal) && Start time relies within γ mins fluctuation && C_{1L} is like $C_{2L'}$ based on an μ string threshold.

$$distance(C_1,C_2) = |StartTimeC_1 = StartTimeC_2 \leq \gamma| * w_1 + |EndTimeC_1 - EndTimeC_2 =< \gamma| * w_2 + |StationListC_1 - StationListC_2| * w_3$$ (equation 1)

Where w_1, w_2, w_3 are empirically (expert-based) derived domain constants and

$$w_1 + w_2 = w_3$$ (equation 2)

Upon successful relevance on similarity 1, a Level 2 similarity can be defined as:

p_1: create relationships => {[S_1, Dur(S_1), Dur(S_2), Meets S_2] ...}

(equation 3)

Where S_1 is a starting point, Dur(S_1) is the time spent on the station, Dur(S_2) the time till the next station, and Meets S_2 the station that follows. A Level 2 similarity is based on equation 3 quadruplets as:

$$distance(C_1,C_2) = |[S_1, Dur_{S1}, Dur_{S2}, S_2]C_1 - [sS_1, Dur_{S1}, Dur_{S2}, S_2]C_2| * w_1 + |StartDayOnlyC_1 = StartDayOnlyC_2| * w_2 + |UN_1 = UN_2| * w_3$$

(equation 4)

Where UN_1 and UN_2 are system identification numbers

EVALUATION

For the needs of evaluation we used data from 159000 trail records approximately over the period of ten months. Workflows were represented as graphs using GTT. Moving windows using level 1 and 2 similarities repsecitively, were used to combine together relevant workflows. Four types of datasets were used including:

1) RCM data from live train journeys
2) Performance data from planned / expected, already ran journeys
3) Timetabling data indicating planned, long-term planned and emergency routes across all networks
4) Spatio-temporal data for any assets (stations, signals, depots) and train location data available from sensors

GTT enabled workflow representation for all datasets starting from structured ones, like: Timetabling and Locations as well as free form ones: Performance and RCM. Level 1 and 2 similarities enabled workflow alignment and match of segments with complementary data provenance and information. Every performance journey was ranked with an indicator of delay which could be

1. Type A: No delay
2. Tybe B: Sub-threshold delay between 1-3'
3. Type C: Recorded Delay between 3-15'
4. Type D: Severe Delays of more than 15'

These classification scale was available just to one type of workflows and not the others. With the workflow unification, industry experts were able to see the journey classification as well as retrace back what happened on that specific case, see relevant information for the underlying family of services, routes as well as any available information on a daily basis. Based on the combined multiple provenance workflow data machine learning tehchniques were used to verify the accuracy of the system in numerical prediction e.g. given a specific trail of data can this be attributed to the right family of workflows and can it be classified accurately against delays of type A-D.

For the first part of the evaluation the aggregation results using GTT enabled graphs and level 1, 2 similarity were encouraging with 93.89% success rate.

Table 1 summarises the results in terms of successful vs. unsuccessful cases.

	Accurate Match	Total records
Workflow records	100%	159000
Matched successfully	93.89%	149282
Unsuccessful match	6.11%	9718

Table 1: Workflow match accuracy

Workflow matching had a high match ration, however still a high number of cases was not able to be connected dud to data inconcistencies, duplicate records and hardware peculiarities that required further processing and filtering. The results from this initial phase were treated as encouraging from industry stakeholders and requested the emphasis of the evaluation work to be placed on delay prediction given partial visibility of real time datasets. For this phase BPW mining techniques in workflow numerical prediction were used by applying generalized linear model, regression and a neural network classifier trained from existing workflow. Target was set as predicting whether a service will experience delay using early available data from the beginning of each route. A typical route can contain any number of stop between the range of 18 - 50 stations approximately. The first three nodes for each workflow graph where used as predictors for a combined workflow journey. For the needs of the evalution just week working days were selected as well as peak times where most delays take place usually.

	Generalised Linear Model	**Regression**	**ANN**
Min Error	-878	-1025	-476
Max Error	1754	1831	1907
Mean Absolute Error (MAE)	56	58	68
Standard Deviation	102	106	96
Linear Correlation	0.756	0.787	0.863
Occurrences	96,671	96,671	96,671

Table 2: Predicution results, journey times in seconds

As shown in Table 2, neural network predictors were shown most accurate in predicting delay. Results were interpreted positively from rail experts, however they expressed views for further workflow segmentation, special cases identification and filtering (for abnormal events) as well as the need for further explainability which will be the focus for further work.

CONCLUSION

This work presents a workflow harmonization approach in a real industrial environment. This work has been promising to domain experts since it is able to collate together workflows originating from different origins and present them under a common ground. There is substantial amount of improvement that can be applied in this field. Further work will focus explicitly on specialized workflow segmentation, algorithmic explanation and enhancement of the workflow auditing results. This approach seems generic and reusable to other domains, work which will be pursued in the future phases of this work.

REFERENCES

Adedoyin, A., Kapetanakis, S., Samakovitis, G., Petridis, M. (2017) Fraud Detection in Mobile Payment Transfer, In proceedings of the 22nd UK CBR workshop, Peterhouse, December 2017,(Ed M. Petridis), Brighton press, pp. 41-44

Agorgianitis, I., Petridis, M., Kapetanakis, S., Fish, A. (2016) Evaluating Distributed Methods for CBR Systems for Monitoring Business Process Workflows. In proceeding of ICCBR 2016, Workshop on Reasoning about time in CBR, Atlanta, GA, October 28-November 2, 2016, pp.122-131

Al Murayziq, T. S., Kapetanakis, S.,Petridis, M. (2015). Towards successful prediction of Dust Storms using Case-based Reasoning and Artificial Neural Networks. In proceedings of the 20th UK CBR workshop, Peterhouse, December 2015,(Ed M. Petridis), Brighton press, pp. 58-67

Al Murayziq, T., S., Kapetanakis, S., Alshammari, G., Petridis, M. (2017) Identifying and Predicting the Dust Events by Using Case Based Reasoning (CBR) , In Proceedings of the 22nd UK CBR workshop, Peterhouse, December 2017,(Ed M. Petridis), Brighton press, pp. 45-46

Alshammari, G., Jorro Aragoneses, J. L., Kapetanakis, S., Petridis, M., Recio-Garcia, J. A., Diaz-Agoudo, B. (2017) A hybrid CBR approach for the long tail problem in recommender systems. In proceedings of The International Conference in Case-based Reasoning (ICCBR 2017), pp. 35-45

Amin, K., Kapetanakis, S., Althoff, K., Dengel, A., Petridis, M. (2019) Building Knowledge Intensive Architectures for heterogeneous NLP workflows, In proceedings of the AI 2019

Amin, K., Kapetanakis, S., Polatidis, N., Althoff, K., Dengel, A. __(2020)__DeepKAF: A Heterogeneous CBR & Deep Learning Approach for NLP Prototyping, International Conference on Innovations in Intelligent Systems and Applications: INISTA 2020. IEEE

Bandis, E, Kapetanakis, S, Petridis, M, Fish, A. 2017. Effective Similarity Measures for Process Mining Using CBR on Rail Transport Industry, *in Proceedings of the 22nd UKCBR worksho*p, Cambridge UK

Bandis, E, Petridis, M, Kapetanakis, S,: Predictive Process Mining Using a Hybrid CBR Approach for the Rail Transport Industry, in RATIC 2018, *Proceedings of the 26th International Conference in Case Based Reasoning*, Stockholm, Sweden 9-12 July 2018

Business Process Management Initiative (BPMI): BPMN 1.1: OMG Specification, February 2008, http://www.bpmn.org/, accessed Feb 2021

Ekpenyong, F., Samakovitis, G., Kapetanakis, S., Petridis, M. (2019) An ensemble method: Case-Based Reasoning and the Inverse Problems in Investigating Financial Bubbles, in Proceedings of the International Conference on Cognitive Computing (ICCC 2019)

Kapetanakis, S., Petridis, M., Ma, J., Bacon, L. (2009). Workflow Monitoring and Diagnosis Using Case Based Reasoning on Incomplete Temporal Log Data. Proceedings of the Workshop on Uncertainty, Knowledge Discovery, and Similarity in Case Based Reasoning UKDS, in Workshop proceedings of the 8th International Conference on Case Based Reasoning, Seattle, USA, 2009

Kapetanakis, S., Petridis, Ma, J., Bacon, L. 2010. Providing explanations for the intelligent monitoring of business workflows using case-based reasoning. In: Roth-Berghofer, T., Tintarev, N., Leake, D. B., Bahls, D. (eds.) *Proceedings of the 5th International Work-shop on explanation-aware Computing Exact* (ECAI 2010), Lisbon, Portugal

Kapetanakis, S.,Petridis, M.,Knight, B.,Ma, J.,Bacon,L. 2010. A Case Based Reasoning Approach for the Monitoring of Business Workflows, *18th International Conference on Case-Based Reasoning,* ICCBR 2010, Alessandria, Italy, LNAI

Kapetanakis, S., Petridis, M., Ma, J., Knight, B., Bacon, L. (2011). Enhancing Similarity Measures and Context Provision for the Intelligent Monitoring of Business Processes in CBR-WIMS, In: Process-oriented Case-Based Reasoning workshop (PO-CBR), ICCBR2011

Kapetanakis, S., Samakovitis, G., Gunasekara, B., Petridis, M. (2012). 'Monitoring Financial Transaction Fraud with the use of Case-based Reasoning', Seventeenth UK Workshop on Case-Based Reasoning (UKCBR 2012), 11th December 2012, Cambridge, UKa

Kapetanakis, S., Samakovitis, G., Gunasekara, B., Petridis, M. (2013). The Use of Case-Based Reasoning for the Monitoring of Financial Fraud Transactions, Journal of Expert Update Vol. 13 (1) pp.75-83

Kapetanakis, S., Petridis, M. 2014. Evaluating a Case-Based Reasoning Architecture for the Intelligent Monitoring of Business Workflows, *in Successful Case-based Reasoning Applications-2*, S. Montani and L.C. Jain, Editors, Springer Berlin Heidelberg. p. 43-54

Lansley, M., Polatidis, N., Kapetanakis, S., Amin, K., Samakovitis, G., Petridis, M. (2019) Seen the villains: Detecting Social Engineering Attacks using Case-based Reasoning and Deep Learning, In Proceeding of the Deep Learning Workshop in Case-based Reasoning, ICCBR 2019

Ma, J., Knight, B. 1994. A General Temporal Theory, the Computer Journal, 37(2), 114-123

O' Connor, D., Kapetanakis, S., Samakovitis, G., Floyd, M., Ontañon, S., Petridis, M. (2018) Autonomous Swarm Agents using Case-based Reasoning In Proceedings of the Thirty-eighth SGAI International Conference on Artificial Intelligence, AI 2018, pp. 210-216

Petridis, M., Kapetanakis, S., Ma, J., Burlutskiy, N. (2014). Temporal Knowledge Representation for Case Based Reasoning Based on a Formal Theory of Time. In: Gundersen, O. E., Montani, S. (eds) Proceedings of RATIC: Reasoning about Time in CBR, ICCBR 2014, pp. 154-164, Springer, Heidelberg(2014)

Reijers, H.A. 2003. Design and Control of Workflow Processes: *Business Process Management for the Service Industry*. Springer, Heidelberg

Tiwari, A., Turner, C. J., & Majeed, B. 2008. A review of business process mining: State-of-the-art and future trends. *Business Process Management Journal*, 14(1),5–22

Van der Aalst, W.M.P., ter Hofstede, A.H.M., Weske, M. 2003. Business Process Management: A Survey. *In: van der Aalst, W.M.P., ter Hofstede, A.H.M., Weske, M. (eds.) BPM 2003. LNCS*, vol. 2678, pp. 1-12. Springer, Heidelberg

Van der Aalst. 2011. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag, Berlin

Workflow Management Coalition. Workflow management coalition glossary & terminology. http://www.wfmc.org/standards/docs/ TC1011_term_glossary_v3, 2021

Zur Muehlen, M. 2004. Workflow-Based Process Controlling: Foundation, Design and Application of Workflow-driven Process Information Systems. Logos

AUTHOR BIOGRAPHIES

Eleftherios Bandis has a BSc in Software Enginering and a PhD in Machine Learning from the Univesity of Brighton, UK. He worked for several years in the Transportation Industry as Data Engineer and Machine Learning Expert. His expertise resides in Real time systems, Spatiotemporal Workflows and Graph Theory. His e-mail address is : e.bandis@brighton.ac.uk

Nikolaos Polatidis has a BSc in Computer Science from Heriot-Watt University, an MSc in Internet Software Systems from the University of Birmingham, UK and a PhD from University of Macedonia. His background is in Artificial Intelligence, Machine Learning and Cyber Security. His e-mail address is : n.polatidis@brighton.ac.uk

Maria Diapouli has a BSc in Software Enginering and an MSc in Enterprise Systems from the Univesity of Greenwich, UK. Her background is in Advanced Databases and Distributed Systems. She has 10 years of experience in the Transportation and Online Marketing Industry. Her e-mail address is : m.diapouli@brighton.ac.uk

Stelios Kapetanakis has a PhD in Artificial Intelligence and an MBA in Knowledge and Innovation Management. He has been a Principal Lecturer in the University of Brighton as well as technical consultant for several startups in Europe , Australia and the US. His work focuses on machine learning solutions in enterprise environments. His e-mail address is : s.kapetanakis@brighton.ac.uk

Predicting Next Touch Point in a Customer Journey: a Use Case in Telecommunication

Marwan Hassani Stefan Habets
Department of Mathematics and Computer Science
Eindhoven University of Technology, The Netherlands
m.hassani@tue.nl s.habets@student.tue.nl

ABSTRACT

Customer journey analysis is rapidly increasing in popularity, as it is essential for companies to understand how their customers think and behave. Recent studies investigate how customers traverse their journeys and how they can be improved for the future. However, those researches only focus on improving the process for future customers by analyzing the historical data. This research focuses on helping the current customer immediately, by analyzing if it is possible to predict what the customer will do next and accordingly take proactive steps. We propose a model to predict the customer's next contact type (touch point). At first we will analyze the customer journey data by applying process mining techniques. We will use these insights then together with the historical data of accumulated customer journeys to train several classifiers. The winning of those classifiers, namely XGBoost, is used to perform a prediction on a customer's journey while the journey is still active. We show on three different real datasets coming from interactions between a telecommunication company and its customers that we always beat a baseline classifier thanks to our thorough pre-processing of the data.

I. Introduction

Nowadays, companies collect all sorts of data from their services and customers. Altogether, data-driven analysis has become more interesting for companies and the collected data is used by companies to analyze all of their products and services. In organizations, such as a telecommunication companies, the data is used to analyze the journey of a customer. Investigating and analyzing often reveals ways to potentially optimize services provided. This investigation is both in favor of the customer and the company, as the service is improved the customer will get a more tailored approach. For example, a customer that can install his newly received modem on his own or with the help of the website and does not have to call the service desk or worse, need a mechanic. Those customers save the company money because the service desk has less work or no mechanic has to be send, decreasing the overall expenses. Also the customer is happier since he does not have to wait in the phone queue of the service desk or wait at home to receive a mechanic, increasing the customer satisfaction.

One kind of such data is the customer journey, the customer journey represents the steps a customer takes with the company. Each step is called a touch point and is defined to be an interaction from the customer with the companies' products or services [3]. These customer journeys are mapped into a customer journey map (CJM) to perform analysis on. The customer journeys are very interesting to analyze since the company can see how customers actually behave. The company has an idea of how a certain process should work, but in practice this might not be the case at all. Reviewing the customer journeys gives insight in how customers follow certain flows. When an important step is missing from the process flow in a majority of the customer journeys then the company can investigate this matter and see why the customers do not perform those steps.

Before applying process mining to it, customer journey analysis was performed to improve processes in hindsight [10]. Historical data is checked and used to iteratively improve the customer experience when interacting with a certain process of the company. Therefore the company is always late in providing immediate support to the customer. When something goes "wrong", the mistake can later be found in the data to help improve the error for the future.

Knowing what a customer will do next, gives the company the ability to proactively provide support to the customer. Helping the customer proactively saves time and therefore also costs, as well as increases the customer satisfaction as he/she is helped faster. This leads to the following research question defined for this paper: Is it possible to predict a customer's next touch point by starting from the historical data from customer journeys in our use case?

This paper is structured as follows: Preliminaries and some existing solutions are discussed in Section II. In Sections III and IV an understanding of respectively the business and the data is shared. The data pre-processing and the applied models are discussed in Section V. The experimental results are discussed in Section VI. Finally, Section VII concludes this paper.

II. Preliminaries and existing solution

The data consists of customer journeys in the form of event logs. An event log $L = (Tr_1, Tr_2, \ldots, Tr_n)$ consists of a collection of n traces. A trace is a sequence of events $Tr_i \in E^*$, where E is a collection of events. An event $e_i = (c_j, a_k, t_i)$ needs to include an uniquely identifiable customer journey c_j, an activity a_k from

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

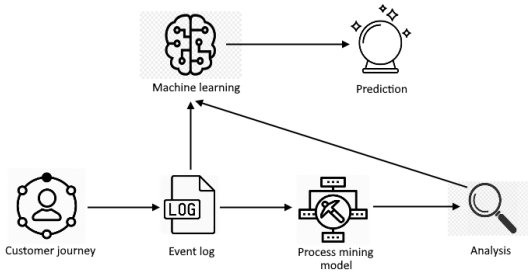


Fig. 1. Methodology overview

the set of l possible touch points $A = \{a_1, a_2, \ldots, a_l\}$ and lastly a timestamp t_i when the event took place [18]. Thus an event log has a collection of traces, which contain events that denote the activity that happened on what time and referring to what customer journey.

We want to predict the activity a of event e_{i+1} given that we are at event e_i belonging to the same customer journey c_j and having the timestamp $t_i < t_{i+1}$. We also have the information of previous events which will help determine the next touch point. Meaning that e_{i+1} is the dependent variable on our independent variables of the previous steps. We do not only use touch points, but also the additional *static* information in the customer journey.

In Figure 1 we can see the steps that can be used to perform predictions on the customer journey.

The main goal is to try and predict what touch point a customer will use next. This prediction is required if the company wants, for instance, to try preventing this next step from happening. This can be done by helping a customer proactively or even better by making sure that the next step is never needed. This does not only save the company resources as the customer needs less attention, the customer will also be more satisfied as the goal of the journey is reached faster.

One example of solutions from the literature addressing relatively similar problems is OARA [6] which this is based on customer journey prediction but also includes a recommendation afterwards, which deviates from the goal in our research. A similar solution is proposed by Terragni & Hassani [17] in an article on analyzing customer journeys for recommendations [16]. This research is useful to investigate as the basis is similar. However we want to predict what a customer will actually do and not what a customer will like (also called the final outcome). Therefore we can investigate their sources and analyze how they tackled the problem and learn from that.

III. Business understanding

The business owner *BO* which owns the scenario and the data preferred not to reveal their identity and is a market leader for telecommunications in the Netherlands. We will refer to them as *BO* in the remainder of this paper. They provide many different types of services for people living in the Netherlands, as well as different services for companies in the business market. The main services provided for customers are mobile telephony, internet for households and interactive television. The different types of services can be combined to get extra benefits on top of the packages. *BO* provides various ways to get in contact with their customers.

A customer can get into contact with *BO* for different reasons such as having a question about a service, an invoice, the installation of a new piece of hardware, reporting malfunctions, acquiring a new subscription with *BO* and a dozen more possible reasons. To facilitate these contacting customers, *BO* has a number of channels to reach them. Customers can contact the service helpdesk of *BO* over the phone, the website, a list of FAQ, a community forum, social media platforms or offline by physically showing up in one of *BO* stores.

A. Customer journey

The customer journey at *BO* is defined to end after a customer has not been in contact with *BO* for at least seven days. Thus, a customer journey can be of arbitrary time duration, the only restriction is that the duration between two consecutive touch points should not be longer than 7 days, otherwise it is assumed to be a new customer journey (new case).

The customer journey is a customer-driven process that starts with a touch point initiated by the customer to interact with *BO*. Most of the customer journeys are quite short, meaning of length one, two and three. The short length is actually a good sign for *BO* because it means that the customer is helped within a few steps.

B. Business Problem

Regarding customer support, there are three high cost factors within *BO*. As such those three are interesting to investigate and in context of this research, they are interesting to predict. Ideally, a customer would never have to call *BO*. All changes, sales, terminations and troubleshooting problems are being resolved using the online portal of *BO*. *BO* offers customer support through their telephone service helpdesk. The service helpdesk can be used in case of a question which could not be answered by the *BO* website and there are always people who prefer to call instead of using self-service options. However keeping the call centre operational is quite expensive. *BO* provides around the clock support for malfunctions, technical support and questions about the theft or lost of a mobile phone. Every call made costs *BO* money, this has already led to the creation of self-service portals online and robotic chat services.

Another big impact factor regarding cost is the mechanic. Sending a mechanic to a customer is costly for both *BO* as for the customer. Mechanics need proper training and planning the mechanics in a way that travel time is minimized is hard and thus costly. The mechanic is unfortunately partly unavoidable, as he needs to fix issues on the customer side. Though there is also a part which is unnecessary, it is hard to distinguish the latter from the former. Improving the quality of manuals might reduce the required number

of mechanics but they will always be needed. Even though, predicting that a mechanic will be needed can certainly help. As preemptive steps can be taken to offer a customer to receive a mechanic. Doing so will save the cost of a customer having to make a repeated call to *BO*.

Last costly defect is swapping of hardware for the customer. When a modem or tv-box has a defect and needs to be replaced, it creates a lot of administration and logistics. Especially keeping track of all aspects of the swap in the logistics is not a trivial task. Even though a large part of the modems returned are not actually broken. The customer will send the malfunctioning modem back to *BO* and *BO* has to send the same model modem from their warehouse to the customer.

These three costs are part of the customer journey and loads of different kinds of research is being done to better understand and improve them. The main costs for these three touch points is the overhead, they are often unnecessarily repeated. The costs of sending two mechanics is substantially higher than one mechanic who spends a bit more time at a customer. Not only those three items are investigated, the whole customer journey is under the loop and is being streamlined more and more.

The objective for *BO* is to resolve the issue of the customer within one contact, as all repeating touch points are considered excessive. To achieve this goal and besides the answer to the first sub-question is to get a prediction on what type of contact (i.e. touch point) a customer will use next and if possible also the subject for which the customer comes into contact with *BO*. As it can make quite a difference if the customer calls for explanation on his bill or to cancel their subscription.

IV. Data understanding

For this research we have two datasets with eight weeks of data, which both already include over a million rows of data and almost half a million customer journeys. The column called *contact_type* is the type of contact a customer has with *BO*, this attribute is what we have defined as a touch point in this paper. Meaning this is the attribute on which we want to do the prediction, we want to predict which type of contact the customer will use next. Therefore all the distinct touch points will be looked at and briefly explained in the list below.

A. Touch Points

- **call**: This touch point means that a customer has called with the service helpdesk of *BO*. The reason behind the phone call can vary a lot, from a malfunction to the theft of a mobile phone.
- **call - dvb**: This is when a call has to be forwarded to another department of the service helpdesk where employees are trained in other skills. For example if a customer wants to buy a product or service, he will be put through to the sales department of the helpdesk.
- **chat**: The chat occurs when a customer is on the *BO* website and uses the online portal to ask a question to the chatbox. First a bot will respond but later a real life employee may continue the conversation if necessary.
- **conversational**: Conversational is when a customer calls the service helpdesk of *BO* but before a real employee is on the phone. A bot will ask the customer to state his question, the bot then tries to classify the question. If the question is general, the bot will send a link to the webpage on which an answer for the question can be found. Conversational helps to reduce the number of calls that have to go through to actual employees.
- **logistiek**: The logistics part is for the swap and distribution of hardware. This is a more static step in the process, as a type of hardware is requested and the warehouse has to perform the logistics to get it to the customer.
- **monteur - levering**: This means a mechanic for delivery. The delivery to a customer who has gotten a new subscription or an upgrade and wants a mechanic to perform the installation of the new modem or box.
- **monteur- ondergrond**: This stands for mechanic - underground. Meaning that there has to be done actual digging or crawling in the crawlspace of the house. The mechanic will then check and replace the cable(s) if necessary. A mechanic - underground is not often needed and mostly only after a regular mechanic - service has come by the house already.
- **monteur - service**: This is the regular mechanic for service. The mechanic is send when a customer calls with a malfunction which cannot be solved by himself or over the phone with the service helpdesk employee.
- **online**: Online is when the customer checks the website of *BO*. This can be for everything on there, even if it is just browsing. Online can be hard to link to the actual customer as most people are not always logged in into their account.
- **order**: In this dataset an order has two categories, namely move and termination. Termination is when a customer cancels his subscription, so the provided services have to be stopped. Move is when a customer changes address and therefore the services, like internet and TV, have to be changed to the new address as well.
- **service ticket**: The service ticket is also like the logistics an intermediate, more static step. The ticket is created by a service helpdesk employee who has a customer with a problem on the phone. The service ticket states all the information needed for later reference if the customer calls again or for the mechanic to be informed about the problem.
- **winkel**: Winkel is dutch for store. So the store is when a customer walks into the store and speaks with an employee. The reason can also vary, it can be to buy a new product or to ask for information or even to report a malfunction.

The next column after *contact_type* differs in the two datasets. The key feature of the first dataset is an important column named *bucket_name*. In this column the bucket in which the touch point is categorized is shown. Not all touch points use the same buckets to

Baseline of touch points	
eind	48%
call	21%
conversational	13%
online	7%
logistiek	3%
service ticket	2%
monteur - service	2%
order	1%
winkel	1%
chat	1%
monteur - levering	0%
monteur - ondergrond	0%

TABLE I: Baseline touch points in the data

be categorized into, which is interesting. We will check which touch points are connected to which bucket.

Touch point occurrence In Table I, the frequency of each touch point in the dataset is shown. All touch points were counted and then normalized, which is the result visible in Table I.

Process overview The journeys are very different for each customer but it could be the case that multiple customers have the same journey as there is only a limited number of touch points. In Figure 2 the process overview, made by the heuristic miner . In this figure we do see some connections. We observe on the left of the figure that a **call** often leads to an **order**, **service ticket** or **logistiek** (logistics). These are interesting observations and when thought about make sense. When a customer calls *BO* it can be about a malfunction or other question, this leads to a **service ticket**. The call can also be about acquiring a service, thus **order**. Moreover it could be a defect piece of hardware, leading to **logistiek** (logistics).

From the **service ticket** a followup is the **monteur - service** (mechanic - service), which is expected. The same holds for an **order** inducing the **monteur - levering** (mechanic - delivery). When a **monteur - service** is not enough to fix the problem a **monteur - ondergrond** (mechanic - underground) is sent, as already stated in the explanation of **monteur - ondergrond** in Section IV-A. So this dependency is also no surprise to see.

The link between **monteur - ondergrond** → **logistiek** and **monteur - ondergrond** → **conversational** is not immediately clear. Possibly for **monteur - ondergrond** → **conversational** the customer calls to ask for an update while the **monteur - ondergrond** is still working on the problem.

B. Distribution of journey types

In the second dataset, we looked at the distribution of the journeys regarding how many types one journey includes, which can be seen in Table II . Meaning in this dataset each customer journey has a category which is one out of the ten types of journeys. These tables show how many different journey types are included in one customer journey. We see that when ignoring customer journeys of length one, the biggest value is when

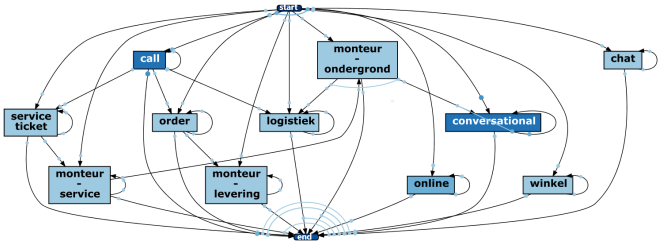


Fig. 2. Process overview using heuristic miner

Number of types	Percentage	Absolute
1	39%	55914
2	40%	57532
3	14%	20731
4	5%	6752
5	1%	2078
6	0%	539
7	0%	91
8	0%	15
9	0%	1

TABLE II: Journey distribution with $length > 1$.

there are two journey types in one customer journey. This is closely followed by just one type per journey. More than two types per journey has a much smaller percentage. Another interesting observation is the fact that there are a lot of customer journeys of length one, there are almost a 100k of them.

V. Data preparation and models

A. Predicted and independent variables

In this research the predicted variable is the data in the column *contact_type_next*. For the independent variables X_i the columns with information have to be selected. For the customer journey this will always be the column containing the touch points, as the touch points are the most important feature in a customer journey. Besides the touch points some additional information regarding the reason of the contact can be added, like a callreason or category. If it would be helpful then also customer information can be added to provide more personal information to the model.

In our research the most important feature is the touch point, i.e. the column *contact_type*. This column contains the current touch point in the journey and has the same values as the outcome variable minus **eind**. Furthermore we also want to include previous touch points belonging to the current journey.

To provide meaning to the touch points, the column with *bucket_name* is used. The bucket provides a sort of category for the current touch point. As we also include previous touch point, we will likewise include the previous buckets in our data.

Another data feature which supports the information surrounding the touch point is the *callreason*. As described previously the *callreason* consist of a written or selected reason made by a *BO* employee. As such there occur many different variations in *callreasons* of

which some are very similar to each other. They can even be as similar as using different punctuation or capital letters. This makes them less suited for the use in prediction, however they contain valuable information. This is why we will perform a basic cleaning performance on the *callreason* data, so it is usable for our models. Because there are still many options available for *callreason*, we will only use the directly previous *callreason* and the current *callreason*.

B. Data cleaning

To reduce the noise in our data we have removed the touch point **call - dvb**. We simply do this by removing the entire row in which **call - dvb** is the *contact_type*, because we will run a script later that reshapes the data in the right way. We made the decision to limit the *callreason* to a maximum of ten characters. We chose ten because it cuts off all too specific parts of the reason while still providing enough room within the first ten characters to be different. Lastly we had the buckets $undefined$, *nog niet toebedeeld* and *unknown*. These are all non-informative buckets and therefore we aggregated them into one bucket instead of three separate buckets. We looped over the data and whenever we encountered one of the three buckets, we replaced the entry with *missing*.

We will use a logistic regression, random forest, boosted trees and a LSTM neural network. Then we will measure the performance of the models with a metric and assess them.

C. Applied models

Logistic regression: The logistic regression is the most basic technique we use. We only have to set a few parameters for this model. As our problem consists of predicting one of multiple outcomes, we need a multi-class classification model. Therefore we have to set the *multi_class* parameter in our logistic regression to *multinomial*, then it will use cross-entropy loss to find the best model. The other parameter we set is the solver, we cannot use liblinear for our multi-class problem as this is only suited for binary problems. We choose for the SAGA solver [5]. This solver performs well in practice and is faster on large datasets than other solvers. We will use these settings to train our logistic regression model.

Random forest: Thousand trees are enough to get an average and reduce overfitting. To further control overfitting we set the *max_depth* parameter to 25.

XGBoost: XGBoost [4] is also based on decision trees. For XGBoost we have to choose the objective function to perform the gradient boosting on, in our case we choose for the $multi{:}softprob$ option. This indicates that we are dealing with a multi-class output and the softprob refers to the softmax function. Instead of returning the label, it returns the probability for each output. Similar to the random forest, here we also choose a *n_estimators* of a thousand and a *max_depth* of 25 to help with memory management and overfitting.

LSTM [9]: Hyperparameter tuning is very important for neural networks. There are a few ways to do the hyperparameter tuning. We can manually tune the hyperparameters but this is very time consuming and you need an expert or else the tuning will not be much of an improvement. The most standard option in parameter tuning is gridsearch. Gridsearch evaluates all the different possible combinations of parameters in a grid-like manner, therefore it is called a gridsearch. However testing all different possible combinations of parameters and finding the best combination takes a lot of computational power and time. This is why a randomized gridsearch was introduced, the randomized gridsearch does not compute all possibly combinations but it randomly chooses a subset of them. By Bergstra and Bengio [1] it is empirically and theoretically shown that randomly chosen trials are more efficient for hyperparameter tuning than trials on a grid.

However there is also a disadvantage in the randomized gridsearch. Randomized gridsearch does not adapt its behavior based on the previous outcomes. This means that a poorly chosen parameter can prevent the model from learning effectively. For example if the dropout rate should be between 0 and 0.5 but we test for values between 0 and 1 then 50% of the tests will return bad results. This is an unnecessary waste of time and therefore the range in which the hyperparameters lie, needs to be chosen carefully. The Bayesian optimization methods by Snoek et al. [13] are capable of learning from the previous trials. Bayesian optimization creates a surrogate objective function to approximate the best hyperparameters for the real model. A study by Bergstra et al. [2] shows that Bayesian optimization methods produce significantly better results whilst also limiting the computation time. Therefore we will use Bayesian optimization in this research. We will now shortly discuss the different hyperparameters we will tune.

- **Gradient descent optimization algorithm**: This optimization technique tries to minimize the loss function after each iteration by tweaking the weights. Some of these optimization algorithms are Momentum, Adagrad, RMSprop, Adam and Adamax [11]. Adam is in general the best performing optimizer [11].
- **Number of neurons in hidden layer**: The number of neurons in the hidden layer determines how well the model learns without underfitting or overfitting.
- **Dropout**: Dropout is a regularization method which drops out random nodes to reduce overfitting and improve overall model performance.
- **Batch size**: The batch size defines the number of samples that will be used every iteration. This hyperparameter is also a balance between not overfitting the model and unable to escape a local minimum. In general it is advised to use a power of two as batch size since this would increase efficiency.
- **Epochs**: The number of epochs is the number of times your model trains on the entire dataset. If this number is too high it will cause overfitting on the opposite side if it is too low then there will be underfitting.

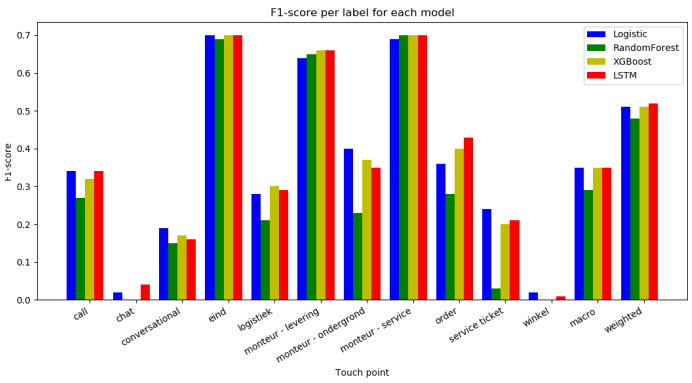


Fig. 3. F_1 of all models w.r.t. the ground truth touch point.

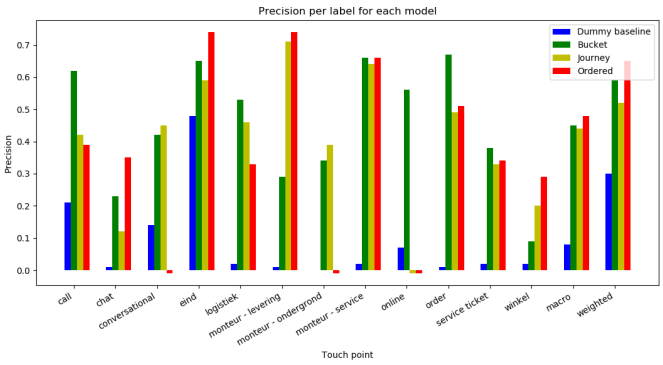


Fig. 4. Precision comparing dummy and three datasets

VI. Experimental Evaluation

To evaluate the prediction quality we used: Precision $= \frac{True\ Positives}{True\ Positives+False\ Positives}$, Recall $= \frac{True\ Positives}{True\ Positives+False\ Negatives}$ and $F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision}+\text{recall}}$.

We have trained the four different models on the two different datasets which were split into three different options. The datasets are split individually into an 80% training set and a 20% testing set .

A. Comparison of applied models

First we will compare the four different kinds of models with each other. For this we show a bar plot depicting the F_1-score for the different models in Figure 3. On the x-axis we plotted the predicted labels, i.e. the touch points. Also a macro average and a weighted average is shown. To compare the different models, it is best to look at the macro average and weighted average. Overall we see that the macro average is equal for the logistic, XGBoost and LSTM models with only the random forest (RF) underperforming. The same observation holds for the weighted average. The models perform similar with only RF underperforming. The reason that RF is worse than the other three could be caused by the fact that we did not tune the parameters of the RF, while we did tune the LSTM and XGBoost boosts itself.

B. Comparison of datasets

In this section we will compare three different types of datasets using the winner model from the previous step: XGBoost. The first is the first dataset used in this research, which we call *buckets*. The second and the third datasets are inferred from the second dataset in this research and differ according to the journey types which are used in two ways. The first way is by just using the dataset trained on all data in a journey, so only sorting by journey id (we call it *journey*). In the other way, we group the journeys on journey id and also on journey type (we call it *ordered*). This creates more journeys and therefore also smaller journeys. However these journeys should all be related to the same subject as they share the same journey type. The precision, recall and F1-score are shown in the Figures 4, 5 and 6 respectively. In the figures the metric is on the y-axis and the different touch points on the x-axis. The bars are the three different datasets as well as a dummy baseline for reference. A dummy baseline randomly predicts one of the labels but with a probability weighted by its relative frequency in the ground truth. For some datasets certain touch points are not available and this is displayed by a small negative value, for example for the ordered dataset there is no **conversational** label. First we will inspect the overall score with the F1-score measure shown in Figure 6. Comparing the macro average and weighted average, we see that on all three datasets, XGBoost outperform the dummy baseline which is good. For the macro average, XGBoost performs the best on bucket dataset and worst on the ordered dataset. While on the ordered dataset it performs the best in the weighted average, on the bucket dataset it also performs well.

We observe two touch points that are very poorly predicted. These touch points are **chat** and **winkel** (store). This could be explained by the fact that these are two of the smaller labels and therefore less tuned on by the models. However thinking about these touch points, they both do not belong to clear processes. A customer can go to the store whenever he wants but he is never expected to go to a store. This makes the store a very unpredictable touch point. The chat has the same issue only when we see a customer online, we could predict that he is going to chat. However there is never a clear indication that the customer will chat with a *BO* employee.

Now we will look at the touch points most interesting to us, namely **call** and **monteur - service**. We observe that **monteur - service** is very good predictable in all datasets. This is likely caused by the fact that a mechanic for service is always sent in a reaction to something and it is not sent out of the blue. The indications are used by the model to predict when a mechanic will be sent. Looking at the **call** touch point only the bucket dataset performs well and the ordered dataset performs very poorly. The poor performance of the ordered dataset could be caused by the fact that there is no **conversational** data in this dataset and **conversational** is one of the biggest indicators that a **call** might follow.

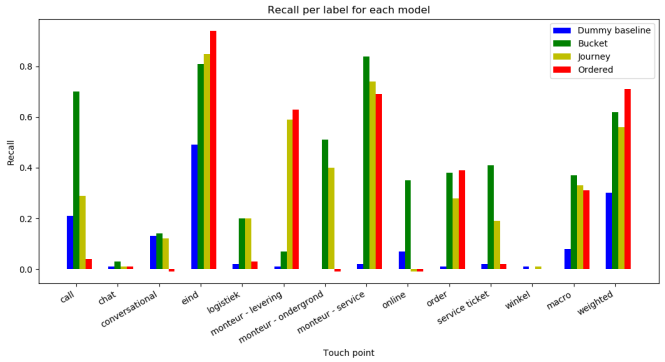


Fig. 5. Recall comparing dummy and three datasets

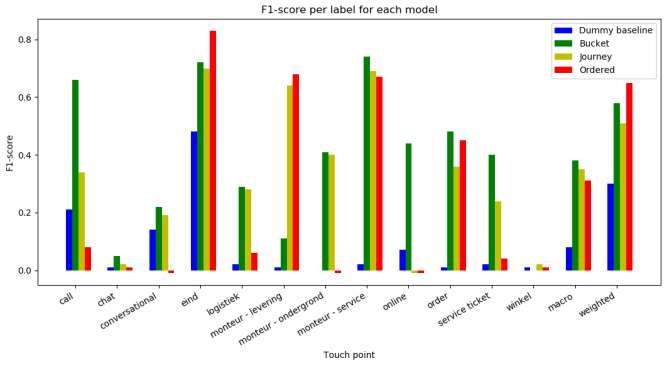


Fig. 6. F_1-score comparing dummy and three datasets

VII. Conclusions

In this paper we have discussed the customer journey and the predictability of its different touch points. First we want to get a better overview of the customer journey. Therefore we started this research by investigating the customer journey. We applied process mining techniques to discover the process model. Then we moved towards predicting the next step in the journey. We have shown the intuition behind each preprocessing step either from the business understanding perspective or from the data analysis perspective. We concluded that among four, carefully-tuned prediction models, XGBoost was the winner so we proceeded with testing it on three datasets. In the results we have shown that we are always able to beat a dummy baseline which predicts randomly one of the labels with a probability weighted by its existence in the ground truth. To have a structured approach to our investigation, we followed a framework similar CRISP-DM [12].

In the future, we would like to address the meaning of the used distance metric in categorizing the journeys into variants by inferring an accurate distance metric that decides the similarities between the journeys [15]. Additionally, we would like to address the predictability under a streaming setting of the customer journeys [8], with varying underlying distributions [7], [14].

References

[1] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *JMLR*, 13(Feb):281–305, 2012.

[2] James S Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for hyper-parameter optimization. In *NeurIPS*, pages 2546–2554, 2011.

[3] Gaël Bernard and Periklis Andritsos. A process mining based model for customer journey mapping. In *CAiSE 2017*, pages 49–56.

[4] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *KDD*, pages 785–794. ACM, 2016.

[5] Aaron Defazio, Francis Bach, and Simon Lacoste-Julien. Saga: A fast incremental gradient method with support for non-strongly convex composite objectives. In *NeurIPS*, pages 1646–1654, 2014.

[6] Joël Goossens, Tiblets Demewez, and Marwan Hassani. Effective steering of customer journey via order-aware recommendation. In *ICDMW'18*, pages 828–837.

[7] Marwan Hassani. Concept drift detection of event streams using an adaptive window. In *ECMS*, pages 230–239, 2019.

[8] Marwan Hassani, Daniel Töws, Alfredo Cuzzocrea, and Thomas Seidl. *BFSPMiner*: an effective and efficient batch-free algorithm for mining sequential patterns over data streams. *Int. J. Data Sci. Anal.*, 8(3):223–239, 2019.

[9] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Comput.*, 9(8):1735–1780, November 1997.

[10] Katherine N Lemon and Peter C Verhoef. Understanding customer experience throughout the customer journey. *J. of market.*, 80(6):69–96, 2016.

[11] Sebastian Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

[12] Colin Shearer. The CRISP-DM model: the new blueprint for data mining. *Journal of data warehousing*, 5(4):13–22, 2000.

[13] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. In *NeurIPS*, pages 2951–2959, 2012.

[14] Yorick Spenrath and Marwan Hassani. Predicting business process bottlenecks in online events streams under concept drifts. In *ECMS*, pages 190–196, 2020.

[15] Yorick Spenrath, Marwan Hassani, Boudewijn F. van Dongen, and Haseeb Tariq. Why did my consumer shop? learning an efficient distance metric for retailer transaction data. In *PKDD*, pages 323–338, 2020.

[16] Alessandro Terragni and Marwan Hassani. Analyzing customer journey with process mining: From discovery to recommendations. In *FiCloud '18*, pages 224–229.

[17] Alessandro Terragni and Marwan Hassani. Optimizing customer journey using process mining and sequence-aware recommendation. In *SAC'19*, pages 57–65.

[18] Wil M. P. van der Aalst. *Process Mining: Data Science in Action*. Springer, April 2016.

Finance and Economics and Social Science

DEMOGRAPHIC AND STATISTICAL MODELLING OF GRANDFATHERHOOD IN RUSSIA

Oksana Shubat
Mark Shubat
Ural Federal University
620002, Ekaterinburg, Russia
E-mail: o.m.shubat@urfu.ru
E-mail: Mark.Shubat@urfu.me

KEYWORDS
Grandfatherhood, grandparents, demographic modelling, statistical demographic model.

ABSTRACT

In recent years, negative demographic trends have been developing in Russia. The most important is a decline in the birth rate. Researchers are actively looking for new determinants of this process, on the basis of which measures of population policy can be developed. One of these determinants may be active grandparenting, which means the active participation of grandparents in the processes of caring for grandchildren. The aim of this study is to create a demographic and statistical model of a typical Russian grandfather, actively involved in childcare. We used the following methods: parametric and nonparametric independent samples tests (t-test, Mann-Whitney U test, median test), regression analysis, indirect method of calculations. As a result, two models were presented – statistical demographic model of the age when Russian men enter grandparenthood and demographic model of a typical Russian grandfather actively involved in childcare. Our study is a preliminary stage for a large-scale survey of grandparenting practices in Russia. The number of older people is growing fast, which makes this socio-economic group increasingly important for addressing the problems of demographic decline in Russia. Therefore, large-scale research of grandparenthood is crucial for more efficient policy-making in this sphere.

INTRODUCTION

Recent demographic trends in Russia present an alarming picture: since 2016, the country has experienced a natural population decline and falling birth rates and the period of 2019-2020 marked a significant population decline (Demographic Indicators 2020).

To address these issues, a number of state measures are being developed and implemented, including the national project "Demography" for the period until 2024, which came into force on 1 January 2019. The project aims at increasing healthy life expectancy to 67 and ensuring a rise in the total fertility rate to 1.7 children per woman. The project encompasses several federal projects dealing with specific goals (Passport of the national project 2019). We believe, however, that these goals can be achieved not only through targeted effort but also as a result of a synergistic effect. One of the factors contributing to the growth in fertility and healthy life expectancy is the involvement of grandparents into the process of caring for their grandchildren.

There is a substantial body of international research on grandparents' role in childcare and upbringing (Sichimba at al. 2017; Nedelcu 2017; Coall at al. 2018). These studies bring to light a number of social and psychological benefits enjoyed by grandparent caregivers: for example, there is evidence that child-raising has a positive impact on grandparents' cognitive functions (Arpino and Bordone 2014), that it enhances their subjective well-being (Mahne and Huxhold 2015), reduces the risk of depression (Grundy 2012), and decreases the mortality rates in elderly people (Hilbrand at al. 2017). There are studies focused on grandparents' positive influence on the well-being of their grandchildren, for instance, on their academic performance (Del Boca at al. 2018). Some studies highlight the role of grandparents in helping families surmount crises (Attar-Schwartz and Buchanan 2018).

At the same time, the level of grandparents' involvement in childcare may differ across countries and regions. For example, there is evidence that it varies significantly across northern and southern European countries (Buchanan and Rotkirch 2018). The intensity of grandparental involvement may also change with time: for instance, as is shown in (Chapman at al. 2017), Finnish children born in 1869 spent on average four years with at least one of their grandmothers and one year with at least one of their grandfathers; for children born in 1950, these figures rose to 24 and 13 years respectively.

In Russia, grandparents have traditionally played an important role in providing childcare and support. Unfortunately, there are currently no studies in Russia that would provide reliable data on grandparenthood, which precludes efficient policy-making in the social and demographic sphere. At the same time, obtaining such assessments could become the basis for the development of those measures of state social and demographic policy that would contribute to a more effective solution of demographic problems in the country.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

In Shubat and Bagirova (2020), we presented a demographic-statistical model of a typical Russian grandmother actively involved in childcare. The aim of this study is to create a demographic and statistical model of a typical Russian grandfather, actively involved in this process.

DATA AND METHODS

In order to model the specific features of Russian grandfathers, we had to address two methodological tasks.

First, we needed to identify the socio-demographic group of grandfathers as men with grandchildren. Despite the fact that the Russian government pays close attention to the problems of senior citizens (measures to support the elderly are specified in the above-mentioned national project "Demography"), there are currently no large-scale national surveys of grandparenthod in Russia. Therefore, there is a perceived lack of statistical data on the size of this group and the criteria that can be applied to identify who belongs to it. Based on the data available in Russian statistics, we found it possible to identify this group based on the age criterion. Therefore, we had to find at what age Russian men and women enter grandparenthood.

The age of grandparenthood is easier to calculate for women with the help of the statistical indicator "Mean age at first birth (the mean age of women at the birth of their first child)". We need to add up these indicators for the two consecutive generations of women to calculate the mean age of entering grandmothering. However, it was impossible to use the same approach to estimate at what age Russian men enter the age of grandfatherhood as there are no data on the mean age at first birth for men. Therefore, we had to build a more complex statistical demographic model. To this end, we used the following statistical data sources:

- mean age at first birth (for women);
- average age at marriage (for women);
- average age at marriage (for men).

The data were provided by the Human Fertility Database, a joint project of the Max Planck Institute for Demographic Research and the Vienna Institute of Demography (The Human Fertility Database 2021). We also relied on the data of the annual demographic report "Population of Russia" (Collection of indicators 2020), whose estimates are based on the Russian and international official statistics.

Second, it is important to note that focusing exclusively on the age criterion makes it possible to identify only the socio-demographic group of potential grandfathers. However, not all Russian men who have entered the age of grandparenthood are actually grandfathers and not all of them take an active part in childcare. Therefore, we had to distinguish between actively involved and disengaged grandparents. As noted above, in Russia no research on the problems of grandparenthood is currently conducted. The only source of valid and reliable data to build a demographic and statistical model of a typical Russian grandfather actively involved in childcare is the federal statistical survey "Comprehensive Monitoring of Living Conditions" (Comprehensive Monitoring of Living Conditions 2018) conducted by the Federal State Statistics Service of Russia. The survey's results are considered representative not only of the country in general but also of specific regions and socio-demographic groups. The most recent data were collected in 2018. We used some of the questions from this survey to build a model of a typical Russian grandfather actively involved into childcare.

The question we used to identify such men was as follows: "Do your daily activities include unpaid care for children, your own or somebody else's?". Grandfathers who gave a positive answer to this question were identified as grandfathers actively involved into childcare.

To build our model we used the following variables:

- Var 1: age (years);
- Var 2: educational level (years spent on education);
- Var 3: marital status;
- Var 4: place of residency (urban or rural area);
- Var 5: social activities – visiting theater, cinema, sports, religious events, cafes and restaurants traveling around the country and abroad in the last year. These variables were used to build a new one (Var 5), reflecting the total number of grandfathers' activities in the last year;
- Var 6-9: health-related variables - objective and subjective health estimations:
  - ✓ Var 6: frequency of health practitioner visits;
  - ✓ Var 7: frequency of ambulance calls;
  - ✓ Var 8: self-assessment of health (from 1 – "very bad" to 5 – "very good");
  - ✓ Var 9: self-assessment of the opportunity to lead an active life.

We suppose that these variables can determine grandfathers' engagement in childcare. For example, we suppose that those grandparents who are more active socially also tend to be more actively involved in raising their grandchildren while healthier grandparents are also more likely to be willing to take care of their grandchildren and so on.

To study the specific characteristics of such grandparents, we analyzed the statistical differences between this group and the group of grandparents that disengaged from childcare. For this purpose we used the following parametric and nonparametric independent samples tests:

- t-test;
- Mann-Whitney U test;
- median test

These statistical tests were chosen for the following two reasons: first, they are suitable for different types of data with different characteristics of distribution. In our study the tested variables were measured in different scales and the distribution of some variables differed

from the normal. Second, in contemporary research literature, there is no universal agreement regarding the applicability or benefits of this or that test, in each case the authors' own experiments and simulations were used to show the effectiveness of the chosen test (see, for example, (Gibbons and Chakraborti 1991; Hollander at al. 2013; Mood 1954; Zar 2018; Zimmerman 1987)). In our analysis, we considered the differences between the groups to be confirmed if at least two of the significance tests we used showed a positive result.

For dichotomous and categorical variables, we used crosstabs to model for differences and computed the Phi coefficient and Cramer's V. We also used econometric modelling, that is, estimated a regression model by the OLS method.

RESULTS

The main results of our study are as follows.

1. We used the available statistical data to build a statistical demographic model of the age when Russian men enter grandparenthood (see Figure.1).

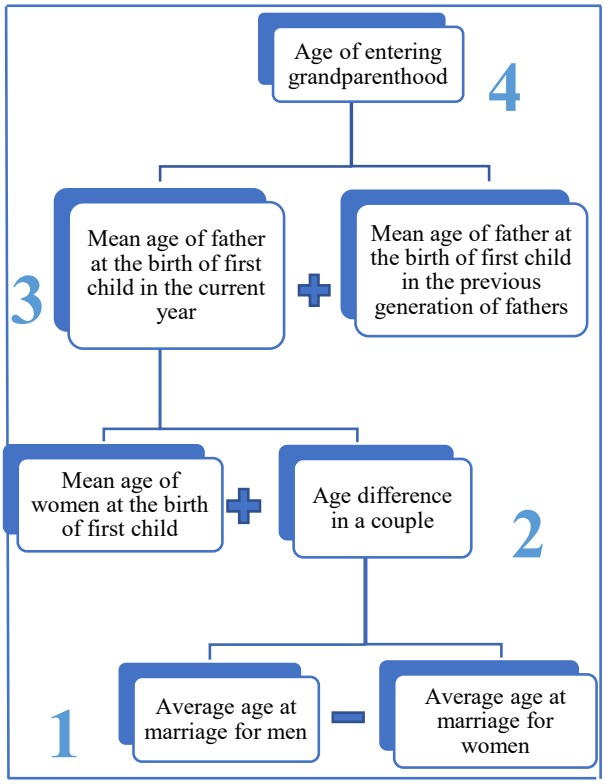


Figure 1: Statistical Demographic Model of the Age When Russian Men Enter Grandparenthood

2. Relying on the above-described model and the available statistical data, we were able to estimate the mean age of entering grandparenthood for Russian men in 2007-2016 (see Table 1).

Table 1: Mean Age at Grandfatherhood in Russia in 2007-2016

Year	2007	2008	2009	2010	2011
Mean age	51,9	52,1	52,3	52,3	52,3
Year	2012	2013	2014	2015	2016
Mean age	52,4	52,6	52,8	53,07	53,1

3. As noted previously, the most recent information that can be used for demographic modelling of a typical Russian grandfather refers to 2018. To calculate the mean age of entering grandparenthood for this year we used econometric modelling, more specifically, we estimated a trend model. Visualization of the primary data highlighted a linear trend. According to the econometric model (see Tables 2-3), the mean age of entering grandfatherhood in 2018 was 53.3.

Table 2: Model Summary

R Square	Adjusted R Square	Std. Error of the Estimate	F	Sig.
0.961	0.956	0.079	196.7	0.000

Table 3: Coefficients

Model		Unstandardized Coefficients		t	Sig.
		B	Std. Error		
1	Constant	-192.349	17.457	-11.018	0.000
	Year	0.122	0.009	14.026	0.000

4. We applied the age-related criterion to form two groups of grandfathers – "active" (engaged in childcare on a daily basis) and "inactive" (disengaged from active childcare). As a result of the selection process, the first group comprised 1,562 respondents, while the second, 16,091. Thus, grandparents taking care of their grandchildren on a daily basis account for only 9% of the whole populatio.

5. We tested for the significance of the differences and found that the major differences between the two groups of grandfathers are related to the following:

- Var 1: age ("active" grandfathers tend to be younger);
- Var 2: education ("active" grandfathers have a higher level of education);
- Var 5: social activity in different spheres ("active" grandfathers are also more prone to be socially active - they more often go to the cinema or theatre, to cafes and restaurants or sports events);
- Var 8: self-assessed health status ("active" grandfathers rate their health higher).

Test results are presented in Tables 4-8. In particular, Table 4 shows mean values of the variables analysed in two groups of grandfathers. Table 5 provides results of testing the significance of these mean values' differences; the equality of variances in the groups compared was verified preliminarily using Levene's Test. The results show that differences for all variables tested are statistically highly significant ($p<0.001$).

Tables 6 and 7 present results of nonpaprametric Mann-Whitney Test, which we used to test whether two samples are likely to derive from the same population. As the data show, null hypotheses (H_0: The two populations are equal) were not confirmed ($p<0.001$), which testifies to the significance of the differences identified by comparing two groups of grandfathers. Table 8 shows results of comparing medians of the variables studied. Tests proved that differences between medians are statistically significant ($p<0.01$).

Table 4: Group Statistics (t-test)

Variable	Is childcare a part of daily activities?	N	Mean
Var 1	Yes	1562	62.91
	No	16091	65.43
Var 2	Yes	1556	12.18
	No	15993	11.80
Var 5	Yes	1562	0.83
	No	16091	0.60
Var 8	Yes	1561	3.03
	No	16075	2.90

Table 5: Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means	
		F	Sig.	t	Sig. (2-tailed)
Var 1	1*	156.368	0.000	-11.691	0.000
	2*			-14.374	0.000
Var 2	1*	15.380	0.000	5.675	0.000
	2*			6.183	0.000
Var 5	1*	43.438	0.000	8.134	0.000
	2*			7.257	0.000
Var 8	1*	89.929	0.000	7.876	0.000
	2*			8.720	0.000

* 1 – Equal variances assumed
2 - Equal variances not assumed

Table 6: Ranks (Mann-Whitney Test)

Variable	Is childcare a part of daily activities?	Mean Rank	Sum of Ranks
Var 1	Yes	7544.27	11784157.00
	No	8951.52	144038874.00
Var 2	Yes	9436.24	14682791.50
	No	8710.67	139309683.50
Var 5	Yes	9740.63	15214857.50
	No	8738.31	140608173.50
Var 8	Yes	9626.57	15027071.00
	No	8740.03	140495995.00

Table 7: Mann-Whitney Test Statistics

	Var 1	Var 2	Var 5	Var 8
Mann-Whitney U	1.06E+07	1.14E+07	1.11E+07	1.13E+07
Wilcoxon W	1.18E+07	1.39E+08	1.41E+08	1.40E+08
Z	-10.429	-5.453	-8.749	-7.926
Asymp. Sig. (2-tailed)	0.000	0.000	0.000	0.000

Table 8: Median Test Statistics

		Var 1	Var 2	Var 5	Var 8
Median		64.0	12.0	0.0	3.0
Chi-Square		77.4	7.7	72.7	9.0
Asymp. Sig.		0.000	0.006	0.000	0.003
Yates' Continuity Correction	Chi-Square	76.9	7.6	72.2	8.7
	Asymp. Sig.	0.000	0.006	0.000	0.003

6. The test results were used to build the following demographic model of a typical Russian grandfather actively involved in childcare (see Figure 2).

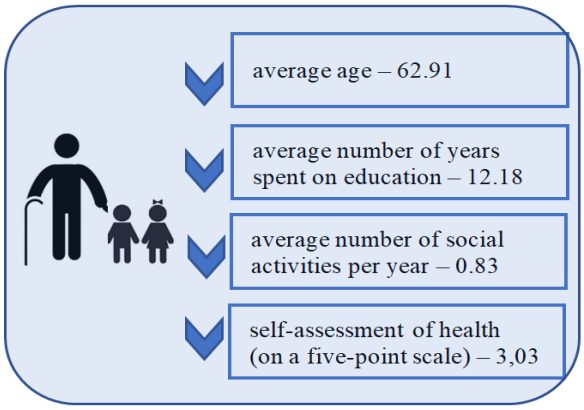


Figure 2: Demograpfic Model of a Typical Grandfather Actively Involved in Childcare

DISCUSSION

The following comments can be made regarding the results of our analysis.

First, it is necessary to comment on the specificity of the proposed statistical demographic model for estimating the age when Russian men enter grandparenthood. The model is based on the available statistical data, that is, the set of indicators assessed by the national statistical services. The model is also designed taking into account the length of the time series used in the study. However, this model can be applied in other countries, where, like in Russia, no specialized surveys of grandparenthood are conducted. The set of indicators used in the model is quite typical of national statistical systems and is by no means unique to Russia.

Undoubtedly, if the set of demographic indicators is expanded, there are more accessible data and large-scale surveys are conducted, especially regarding fatherhood statistics, the above-described indirect estimations may be subject to some corrections.

Second, the proposed model of an "active" grandfather - the one actively engaged in daily care for their grandchildren – also has room for improvement. In our study we used the data of the federal statistical survey "Comprehensive Monitoring of Living Conditions", which is conducted once in every two years and the currently available data cover four time periods. Further research avenues may include comparative analysis of demographic and statistical models of active Russian grandfathers based on the data for different years. Such analysis will make the model more reliable and robust and will help reveal the model's changes over time.

Third, in Russia there are long-standing social, demographic and economic disparities across regions, which means that there might be also specific regional models of grandparenthood. Therefore, special studies investigating such regional models are necessary. It should be noted that those information resources that are currently available in Russia are suitable for regional-data analysis.

Fourth, our results have shown that the group of grandparents actively involved in childcare is quite small. At the same time, there are a number of studies we mentioned earlier, that confirm the high importance of this socio-demographic group. Therefore, in order to tackle the demographic (and other) issues more efficiently, the Russian government needs mechanisms to encourage active grandparenting, especially financial incentives such as payments to grandparents for the time spent looking after their grandchildren while their parents work or study. Such possibilities were explored in our previous publications (see, for instance, [14]).

It should be noted that the actual scale of grandparents' involvement in childcare is obviously greater. Indeed, grandparents can help parents from time to time or even regularly, but not every day. Unfortunately, there are no specialized studies that would estimate the scale of such involvement. The currently available statistical and demographic resources do not provide such data.

Fifth, there is a number of issues concerning the specific characteristics of Russian grandfathers that our analysis has revealed. The fact that they are more active socially and, according to their own assessments, enjoy good health can be a reason for their active involvement in childcare or, vice versa, its result. This issue, therefore, requires further research. Nevertheless, the very fact of this correlation (regardless of its direction) supports the existing evidence regarding grandparenting practices in other countries and shows certain positive socio-psychological effects of grandparent-grandchild communication. This consideration can serve as one more argument to support the view that more incentives are necessary to encourage grandparents to engage in childcare.

CONCLUSIONS

This study proposes a statistical demographic model for estimating the mean age of entering grandparenthood for Russian men. This model relies on the indirect estimation of the Russian statistical data. It can also be applied for research in other countries whose national statistical organizations use a similar set of demographic indicators.

Our demographic model of a typical Russian grandfather who is actively engaged in childcare shows that such men are generally better educated, are more active socially and, according to their self-reported health status, enjoy better health.

The proposed models can be further improved by developing national demographic statistics, expanding the range of indicators and conducting special grandparenthood surveys.

Our findings demonstrate why more active involvement of grandparents into childcare is such a pertinent task and how the government can stimulate such involvement.

It should be noted that this study is a preliminary stage for a large-scale survey of grandparenthood practices in Russia. The number of older people is growing fast, which makes this socio-economic group increasingly important for addressing the problems of demographic decline in Russia. Therefore, large-scale research of grandparenthood is crucial for more efficient policy-making in this sphere.

ACKNOWLEDGMENTS

The reported study was funded by RFBR, project number 20-011-00280.

REFERENCES

Attar-Schwartz, S. and A. Buchanan. 2018. "Grandparenting and adolescent well-being: evidence from the UK and Israel". *Contemporary Social Science*, Vol 13(2), 219-231.

Buchanan, A. and A. Rotkirch 2018. "Twenty-first century grandparents: global perspectives on changing roles and

consequences". *Contemporary Social Science*, Vol 13(2), 131-144.
Chapman, S.; M. Lahdenperä; J. Pettay; and V. Lummaa. 2017. "Changes in length of grandparenthood in Finland 1790-1959". *Finnish Yearbook of Population Research*, No. 52, 3-13.
Coall, D.A.; S. Hilbrand; R. Sear: and R. Hertwig. 2018. "Interdisciplinary perspectives on grandparental investment: a journey towards causality". *Contemporary Social Science*, No. 13(2), 159-174.
Collection of indicators of the Annual Demographic Report "Population of Russia". Moscow: Demoscope. URL: http://www.demoscope.ru/weekly/edd/edd_tab.php (access date 17.10.2020).
Comprehensive monitoring of living conditions. 2018. Moscow: Rosstat. URL: https://gks.ru/free_doc/new_site/KOUZ18/index.html (access date 10.01.2021).
Del Boca D.; D. Piazzalunga; and C. Pronzato. 2018. "The role of grandparenting in early childcare and child outcomes". *Review of Economics of the Household*, Vol. 16(2), 477-512.
Demographic Indicators of the Federal State Statistics Service of Russia. 2021. Moscow: Rosstat. URL: https://rosstat.gov.ru/folder/12781 (access date 20.01.2021).
Gibbons, J.D. and S. Chakraborti. 1991. "Comparisons of the Mann-Whitney, Student's t, and Alternate t Tests for Means of Normal Distributions". *The Journal of Experimental Education*, 59:3, 258-267, DOI: 10.1080/00220973.1991.10806565
Grundy, E. M.: C. Albala; E. Allen; A. D. Dangour; D. Elbourne; and R. Uauy. 2012. "Grandparenting and psychosocial health among older Chileans: A longitudinal analysis". *Aging & Mental Health*, No. 16(8), 1047-1057.
Hilbrand, S.; D. A. Coall; D. Gerstorf; and R. Hertwig. 2017. "Caregiving within and beyond the family is associated with lower mortality for the caregiver: A prospective study". *Evolution and Human Behavior*, No. 38(3), 397-403.
Hollander, M.; D. A. Wolfe; and E. Chicken. 2013. "*Nonparametric Statistical Methods*". New York: John Wiley & Sons.
Mahne, K. and O. Huxhold. 2015. "Grandparenthood and subjective well-being: Moderating effects of educational level". *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, No 70(5), 782-792.
Mood, A. M. 1954. "On the asymptotic efficiency of certain non-parametric 2-sample tests". *Annals of Mathematical Statistics*. 25(3), 514-522.
Nedelcu, M. 2017. "Transnational grandparenting in the digital age: mediated co-presence and childcare in the case of Romanian migrants in Switzerland and Canada". *European Journal of Ageing*, No. 14(4), 375-383.
Passport of the national project "Demography". 2019. Moscow: Rosmintrud. URL: https://rosmintrud.ru/ministry/programms/demography (accessed date 02.02.2021).
Shubat, O. and A. Bagirova. 2020. "Russian grandparenting: demographic and statistical modelling experience". *Communications of the ECMS,* Vol. 34, Issue 1 (June), 78-83.
Sichimba, F.; H. Mooya; and J. Mesman. 2017. "Predicting Zambian Grandmothers' Sensitivity Toward Their Grandchildren". *International Journal of Aging and Human Development*, No. 85(2), 185-203.
The Human Fertility Database. URL: https://www.humanfertility.org/cgi-bin/main.php (accessed date 10.01.2021).
Zar, J. H. 2018. "Biostatistical Analysis". Essex: Pearson Education Limited.
Zimmerman, D.W. 1987. "Comparative Power of Student T Test and Mann-Whitney U Test for Unequal Sample Sizes and Variances". *The Journal of Experimental Education*, 55:3, 171-174, DOI: 10.1080/00220973.1987.10806451

AUTHOR BIOGRAPHIES

OKSANA SHUBAT is an Associate Professor of Economics at Ural Federal University (Russia). She received her PhD in Accounting and Statistics in 2009. Her research interests include demographic processes, demographic dynamics and their 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 webpage can be found at http://urfu.ru/ru/about/personal-pages/O.M.Shubat/

MARK SHUBAT is a student of the Ural Federal University. He is receiving his bachelor's degree from the Engineering School of Information Technologies, Telecommunications and Control Systems. His research interests are related to the application of methods of mathematical statistics to the study of socio-demographic processes. He is engaged in modeling and simulation of these processes.

MODELS FOR FORECASTING THE NUMBER OF RUSSIAN GRANDPARENTS

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

KEYWORDS
Forecating, regression forecast model, number of grandparents, the age of entry into the grandparenthood, grandparents' labor

ABSTRACT

Russian demographic statistics does not provide information about the number of grandparents. The aim of our study is to present models for forecasting their number. We used data from the Human Fertility Database to estimate the average age of a mother at the birth of her first child. Based on the simulated age of Russian women's entry into grandparenthood, the time series of the number of Russian grandmothers was created. To obtain prospective estimates of the number of Russian grandmothers, we tested various models used in demography to forecast population size – mathematical (based on exponential and logistic functions) and statistical (based on statistical characteristics of time series). To estimate the number of grandmothers who are significantly involved in caring for grandchildren, we used data from the Federal statistical survey. Our results are as follows: 1) there is an increase in the age of entry into grandparenthood; 2) we estimated the size of potential grandmothers in different years and we found two models which are more appropriate for forecasting: linear trend model and average absolute growth model; 3) using these models, we predicted an increase in the number of both potential and active grandmothers in the next 5 years.

INTRODUCTION

Depopulation in Russia, as in many countries of the world, is associated with a number of demographic processes. The population decline is ocurring amid a decrease in the birth rate (despite increased government measures to support families with children) and an increase of population aging. An increase in the average age of the population can be facilitated by both an absolute increase in the number of older people and an increase in the share of older people in the total population.

Traditionally, demographic data provide information about the population of certain age groups: young, working age and elderly population. For these categories of the population, both actual and forecast data are published - the latter are used to forecast the development of the labor market. However, the family-role aspect of presented data on population size is also interesting, especially considering the priority of the state task of increasing the birth rate, the complexity and multifactorial nature of demographic processes and their mutual influence, as well as the variety of family models existing in modern societies.

The family roles that people implement affect different types of their behavior, such as labor, consumer, leisure, political behavior, etc. Therefore, data on population size in the family-role context can have important applied value. For example, the number and the share of parents in society can affect the marketing of certain product groups, the development of leisure infrastructure. The number of grandparents can affect the development of the institution of caring for this category of the population (Raišienė et al., 2019), the distribution of programs to improve their grandparent competencies, the development of special banking products, the development of medical services related to the provision of medical care to people with direct relatives, etc. It should be noted that "the intra-family organization of care for elderly relatives is positively portrayed in Russian society as the fulfillment of intergenerational moral obligations" (Tkatch 2015). The desirability of an intrafamilial organization of care for the younger generation (primarily through the involvement of grandparents) is interpreted in a similar way.

However, the social norms prevailing in Russian society, which are associated with the desirability of the participation of grandparents in the life of their grandchildren, may encounter a number of objective and subjective factors that impede the implementation of these norms. For example, one of such objective factors is a some specificity of Russian economics and society: people's entry into grandparenthood does not always mean that they leave the labor market or have enough time for grandchildren.

Stereotypes about old age can act as a subjective factor that prevents the revitalization of grandparental care for grandchildren. According to the Russian traditional understanding, grandparenthood is associated with old age - it is a trigger for age awareness, a marker of aging (Zelikova 2020). As sociologist Zelikova notes, "older women in Russia have only one role – a grandmother. So when a woman has grandchildren, she realizes that this event (the birth of a grandchild) fundamentally changes her status. All the events that

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

took place in her life before, events that were associated with age-related changes, could be described as maturing or developing. Marriage, birth of children, career changes are age-related roles, but they contain attractive images and behaviors, they are not associated with the discourse of decline and withering. The image of a grandmother is completely different" [*completely negative - authors*].

Forecasting the number of grandparents is a rather interesting issue from a methodological point of view, and it is not covered enough in modern demography. The solution to the issue involves the identification of other indicators - the age of entry into grandparenthood, the duration of grandparenthood. There is no unified approach to such assessments in demography (Margolis and Verdery 2019; Leopold and Skopek 2015; Margolis and Wright 2017; Margolis 2016; Yahirun, Park and Seltzer 2018).

Thus, it is necessary to highlight the following factors that hinder the forecasting of the number of grandparents in Russia: 1) lack of statistical data on the number of grandparents; 2) influence of numerous factors on the performance of grandparental functions by biological grandparents; 3) lack of a uniform approach in demography to assessing the age of entry into grandparenthood, the duration of grandparenthood.

These factors determine substantiation of the forecasting period. In demographic studies, forecasts are traditionally subdivided into short-term (up to 5 years), mid-term (up to 25 years), and long-term (more than 25 years). Considering the restrictions specified, we believe we can now undertake only short-term forecasts of the Russian grandparents' number. Additionally, over recent years, Russia has actively implemented demographic policy measures aimed at improving the unfavourable demographic situation. These measures may significantly influence the trends being formed in the population dynamics; consequently, mid-term and long-term evaluations of socio-demographic groups' sizes in Russia would not be accurate enough.

Therefore, the purpose of our study is to forecast, under these restrictions, the number of grandparents in modern Russia in the short term.

DATA AND METHODS

The specificity of the Age-Sex structure of the Russian population requires a construction of two different forecast models of the number of grandparents - separately for grandmothers and grandfathers. Historically, there is a significant difference in the life expectancy of these population groups in Russia – more than 10 years (Russian Statistical Yearbook-2019). Obviously, generalized forecast models will be rather arbitrary and approximate. This study presents models for forecasting the number of Russian grandmothers.

To construct such models, we needed data on the age of Russian women's entry into grandparenthood. Since Russian official statistics do not have this kind of data, we modeled this age based on data on the average age at which a woman gives birth to her first child – in the current year, i.e. in the current generation of mothers, and in the previous generation of women:

$$\begin{array}{c}\text{women's age of entry into grandparent-hood}\end{array} = \begin{array}{c}\text{average age of a mother at the birth of her first child in the current year}\end{array} + \begin{array}{c}\text{average age of a mother at the birth of her first child in the previous generation of women}\end{array}$$

We used data from the Human Fertility Database developed by the Max Planck Institute for Demographic Research (Rostock, Germany) and the Vienna Institute of Demography (Vienna, Austria) to estimate the average age of a mother at the birth of her first child. The database on the Russian Federation contains estimates of the average age of a mother at the birth of her first child until 2018 inclusive (The Human Fertility Database).

Based on the simulated age of Russian women's entry into grandparenthood, the time series of the number of Russian grandmothers was created. The period from 2010 to the present was chosen for the analysis, since 2010 marked the beginning of a stable population growth in the country after a long period of depopulation.

To obtain prospective estimates of the number of Russian grandmothers, we tested various models used in demography to forecast population size – mathematical (based on exponential and logistic functions) and statistical (based on statistical characteristics of time series).

The described technique allows us to obtain estimates of only the potential number of Russian grandmothers. Indeed, not all Russian women in the simulated age range are, in fact, grandmothers, and not all of them are actively involved in caring for their grandchildren. To estimate the number of "active" grandmothers (those who are significantly involved in caring for grandchildren), we used data from the Federal statistical survey "Comprehensive monitoring of living conditions" (Comprehensive monitoring of living conditions 2018). This survey was conducted by the Federal State Statistics Service of Russia in 2011, 2014, 2016 and 2018. Its results are representative for the country as a whole, as well as for individual regions and socio-demographic groups of the population. Despite the fact that this study is not specialized for the study of grandparenthood, some of the questions still allow us to analyze the degree of grandparents' involvement in caring for grandchildren. As an indicator of this involvement, we used the question "Is caring for children, your own or someone else's (without payment), included in your daily activities?" with answer options: "yes", "no", "find it difficult to answer", "no answer". The grandmothers who chose the first answer were identified as actively involved in the

process of caring for their grandchildren. Assessment of the share of such "active" grandmothers in the total number of grandmothers allowed us to adjust the forecast with regard to potential Russian grandmothers.

RESULTS

We obtained the following results in the research process.

Based on data on the age at which a woman gives birth to her first child, we modeled the age at which a Russian woman becomes a grandmother (Table 1). In Russia, as in many other European countries, the age at which a woman first becomes a mother is increasing. This, in turn, leads to an increase in the age of entry into grandparenthood.

Table 1: Estimates of the age of entry into the grandparenthood of Russian grandmothers

year	average age of a mother at the birth of her first child	year of birth of the previous genera-tion of mothers	average age of a mother at the birth of her first child in the previous generation of mothers	average age of entry into grand-parent-hood
2000	23.54	1976	23.22	46.76
2001	23.66	1977	23.18	46.84
2002	23.76	1978	23.10	46.86
2003	23.85	1979	23.03	46.88
2004	23.96	1980	22.99	46.95
2005	24.11	1981	23.01	47.12
2006	24.21	1982	23.00	47.21
2007	24.34	1983	22.96	47.30
2008	24.44	1984	22.92	47.36
2009	24.67	1984	22.92	47.59
2010	24.90	1985	22.91	47.81
2011	24.91	1986	22.95	47.86
2012	24.96	1987	22.92	47.88
2013	25.14	1988	22.90	48.04
2014	25.25	1989	22.78	48.03
2015	25.45	1990	22.65	48.10
2016	25.63	1990	22.65	48.28
2017	25.77	1991	22.60	48.37
2018	25.91	1992	22.60	48.51
2019	–	–	–	48.58*
2020	–	–	–	48.68*

* - *preliminary estimates*

It is important to note that raw statistical data allow us to simulate the age of entry into parenthood only for data until 2018 inclusive. However, increased volatility has been observed in the natural movement of the Russian population in recent years. Therefore, we considered it necessary to include the 2019-2020 data in the analysis to increase the reliability of forecast. To do this, we estimated the average age of Russian women's entry into grandparenthood for these years based on the trend model and extrapolation.

In the process of further analysis, estimates of the size of age groups of women - potential grandmothers in different years - were obtained on the basis of estimates of the age of entry into grandparenthood. The visualization of these data (Figure 1) allowed us to select the following statistical models for forecasting:

- linear trend model;
- average absolute growth model:

$$S_t = S_0 + t \times \overline{\Delta_t},$$

where $\overline{\Delta_t}$ – average absolute growth;
S_0 – the initial number of grandmothers in 2018;
S_t – projected number of grandmothers in year *t*.

Both models assume a steady annual increase in the number of Russian grandmothers. Their assessments are complementary, and they provide a more complete picture of the prospective dynamics of this age group of women.

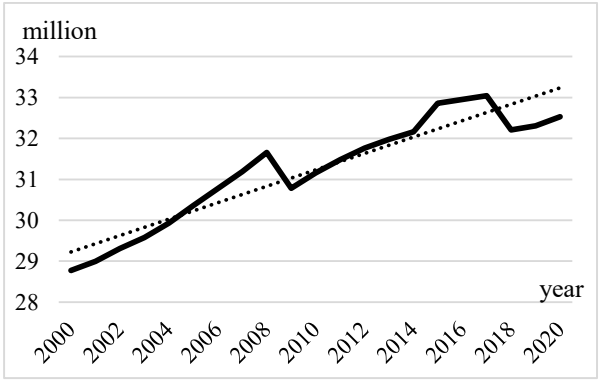


Figure 1: The number of potential grandmothers in Russia in 2000-2020 (actual data and trend)

Other models are also used in demographic statistics for the population forecast. Thus, mathematical models for forecasting the population size are most often based on exponential and logistic functions. Acceptance of the hypothesis about the model of population dynamics by exponential function implies the recognition of exponential population growth. The logistic function characterizes the growth which initially occurs at an accelerated pace, continues until a certain point, then decreases, and finally reaches zero. At the same time the trends emerging in the Russian demographic dynamics do not give us reason to consider such models suitable for forecasting the number of grandparents.

Table 2 presents the most important parameters of the forecast regression model (linear trend), and Table 3 presents parameters of the model of the average absolute growth. We also considered it necessary to obtain two estimates of such growth - for the entire study period and for the period since 2010. The former allows us to make forecasts by focusing on long-term, stable trends, while the latter allows us to consider the increased volatility of recent years and a previous

slowdown in the growth rate of the number of potential Russian grandmothers. Considering the noted features, the estimate of the average absolute growth for this period is slightly lower than the estimate for the entire period of the study.

Table 2: Model Summary and Coefficients of the Forecasting Model (Linear Trend) of the Number of Russian Grandmothers

	Constant	Years
Coefficient	-371875187.1	200549.7
T	-11.09	12.03
P-value	9.5E-10	2.5E-10
F	144.71	
Significance	2.4E-10	
R^2	0.88	

Table 3: Parameters of the regression forecast model (linear trend) of the number of Russian grandmothers

Period of model estimates	Value of the average absolute growth
from 2000 to 2020	188162.5
from 2010 to 2020	159458.0

In general, the parameters of the evaluated models allow us to forecast the growth in the number of Russian grandmothers in the next 5 years (Table 4). At the same time, a more reliable interval forecast (with a significance level of 95%) was presented on the basis of the regression model.

Table 4: Forecast of the number of Russian grandmothers

Interval forecast based on the regression model (linear trend)		
	lower bound of the forecast	upper bound of the forecast
2021	32 372 999	34 498 535
2022	32 560 455	34 712 179
2023	32 746 950	34 926 783
2024	32 932 522	35 142 311
2025	33 117 207	35 358 725
Forecast based on the model of average absolute growth		
	for data from 2000 to 2020	for data from 2010 to 2020
2021	32 722 482	32 693 777
2022	32 910 644	32 853 235
2023	33 098 807	33 012 693
2024	33 286 969	33 172 151
2025	33 475 132	33 331 609

Using the data from a survey of the Federal State Statistics Service "Comprehensive Survey of the Living Conditions of the Population" to forecast the number of active Russian grandmothers (those who are involved in the process of caring for their grandchildren daily), we obtained estimates of the share of such grandmothers among potential grandmothers (Table 5). The data show no clear trend in the dynamics of this share. Therefore, we used the average estimate in forecasting, which was 19%.

Based on the data, we made forecasts of the number of Russian grandmothers who are actively involved in the process of caring for their grandchildren (Table 6). In the next 5 years, an increase in the number of both potential grandmothers and active grandmothers is predicted.

Table 5: The number and the share of active Russian grandmothers, calculated on the basis of Comprehensive Survey of the Living Conditions of the Population (Comprehensive monitoring of living conditions 2018)

Years	The number of potential grand-mothers	The number of active grand-mothers	The share of active grand-mothers
2011	6024	1304	0.22
2014	35252	6765	0.19
2016	35879	7074	0.20
2018	34479	5462	0.16
average	27908.5	5151.3	0.19

Table 6: Forecast of the number of active Russian grandmothers

Interval forecast based on the regression model (linear trend)		
	lower bound of the forecast	upper bound of the forecast
2021	6 150 870	6 554 722
2022	6 186 486	6 595 314
2023	6 221 921	6 636 089
2024	6 257 179	6 677 039
2025	6 292 269	6 718 158
Forecast based on the model of average absolute growth		
	for data from 2000 to 2020	for data from 2010 to 2020
2021	6 217 272	6 211 818
2022	6 253 022	6 242 115
2023	6 288 773	6 272 412
2024	6 324 524	6 302 709
2025	6 360 275	6 333 006

DISCUSSIONS

The projected increase in the number of grandmothers entails the need to expand research on this category of the population. In our opinion, the studies of this topic in Russia should consider the following principles:

1) Selecting grandparents who perform grandparent' functions and participate in the care for their grandchildren from the total number of

grandparents; highlighting the labor nature of the grandparents' participation in the upbringing of their grandchildren. We propose to apply the approach to parenthood as a labor activity, which is quite common in the scientific literature, to the study of grandparenthood (Erickson 2005; Oakley 1974; Daniels 1987; Pedersen et al. 2011);

2) Applying an interdisciplinary approach to the study of grandparenthood and grandparent labor. It is advisable to study the grandparent labor from different angles: demography studies demographic characteristics of the actors of grandparent labor and demographic processes affecting them; sociology studies the motives of this type of labor, the satisfaction of various participants in the labor process, the attitudes in society; economic science studies labor costs and the organization of this type of labor, assesses the possibilities of its stimulation;

3) Differentiated study of grandparent labor by the actors of this labor – grandmothers and grandfathers. This principle is introduced due to a significant differentiation in the life expectancy of Russian men and women, which affects the duration of grandparenthood and, as a consequence, the performance of functions of the grandparent labor;

4) Considering regional differences in grandparent labor. Historically, Russia has developed a high degree of regional differentiation in many social and economic indicators (for example, Vlasov and Panikarova 2017). The regional specificity of grandparent labor may be due to the differentiation of life expectancy, fertility, divorce rate, morbidity, migration flows etc.;

5) Detailing the grandparenthood forecast in terms of age. Consideration of the age structure of grandparents can be important when assessing the potential for their participation in the upbringing and caring for grandchildren. The older the grandparents, the less they actively participate in the life of their grandchildren and the more care they require from their children. An increase in the share of grandparents of older age groups in the structure of Russian grandparents will lead to an increase in the number of the Sandwich Generation (Urick 2017). It will inevitably reduce the volume and quality of parental functions realization by the middle generation, which will simultaneously have to take care of both children and elderly parents;

6) Studying the prevalence of grandparent labor in society as one of the most important elements of the active longevity index calculated by the World Health Organization (Active ageing: A policy Framework 2002). Creating conditions for active grandparent labor in the regions of Russia would make it possible to increase the values of the active longevity index there and obtain all the positive effects that arise in society when the grandparents are more involved in the life of their grandchildren;

7) Using a set of indirect estimates to determine the main indicators of the of the demography of grandparenthood, which serve as the basis for the study of sociological and economic aspects of grandparenthood. The introduction of this principle is associated with several methodological difficulties. Firstly, Russia does not have a statistical record of direct family ties of the population that goes beyond two generations. For example, the population census does not have questions about grandchildren and grandparents. Thus, there is no "direct" way to select people with grandchildren from the older population. Secondly, it is difficult to select grandparents who perform their grandparents' functions on a regular basis from the total number of grandparents. It is also difficult to select grandparents who bear the time costs associated with grandparent labor.

CONCLUSIONS

Thus, in our study:

1) we substantiated the need to forecast the size of a special category of the Russian population - grandparents. Considering the strength of family ties in Russian families, as a hypothesis we assume that the number of grandparents who help raise grandchildren is a specific resource for increasing Russian birth rate;

2) we proposed a methodology that allows us to forecast the number of grandparents in Russia based on statistical models of a linear trend and absolute growth, in conditions of limited information resources;

3) based on the proposed methodology, we obtained estimates of the number of potential grandmothers; based on these estimates and the data of the national population survey, we forecasted the number of grandmothers who actively help their children in raising their grandchildren;

4) we proposed methodological principles for studying the phenomenon of grandparenthood, which is rarely studied in Russia.

From an economic point of view, our results may be of relevance to those organisations providing social services to elderly people when planning their activities. Forecasts may be also of interest to businesses producing goods and services for multi-generational families. Further development of the study lies in improving the models identified: firstly, the construction of models in the context of regions and age groups; secondly, verification and comparison of estimates we obtained on the basis of statistical models with estimates that will be obtained using other demographic forecasting methods (e.g. estimates derived from the age-shifting method) - after that, the final forecast will be possible; thirdly, the refinement of input parameters of the forecast models and the correction of forecast estimates based on the data of the upcoming population census in Russia and the next round of study of the population's living conditions.

We also see room for advancing our study in sociological and demographic domains. For instance, active grandparenting with its types, reasons for fulfilling, consequences, and regional diversity should be studied separately.

ACKNOWLEDGMENTS

The reported study was funded by RFBR, project number 20-011-00280.

REFERENCES

Active ageing: A policy Framework. 2002. Geneva: World Health Organization. URL: http://whqlibdoc.who.int/hq/2002/WHO_NMH_NPH_02.8.pdf (access date 12.02.2021).

Comprehensive monitoring of living conditions. 2018. Rosstat, Moscow. URL: https://gks.ru/free_doc/new_site/KOUZ18/index.html (access date 12.02.2021).

Daniels, A.K. 1987. "Invisible Work". *Social Problems*, No 34, 304-415.

Erickson, R.J. 2005. "Why emotions work matters: Sex, gender, and the division of household labor". *Journal of Marriage and Family*, No 67, 337-351.

Leopold, T. and J. Skopek. 2015. "The Demography of Grandparenthood: An International Profile". *Social Forces*, Vol 94 (2), 801-832. DOI: 10.1093/sf/sov066

Margolis, R. 2016. "The changing demography of grandparenthood". *Journal of Marriage and Family*, No 78, 610-622.

Margolis, R. and A.M. Verdery. 2019. "A Cohort Perspective on the Demography of Grandparenthood: Past, Present, and Future Changes in Race and Sex Disparities in the United States". *Demography*, Vol 56 (4), 1495-1518, DOI: 10.1007/s13524-019-00795-1

Margolis, R. and L. Wright. 2017. "Healthy grandparenthood: How long is it, and how has it changed?" *Demography,* No 54, 2073-2019.

Oakley, A. 1974. *"The sociology of housework*". New York: Pantheon.

Pedersen, D.E., Minnotte, K.L., Susan, E. and G. Kiger. 2011. "Exploring the relationship between types of family work and marital well-being". *Sociological Spectrum*, No 31, 288-315.

Raišienė, A.G., Bilan, S., Smalskys, V. and J. Gečienė. 2019. "Emerging changes in attitudes to inter-institutional collaboration: The case of organizations providing social services in communities". *Administratie si Management Public,* No 33, 34-56. DOI: 10.24818/amp/2019.33-03.

Russian Statistical Yearbook-2019 (Appendix). Rosstat. URL: https://rosstat.gov.ru/folder/210/document/13396 (access date: 12.02.2021)

The Human Fertility Database. URL: https://www.humanfertility.org/cgi-bin/main.php (access date: 12.02.2021).

Tkatch, O. 2015. "Caring Home": Caring for Elderly Relatives and Problems of Living Together". *Sociological research,* Issue 10, 94-102.

Urick, M. J. 2017. "The Aging of the Sandwich Generation". *Generations-Journal of the American society on aging,* Vol 41 (3), 72-76.

Vlasov, M and S. Panikarova. 2017. "Characteristics of the economic development of the multi-ethnic regions of Russia". In *Proceedings of the 5th International Conference on Management Leadership and Governance ICMLG 2017* (Johannesburg, South Africa, March 16th-17th, 2017), 472-477.

Zelikova, J. 2020. "I Can Only Perceive Myself as a Babushka": aging, ageism, and sexism in contemporary Russia". *Laboratorium: Russian Review of Social Research,* 12(2), 124-145, DOI: 10.25285/2078-1938-2020-12-2-124-145. URL: https://soclabo.org/index.php/laboratorium/article/view/966/2403

Yahirun, J. J., Park, S. S. and J. A. Seltzer. 2018. "Step-grandparenthood in the United States". *Journals of Gerontology, Series B: Psychological Sciences & Social Sciences,* No 73, 1055–1065.

AUTHOR BIOGRAPHIES

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 the sociology of labour. She is a doctoral supervisor and a member of the International Sociological Association. Her email address is a.p.bagirova@urfu.ru and her webpage can be found at http://urfu.ru/ru/about/personal-pages/a.p.bagirova/

OKSANA SHUBAT is an Associate Professor of Economics at Ural Federal University (Russia). She received her PhD in Accounting and Statistics in 2009. Her research interests include demographic processes, demographic dynamics and their 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 webpage can be found at http://urfu.ru/ru/about/personal-pages/O.M.Shubat/

FACTOR MODELING OF RUSSIAN WOMEN'S PERCEPTIONS OF COMBINING FAMILY AND CAREER

Natalia Blednova
Anna Bagirova
Ural Federal University
Graduate School of Economics and Management
Mira st., 620002, Ekaterinburg, Russia
E-mail: n.d.blednova@urfu.ru
E-mail: a.p.bagirova@urfu.ru

KEYWORDS
Career, motherhood, Russian women, factor analysis, correlation analysis

ABSTRACT

Sociologists and demographers explain late childbearing by the transformation of the life values of modern women. This is considered as one of the reasons for the decline in the birth rate. Our study aims to reveal perceptions of the relationship between career and family in the life strategies of working Russian women by using factor analysis.
We collected data in a sociological survey of working women living in the Ural region. We asked respondents to rate 10 statements about work, family and children. We constructed 3-factors model of Russian women's perceptions of combining family and career. Then we used correlation analysis to assess the relationship between these factors and the social and demographic parameters of the respondents.
We concluded that the use of factor analysis made it possible to model a wide range of Russian women's perceptions of combining family and career.
Considering the results obtained may contribute to improving the regulation of interaction of two important societal spheres of professional and parental activities and create conditions for increasing the birth rate in Russia.

INTRODUCTION

The average age of a mother at childbirth has been increasing in several developed countries in recent decades. According to Eurostat, the average age at birth of the first child in Europe was 26 in 1980s, but it had increased to 30.8 by 2020 (Eurostat: Mean age of women at childbirth and at birth of first child). Russia follows the European trend. In 1990, the average age of a mother at childbirth was 25.3, but it had increased to 28.7 by 2018 (Rosstat: Average age of a mother at childbirth).
According to sociologists and demographers, late childbearing can be explained by the transformation of the life values of modern women, who focus on meeting personal needs for professional development, education and achieving a certain level of financial well-being (Savinov et al. 2020; Goldstein et al. 2009; Lesthaeghe, 2010). Many women prefer work to family when there are clear career prospects, the child in this case can be perceived as an obstacle for further professional self-realization (Betz 1993; Metz and Tharenou 2001; Procter 1998). Interestingly, employers still pay particular attention to candidate's family status. Those having children are often rejected (Doris and Oliver, 2019; Miller, 2019; Henle et al., 2020).
The work-family balance issue results in more stress, which mothers feel in a competitive labour market (Nair et al., 2019; McCanlies et al., 2019). Being involved in both professional and parental labour, women adjust their life strategies to mitigate the work-family conflict (Borgmann et al., 2019). When employed, women tend to take specific attitude to reproduction—they more often choose the smaller number of children (Greulich et al., 2017; Cools et al, 2017). However, some researchers note the positive impact of having children on the overall well-being of parents. In particular, children have a beneficial effect on the emotional state of mothers and help them cope with problems and psychological distress more easily (Rao 2020).
We used factor analysis to examine women's perceptions of the impact of having children on a career. It is often used in social sciences (for example, Popov et al. 2018; Bork and Moller 2018; Lifshits and Neklyudova 2018). We chose this statistical modeling method because we sought to use the answers to direct questions to identify those latent variables that may affect the reproductive decisions of Russian women. It allowed a wide variety of variables to be used during the initial measurement, and then it reduced the dimension of the problem by switching to latent variables. The purpose of our paper is to reveal perceptions of the relationship between career and family in the life strategies of working Russian women by using factor analysis.

DATA AND METHODS

1. We used data collected in a sociological survey of working women living in the Ural region. We surveyed them in February-April 2020. For our analysis we filtered the answers of women aged 18-45 with children - there were almost 200 such women among the respondents. The sample was calculated according to the data of the All-Russian Population Census and Regional Official Statistics.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

2. To study the mothers' perceptions of the impact of having children on their careers and the children's place in the life strategies of women, we asked respondents to rate 10 statements about work, family and children on a five-point scale:

– V1 – Workers with children have a higher financial position than workers without children;
– V2 – It is easier to advance the career after having a child;
– V3 – Job seekers with children are in demand in the labor market;
– V4 – A worker with children is less likely to find a job than a worker without children;
– V5 – It is more difficult to advance the career after having a child;
– V6 – Workers with small children cannot expect high wages;
– V7 – It is more difficult for workers with children to fulfill themselves in life;
– V8 – Workers with children are more stressed than workers without children;
– V9 – A person must build a career first, then have a family and children;
– V10 – The most important thing in life is family and children, career is secondary.

3. To assess the suitability of variables for factor analysis, we used Bartlett's Test of Sphericity and Kaiser-Meyer-Olkin Measure of Sampling Adequacy. We used principal component analysis as the extraction method. We determined the number of factors using two methods: Kaiser Criterion and Scree Plot. The orthogonal rotation of the components was determined by the Varimax solution. We assessed the relationship between the obtained components and respondents' social and demographic characteristics by using the Spearman correlation, scatter plots and boxplots. We used SPSS 22.0.

RESULTS

1. Means, medians, mode and standard deviations pertaining to the ten variables of interest are presented in Table 1.

Table 1: Descriptive statistics

Variables	Mean	Median	Mode	Std. Deviation
V1	2.05	2	1	1.079
V2	2.07	2	1	1.046
V3	2.01	2	1	1.065
V4	3.35	3	5	1.414
V5	3.20	3	5	1.470
V6	2.69	3	1	1.447
V7	2.51	2	1	1.346
V8	2.98	3	3	1.355
V9	2.33	2	1	1.224
V10	3.92	4	5	1.188

2. The correlation matrix is presented in Table 2. It shows the relationship between the respondents' assessments which indicate the positive impact of having children on a career. For example, the share of women who believe that workers with children are in demand in the labor market among those who believe that it is easier to advance the career after having a child. On the other hand, we can see the strongest correlation between the assessments of statements that view children as a hindering factor in building a career. For example, the share of women who believe that it is more difficult to advance the career after the birth of a child is higher among those who believe that workers with small children cannot expect high wages. In addition, an idea that employees with children will not receive high remuneration is related to an opinion that children interfere with self-realisation.

Table 2: Correlation matrix of the respondents' opinions about family and career

	V1	V2	V3	V4	V5	V6	V7	V8	V9
V2	.38**	-							
V3	.32**	.42**	-						
V4	-.01	-.16**	-.19**	-					
V5	.03	-.09	-.10	.46**	-				
V6	.11*	.03	.00	.39**	.54**	-			
V7	.03	.01	.02	.39**	.54**	.49**	-		
V8	.07	-.06	.05	.30**	.37**	.24**	.35**	-	
V9	.05	.02	.02	.17**	.16**	.12*	.12*	.15**	-
V10	.11*	.07	.10	.06	.02	.12*	.03	.02	-.33**

*p<0.05. **p<0.01.

3. The possibility of using factor analysis for the ten indicated variables was confirmed by two statistical tests: Bartlett's Test of Sphericity, which is calculated from the sample data, is 477.022 (df = 45; $\alpha = 0.000$); Kaiser-Meyer-Olkin Measure of Sampling Adequacy is 0.724. Table 3 presents an analysis of the communalities. There are no values close to zero, therefore, we included all variables in factor analysis.

Table 3: Communalities

Variables	Initial	Exrtraction
Workers with children have a higher financial position than workers without children	1.000	.577
It is easier to advance the career after having a child	1.000	.729
Job seekers with children are in demand in the labor market	1.000	.648
A worker with children is less likely to find a job than a worker without children	1.000	.552
It is more difficult to advance the career after having a child	1.000	.729

Workers with small children cannot expect high wages	1.000	.608
It is more difficult for workers with children to fulfill themselves in life	1.000	.629
Workers with children are more stressed than workers without children	1.000	.365
A person must build a career first, then have a family and children	1.000	.709
The most important thing in life is family and children, career is secondary	1.000	.749

Extraction Method: Principal Component Analysis.

4. Data from both Table 4 and Scree Plot (Figure 1) indicate the advisability of identifying three components.

Table 4: Total Variance Explained

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.956	29.561	29.561	2.956	29.561	29.561
2	2.034	20.340	49.901	2.034	20.340	49.901
3	1.305	13.053	62.954	1.305	13.053	62.954
4	.801	8.009	70.962			
5	.771	7.708	78.670			
6	.575	5.752	84.422			
7	.457	4.568	88.990			
8	.403	4.035	93.025			
9	.368	3.678	96.703			
10	.330	3.297	100.00			

Extraction Method: Principal Component Analysis.

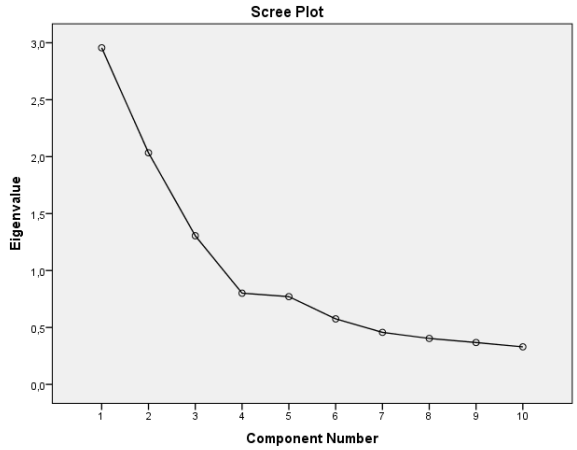


Figure 1: Scree Plot

5. Table 5 shows the loadings of all variables for three factors. Each variable is uniquely associated with only one of three factors.

Table 5: Rotated Component Matrix

	Component		
	1	2	3
Workers with children have a higher financial position than workers without children	.109	**.745**	-.099
It is easier to advance the career after having a child	-.103	**.847**	.027
Job seekers with children are in demand in the labor market	-.064	**.802**	.016
A worker with children is less likely to find a job than a worker without children	**.724**	-.169	.019
It is more difficult to advance the career after having a child	**.850**	-.056	.051
Workers with small children cannot expect high wages	**.775**	.084	-.032
It is more difficult for workers with children to fulfill themselves in life	**.792**	.042	.011
Workers with children are more stressed than workers without children	**.596**	.019	.100
A person must build a career first, then have a family and children	.248	.160	**.789**
The most important thing in life is family and children, career is secondary	.116	.212	**-.831**

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
Rotation converged in 5 iterations.

6. As a result of factor analysis, we obtained the following three components (Figure 2):

– Component 1: perception of a child as a factor hindering professional activity (29.6% from the total variance);

– Component 2: perception of a child as a factor motivating professional activity (20.3% from the total variance);

– Component 3: perception of a child's place in the life strategy of women (13.1% from the total variance).

7. We revealed significant correlation between the values of components and the social and demographic characteristics of the respondents (Table 6). Figure 3 presents diagrams showing the discovered relationships.

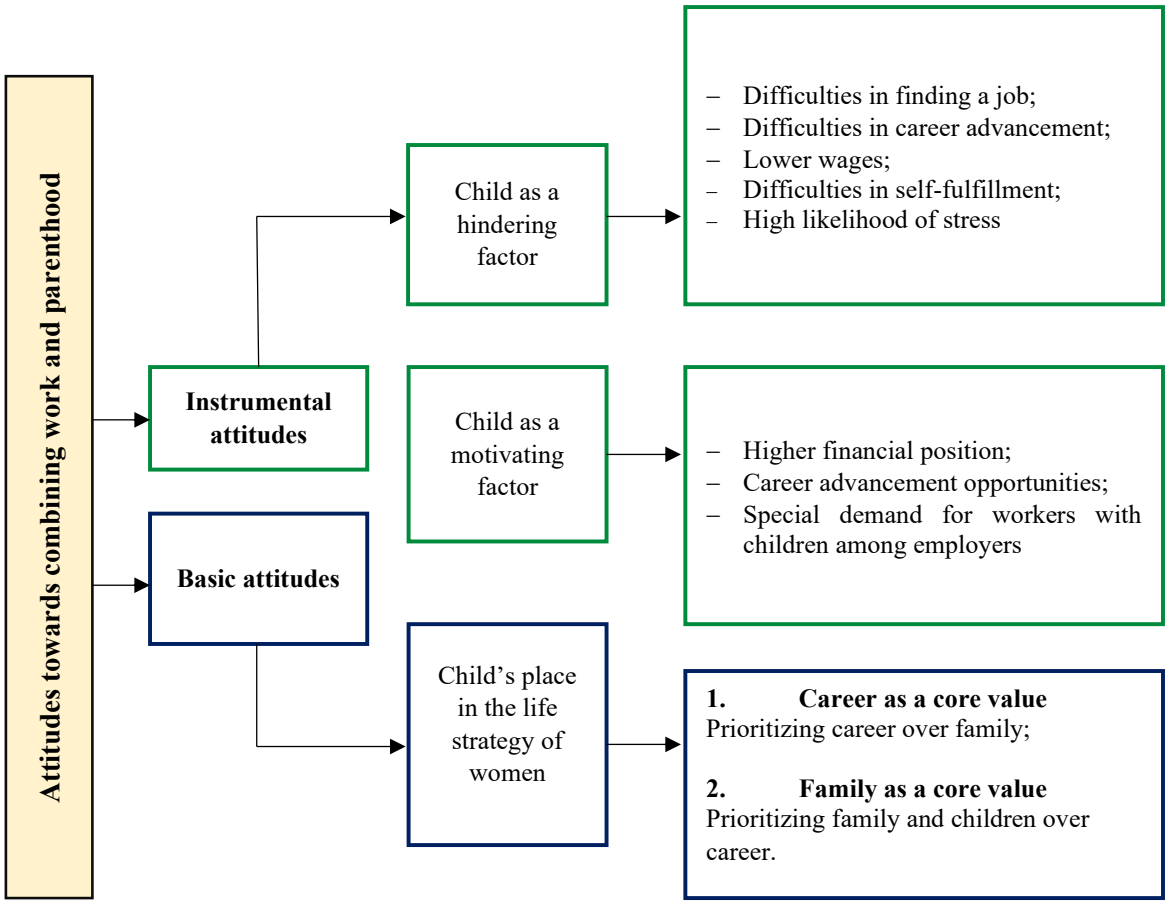


Figure 2: Factor analysis results

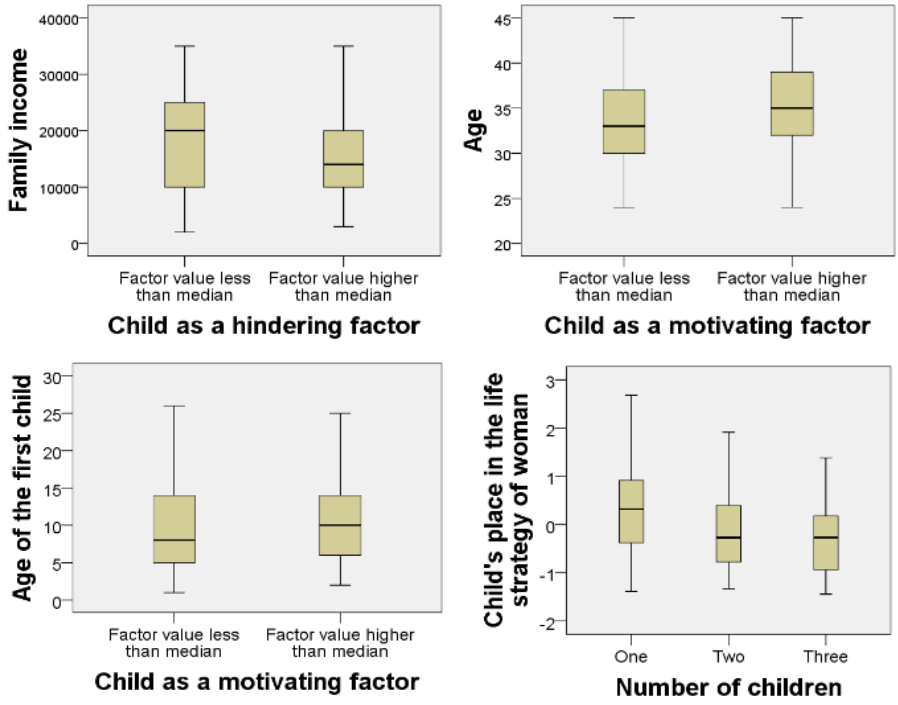


Figure 3: Diagrams showing the relationship between social and demographic characteristics of respondents and the values of factors

Table 6: Spearman's ρ between components and social and demographic characteristics of respondents

Components	Social and demographic characteristics of respondents			
	Income	Age of respondents	Age of first child	Number of children
Component 1: perception of a child as a hindering factor	-.270**			
Component 2: perception of a child as a motivating factor		.193**	.178*	
Component 3: perception of a child's place in the life strategy of women				-.197*

*p<0.05. **p<0.01

DISCUSSIONS

The results allow us to establish the following influence of having children on women's life attitudes:

1. Women with low income are more likely to perceive children as a hindering factor. Most likely, this is due to the large material costs that are required to raise a child and provide him or her with all the necessary things. For parents with low income, having a child creates strong financial barriers and restrictions.
2. With age, women more often perceive children as a motivating factor for new achievements. Perhaps this is due to the transformation of the person's basic attitudes over time. A woman who has just had a child does not yet have the experience of life with children, therefore she cannot realize all the benefits of motherhood. As the child grows up, the mother gains parenting experience and reevaluates the opportunities that open up after having a child. Notably, we found a positive correlation between the Component 2 value and the age of the respondent's first child. This means that our hypothesis may be correct: as the child grows up, women are more likely to perceive children as a motivating factor.
3. The more children there are in the family, the more likely it is for women to prioritize family and parenting over professional growth. There are various causal relationships in this case. On the one hand, an increase in the number of children can change the person's life values. Perhaps when there are more than one or two children in a family, parents begin to perceive them as a source of their well-being. Mothers with one child, on the contrary, want to realize themselves not only in the family, but also in the professional sphere. As a result, they prefer a small number of children to be able to realize their career goals. On the other hand, a basic attitude of prioritizing family over career can lead to reproductive decisions associated with having more children. Perhaps women who have a family as a basic life value initially plan to have more than two children, sharing the opinion that a happy family should have many children.

CONCLUSIONS

The main conclusions of our study are as follows:

1. The use of factor analysis made it possible to model a wide range of Russian women's perceptions of combining family and career. This analysis helped us to identify groups of opinions that indicate the multidirectional influence of having children on the mother's career. Further correlation analysis showed which social and demographic characteristics of women are associated with these different assessments.
2. The results of our analysis showed that it was mostly young parents who perceive children as a hindering factor of professional activity. With age, the attitude towards children and career changes – people prioritize family and parenting over professional self-realization. The implementation of this result is as follows: improving the regulation of interaction of two important societal spheres of professional and parental activities will create conditions for increasing the birth rate in Russia. This can be achieved through the realization of state and corporate demographic policies aimed at stimulating the employment of workers with children.

We see the development of our research in: applying nonlinear factor analysis based on a neural network; analyzing the relationship between components and the strength of women's reproductive attitudes; seeking other factor models, which will help explain how children affect women's career; including working fathers in the study, as well as in the comparative analysis of the results of factor analysis of working mothers and fathers.

REFERENCES

Betz, N. 1993. Women's career development. In F. L. Denmark & M. A. Paludi (Eds.), *Psychology of women: A handbook of issues and theories* (pp. 625–684). Westport, CT.

Borgmann, L.S, Kroll, L.E, Müters, S., Rattay, P. and Lampert, T. 2019. "Work-family conflict, self-reported general health and work-family reconciliation policies in Europe: Results from the European Working Conditions Survey 2015". *SSM Popul Health*, 9, 100465, doi: 10.1016/j.ssmph.2019.100465.

Bork, L. and Moller, S.V. 2018. "Housing Price Forecastability: A Factor Analysis*". REAL ESTATE ECONOMICS, 46*, 3, 582-611.

Doris, H. and Oliver, L. 2019. "Job insecurity and parental well-being: The role of parenthood and family factors". *Demographic Research*, 40, 897–932 doi: 10.4054/DemRes.2019.40.31.

Goldstein, J.R., Sobotka, T. and A. Jasilioniene. 2009. "The end of 'lowest-low fertility?' ". *Population and Development Review*, 35(4), 663-699.

Greulich, A., Guergoat-Larivière, M. and Thévenon, O. 2017. "Emploi et deuxième naissance en Europe". *Population*, 72, 625-647, doi: 10.3917/popu.1704.0653.

Henle, C.A., Fisher, G.G. and Clancy, R.L. 2020. "Eldercare and Childcare: How Does Caregiving Responsibility Affect Job Discrimination?". *Journal of Business and Psychology*, 35(1), 59-83 doi: 10.1007/s10869-019-09618-x.

Lesthaeghe, R. 2010. "The unfolding story of the Second Demographic Transition". *Population and Development Review*, 36(2), 211-251.

Lifshits, M. L. and N.P. Neklyudova. 2018. "Factor Analysis Reflecting the Impact of Labor Migration on the Spread of Socially Dangerous Diseases in Russia". *Economic and Social Changes-facts Trends Forecast*, 11, 6, 229-243.

McCanlies, E., Mnatsakanova, A., Andrew, M., Violanti, J. and Hartley, T. 2019. "Child care stress and anxiety in police officers moderated by work factors". *Policing: An International Journal,* 42(6), 992-1006, doi: 10.1108/PIJPSM-10-2018-0159

Mean age of mothers at childbearing (years). In *The Demographic Yearbook of Russia 2019*. Moscow: Rosstat. URL: https://gks.ru/bgd/regl/B19_16/Main.htm. (access date 25.01.2021).

Mean age of women at childbirth and at birth of first child. 2020. Eurostat. URL: https://ec.europa.eu/eurostat/web/population-demography-migration-projections/data/main-tables (access date: 25.01.2021).

Metz, I. and P. Tharenou. 2001. "Women's career advancement". *Gender and Organizational Management*, 26, 312–342.

Miller, A. 2019. "Stereotype threat as a psychological feature of work–life conflict". *Group Processes & Intergroup Relations*, 22(2), 302–320, doi: 10.1177/1368430217735578.

Nair, D., Millath, M. 2019. "Identifying family-work conflict among employees of the travancore cements limited, Kottayam, Kerala". *International Journal of Recent Technology and Engineering (IJRTE),* Vol. 8, 2S6, 718–726, doi:10.35940/ijrte.B1135.0782S619.

Popov, E., Simonova, V. and M. Maksymchik. 2018. "Factor model of the network capacity of a firm". In *Proceedings of the 14th European Conference on Management, Leadership and Governance ECMLG 2018* (Utrecht, Netherlands, October 18th-19th, 2018), 221-230.

Procter, I. and M. Padfield. 1998. "*Young Adult Women, Work and Family: Living a Contradiction*". London & Washington: Mansell Publishing Ltd.

Savinov, L., Solovyeva, T., Bistyaykina, D. and A. Karaseva. 2020. "Socio-cultural determination of late fertility measures and family-demographic policy of birth rate (On materials of the Republic of Mordovia)". *Woman in Russian Society*, 1, 101-112, doi: 10.21064/WinRS.2020.1.8

Rao, A. 2020. "From Professionals to Professional Mothers: How College-educated Married Mothers Experience Unemployment in the US". *Work, Employment and Society*, Vol. 34, 2, 299-316, doi: 10.1177/0950017019887334.

ACKNOWLEDGMENT

The article is one of the outputs of the research project "Russian pro-natalist policy: resources, effects, optimization opportunities", supported by the Council on grants of the President of the Russian Federation, project no. NSh-2722.2020.6

AUTHOR BIOGRAPHIES

NATALIA BLEDNOVA is an analyst at the Center for Regional Economic Research of Ural Federal University (Russia). Her research interests include balancing professional and parental labour. Her email address is n.d.blednova@urfu.ru

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 the sociology of labour. She is a doctoral supervisor and a member of the International Sociological Association. Her email address is a.p.bagirova@urfu.ru and her webpage can be found at http://urfu.ru/ru/about/personal-pages/a.p.bagirova/

CLEARINGHOUSES VERSUS CENTRAL COUNTERPARTIES FROM MARGIN CALCULATION POINT OF VIEW

Melinda Friesz
Department of Finance
Corvinus University of Budapest
Fővám square 8. Budapest, 1093, Hungary
KELER Ltd
Rákóczi street 70-72. Budapest, 1074, Hungary
E-mail: szodorai.melinda@keler.hu

Kata Váradi
Department of Finance
Corvinus University of Budapest
Fővám square 8. Budapest, 1093, Hungary
E-mail: kata.varadi@uni-corvinus.hu

KEYWORDS

Initial margin, maintenance margin, variation margin, margin call

ABSTRACT

Clearinghouses and central counterparties (CCPs) have a notable role in financial markets, namely facilitating securities trading and derivative transactions on exchanges and over-the-counter markets. They have to clear the transactions and carry out their settlements to decrease costs and settlement risk. To efficiently carry out this activity, they need to collect adequate collateral from the trading parties as guarantees. Two main elements of these guarantees are the margin requirement and default fund contribution. Our paper focuses on the margin calculations and emphasizes their notable difference in the case of clearinghouses and CCPs. Our main result is that clearinghouses' margin requirement is better from a procyclicality point of view; however, CCP margining is more prudent based on our results.

INTRODUCTION

Following the global financial crisis (GFC) of 2007-2009, the attention turned to the over-the-counter (OTC) derivatives markets – trades outside the exchanges – and the risks associated with them. OTC transactions carried a considerable counterparty risk during the GFC, which had transformed easily into a systemic risk throughout large financial institutions' bankruptcy, e.g., the Lehman Brothers. As a result of the GFC, regulators realized the importance of decreasing counterparty risk during trading. Counterparty risk can be managed through clearing bilaterally or centrally (Gregory, 2014). Bilateral clearing means that the two trading partners enter into a master agreement without a CCP covering all of their trades. This agreement has an annex, called the credit support annex (CSA), that requires both parties to provide collateral (Hull, 2018). While central clearing means that every trading partner is trading with the CCP, as shown on the right-hand side of Figure 1, the left-hand side shows the absence of a CCP, representing the bilateral clearing case.

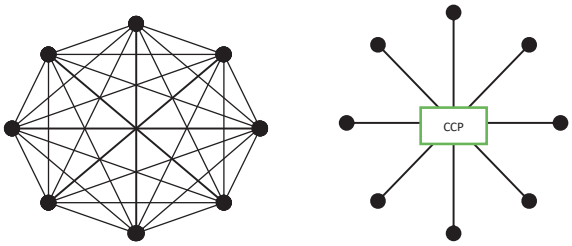


Figure 1: Bilateral versus central clearing on OTC markets (Hull, 2018, pp. 57.)

During the summit in Pittsburgh on 26th September 2009, the participating G20 leaders were in unison that all standardized OTC derivative contracts should be cleared through CCPs by the end of 2012. The other aim of the financial reform was to enhance transparency, as well as to make these contracts reported to trade repositories (EMIR (5), 2012). Finally, the capital requirements should be higher for the non-centrally cleared OTC derivatives (Gregory, 2014). As a result of this agreement, the Dodd-Frank Wall Street Reform and Consumer Protection Act (DFA) enacted in July 2010 in the USA, while in July 2012, the European Market Infrastructure Regulation (EMIR) in the European Union (EU).

In April 2012, the Principles of Financial Market Infrastructures (PFMI, 2012) was issued by the Committee on Payment and Settlement Systems and the International Organization of Securities Commissions (CPSS-IOSCO), which became the global benchmark for the regulatory requirement for CCPs (RTS (2), 2013). As a result, all of the standardized OTC trades between financial institutions must be cleared through a CCP. This regulation's exceptions are the non-financial institutions if their position does not exceed the clearing threshold (Doyle et al. 2016) in the framework of the EU regulation; while in the case of the US regulation, the non-financial firms are exempted if their transaction is entered in order to hedge commercial risk (Gregory, 2014). Finally, the foreign exchange transactions are exceptions, too (Hull, 2018).

Based on the DFA and the EMIR, the CCPs should manage a multilevel guarantee system. The traders have to pay two significant types of guarantees within this guarantee system, the margin requirement and the

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

default fund contribution, which is called guaranty fund in the case of the clearinghouses. Our primary focus in this paper is the margin requirement calculation. We will analyze how clearinghouses and CCPs are calculating the value of margin, what the differences are, and what the similarities are. We will also explain the different margin-related notions, like the variation margin, initial margin, maintenance margin, and the margin call.

LITERATURE

The primary role of clearinghouses and central counterparties is the clearing and the settlement of trades. A clearinghouse is operating on exchange markets, while CCPs can operate on exchange markets and also on OTC markets. The main difference between the two institutions is that the CCP takes over the counterparty risk during trading, namely becomes the seller to every buyer, and the buyer to every seller, while a clearinghouse usually does not do this. So in the case of OTC CCPs, the two trading parties are, therefore, no longer exposed to each other, but only to the CCP, which provides insurance against bilateral default risk (Biais et al., 2016). Also, a CCP always nets transactions, while clearinghouses not necessarily. The final difference is that in OTC CCPs, the trades are not necessarily cleared daily (Berlinger et al., 2016). So we can state that every CCP is a clearinghouse, while not every clearinghouse can be regarded to be a CCP (DNB, 2013).

There are several margin notions related to trading: initial margin (IM), maintenance margin (MM), variation margin (VM), and margin call. These notions have a different meaning, depending on which margin we are talking about: securities margin, futures margin, or the CCP margin. Securities margin means a partial downpayment – usually up to 50%, regulated by Regulation T (Reg T, 2021) – of the financial asset's price. The trader has to pay this amount to his broker and borrow the remaining amount also from his broker to buy the financial assets. This is what is called "buying on margin" (CFA, 2017). The notion of securities margin is different from the margin notions we are analyzing in this paper, although it is also related to trading with the financial assets but has nothing to do with the clearing activity.

Futures margin is the margin calculated by clearinghouses only in exchange trading, while the CCP margin is calculated by CCPs and can also be used in exchange- and OTC trading. In both cases, the trader (a.k.a. the clearing member) has to pay both the variation margin and the initial margin.

The initial margin aims to cover the potential closeout cost of the traders' positions to cover potential future costs a CCP or a clearinghouse may face in normal market conditions if the trader defaults. The value of the IM is usually based on a risk measure. For example, in the EMIR framework, the CCP IM on the OTC market should be enough to cover losses on a 99,5% significance level, with a 5-day liquidation period, while for the exchange-traded asset, 99% significance level, and 2-day liquidation period should be applied. The model parameters should be estimated from a 12 months look-back period, which contains a stressed time period (EMIR, 2012, RTS, 2013). The DFA is not as detailed as the EMIR regarding the IM model's parameters, and it just quantifies the application of the 99% significance level (SEC, 2021a). Moreover, another notable difference between the two regulations from the IM point of view is that EMIR emphasizes that IMs should not be procyclical, while DFA does not.

The process of how the IM is handled in the futures margining and the CCP margining is different. Regarding the futures margining, it has to be paid before the trade is entered, so the trader cannot start trading without it (CME Group, 2021a). Meanwhile, in the CCP margining case, the initial margin has to be paid after the first trading day is over, when the transactions are cleared (Hull, 2018).

Meanwhile, the initial margin is not applied to the bilateral clearing in most of the trades. Basel Committee on Banking Supervision IOSCO (BCBS-IOSCO) (2015) states that the total amount of initial margin on bilateral transactions not cleared by a CCP represents only 0.03% of the gross notional exposure in 2012. In 2011 a Working Group on Margin Requirements (WGMR) had been formed by the BCBS, the IOSCO, the CPSS, and the Committee on Global Financial Systems (CGFS) to work out a margining framework on the bilateral trades as well. Based on their work, financial firms and systematically important non-financial institutions should use initial margins above a certain threshold, applying a 99% significance level and a 10-day liquidation period (BCBS-IOSCO, 2015).

Variation margin has to be paid after the mark-to-market valuation of the open positions, so after the losses/gains the trader had on a certain trading day. In the case of CCP margin, if the trader gains on his open position on a particular day, he has access to this amount and could withdraw from his collateral account. Contrary, if the trader loses on his position, he has to pay this loss to the CCP as a variation margin. It is important to note that in bilateral trading without a CCP, this variation margin also has to be paid. So the losses will increase the amount of collateral the trader has to pay, while the gains will decrease is.

In the futures margin case, the clearing of the actual loss/profit works differently. If the trader has a loss/profit when his position is being marked-to-market, this loss/profit is being subtracted/added to his actual margin account balance. In case the loss is so extensive that this balance falls to the level – or below – the so-called maintenance margin level, the trader will get a margin call. A margin call means that the margin balance level has to be increased to the level of the initial margin. In sum, the variation margin is the amount that is needed to increase the margin account balance to the initial margin level (CME Group, 2021a). In the case of the CCP margin, the notion of maintenance margin is not applied.

The notions we have described here are the basics of futures margining and CCP margining on the individual financial assets' level. The actual margin calculation process can be different in the case of clearinghouses and CCPs. For example, the parameters of the IM model can change, or the risk measure they use. Also, the application of the calculated IM value to define the final collateral value may differ. An example of this is the CME Group's Standard Portfolio Analysis of Risk (SPAN) methodology, which defines the margin on a portfolio level (CME Group, 2021b). Several CCPs are using this approach besides CME Group, e.g., KELER CCP Ltd, the CCP of Hungary (KELER CCP, 2021).

Pros and cons of central clearing

Bilateral and central clearing are similar in that both are netting the transactions, and both use margining for the same purposes. However, in several other aspects, they are very different, which we aim to show by highlighting the advantages and disadvantages of CCP clearing from the perspective of the CCP.

The most important advantages of central clearing through CCPs are transparency, offsetting, loss mutualization through the default fund, legal and operational efficiency, liquidity and default management (Gregory, 2014), risk mitigation, and capital efficiency (ICE, 2021). While the main disadvantages include moral hazard; adverse selection; bifurcation; procyclicality; assets are less effective for hedging if they have to be centrally cleared; more costly than bilateral clearing without a CCP; only highly liquid assets can be used as collaterals. At the same time, several factors are undecidable whether they are advantages or disadvantages of central clearing. The most important is the CCPs contribution to systemic risk. Interoperability among CCPs increases risk without enhancing the financial resources of each interoperating CCP (Turing (2012)). Duffie and Zhu (2011), Amini et al. (2016), Lopez and Saeidinezhad (2017), Health et al. (2016), and several others are primarily concerned with the potential for contagion due to their high level of interconnectedness. King et al. (2020) address the problem from a procyclicality point of view: CCP resource demands are inherently procyclical concerning the market, thus threatening the ability of CCPs to fulfill their obligations in stressful periods. Gregory (2014) states that CCPs convert counterparty risk into liquidity, operational and legal risk.

Other vital reasons against central clearing through CCPs on the OTC markets are the assets maturity, liquidity, and complexity. Although the CCPs are efficient in handling counterparty risk on futures and spot markets, where usually the positons are open for only a short period of time (weeks to months), CCPs are not as efficient in case of the OTC market's assets, which are financial assets that usually has a considerable maturity which can even last for decades (Gregory, 2014). For example, a ten-year credit default swap is not uncommon in these markets (Murphy, 2013). Moreover, the exchange-traded assets are standardized, not too complicated, and liquid. So handling counterparty risk on OTC markets, where the assets are complex, traded volumes are not concentrated in highly liquid assets, is inefficient and too expensive to clear through a CCP. For example, in a stressed market condition to close down a position can take some days because of illiquidity. It can also happen that for non-standardized and exotic OTC derivatives, central clearing is just not feasible. However, the most important reason against CCP central clearing is that the OTC markets are the central place of financial innovations and offer cost-effective and well-tailored risk reduction products. Nevertheless, these new, non-standardized, or exotic products cannot be cleared by CCPs (Gregory, 2014).

The most convincing reason for the CCP central clearing is the case of the global financial crisis where the Lehman default on 15th September 2008 was the biggest default in CCP history (Fleming and Sarkar, 2015; Bernstein et al., 2019). LCH.Clearnet's SwapClear service provided nearly half of the world's interest rate swap positions at the time of the default. It could handle the default of Lehman efficiently within hours, by suspending insolvent Lehman entities and by having around USD 2 billion as initial margin account from Lehman already by that time (Gregory, 2014, Norman, 2011). LCH.Clearnet faced massive failures before as well, like the default of Barings in 1995, which it could also handle without any severe problems (Gregory, 2014). These examples show the most critical advantages of CCPs, and how shock resistant they are. The critical nature of their role was endorsed in September 2018, when the default of a clearing member at the Swedish CCP, Nasdaq Clearing, reached losses causing the use of the CCP's prefunded buffer, and surviving members were required to replenish USD 107 million of that buffer within a few days (King et al. 2020). The event did not end with the mass default of the market participants.

Clearing in the USA and the EU

In the USA, the clearing services are provided by a few major actors:

- Subsidiaries of the Depository Trust & Clearing Corporation (DTCC), which are the world's largest clearinghouses (DTCC, 2021), both administered and supervised by the US Securities and Exchange Commission (SEC) (CFI, 2021): 1) National Securities Clearing Corporation (NSCC) and 2) Fixed Income Clearing Corporation (FICC).
- Option Clearing Corporate (OCC): the world's largest equity derivatives clearing organization. It is under the jurisdiction of the SEC and the CFTC (Commodities Futures Trading Commission) because it is also registered as a derivatives clearing organization (DCO) (OCC, 2021).
- CME Clearing: it is a subsidiary of the CME Group Inc. exchange and clears and settles exchange-traded futures and options contracts, and also OTC

derivative contracts (BIS, 2012). Since it is considered a DCO, it is regulated by the CFTC.
- ICE Clear Credit LLC is an ICE subsidiary (Intercontinenal Exchange Inc.) and the world's largest credit default swap clearinghouse. The CFTC and SEC regulate it (ICE Clear Credit, 2021).
- There are other domestic (e.g., MGE Clearing, New York Portfolio Clearing) and foreign clearing which operate as DCOs (e.g., LCH.Clearnet Ltd.).

Besides the already mentioned CCPs, the following belong under the SEC's supervision: ICE Clear Europe Limited and the LCH SA (SEC, 2021b).

There are 19 CCPs in the EU from 15 different countries (EACH, 2021) with the EMIR recognition for providing clearing and settlement services on exchanges and OTC markets in the European Union. The supervisor for all of them is their national supervisor, e.g., the national bank and the European Securities Market Authority (ESMA) on the EU level. It is important to note that not only an EU-based country can get EMIR recognition (ESMA, 2021). The CCP just has to prove that its operation and risk management process is prudent enough. Moreover, the European Commission and the CFTC agreed on a common approach to cross-border processes. The announcement made on 10th February 2016 permits DCOs and CCPs to clear derivatives for counterparties abroad. (Doyle, 2016).

MODEL

In our simulation, our main goal was to show how the margin calculation in the case of the futures- and CCP margining work for a stock position. We used the following assumptions for the simulation:

The logreturn of the stock follows arithmetic Brownian motion (ABM) based on Equation 1.

$$dY = \alpha \cdot dt + \sigma * \sqrt{dt} \cdot N(0,1) \qquad (1)$$

where ′dY′ is the change in the logreturn during ′dt′ period, $'\alpha'$ is the expected value of the logreturn, $'\sigma'$ is the standard deviation for the logreturn, and 'N(0,1)' is a standard normal random variable. We have estimated the expected value of the logreturn (7.71%) and the standard deviation (22.37%) from the time series data of the DAX index in the period of 12th January 1991 and 11th January 2021.

We simulate also stresses into the simulated logreturn time series. The occurrence of the stress is modeled with a Poisson process with a lambda parameter of 0.005, while the extent of the shock is modeled with a lognormal distribution with a mean of -10 and a standard deviation of 2.25. The decay of the shock is modeled with a 0.97 parameter. Finally, the stock price is determined by Equation 2, where 't' stands for time, and 'S' stands for the asset's price.

$$S_t = S_0 \cdot e^{Y_t} \qquad (2)$$

We simulate the prices for 500 days, from which the first 250 is used to define the initial margins input parameters, and the remaining 250 days will be used to calculate the IM, VM, MM on a daily basis.

We define the initial margin with the model of Béli-Váradi (2017), which model is based on the Value-at-Risk model and applies a 25% procyclicality buffer, which is exhausted if the exponentially weighted moving average (EWMA) standard deviation of the stock's logreturn is greater than its equally weighted standard deviation. This IM value will be the same for both (futures and CCP) margining methods.

We have used the following parameters for the IM calculation: the look-back period is 250 days; the significance level is 99%; the liquidation period is 2 days; the lambda parameter of the EWMA standard deviation is 1%.

The maintenance margin will be 75% of the actual day's initial margin value.

Our assumptions regarding handling the gains and losses of the marked-to-market valuation will be different, based on how it works in practice.

- CCP margining: the variation margin requirement will be paid to the CCP if the trader had a loss on that day, while he will receive money – collateral – back if he had a gain. The only exception is if his overall margin account balance (so the sum of the IM and VM) would go below the IM's value. In that case, the trader can take away only that amount from the gain to have at least the IM value on the margin account.
- Futures margining: if the trader makes a loss, it will be subtracted from the margin account balance till the maintenance margin is lower than this balance. If its MM is higher, the trader will get a margin call, and has to increase the margin account balance to the level of the IM. If the trader makes a gain, it will increase the margin account balance, even if the balance is already greater than the initial margin. This means that he won't take away the gain from the balance.

A 10-day sample of a simulated margin calculation series can be seen in Table 1 and Table 2, which shows the IM, MM, VM, and margin call dynamics.

Table 1: Simulated margin account balances in case of the futures margining

Futures margin							
date	S	initial margin	Daily gain	Margin account balance	Maintenance margin	margin call	Daily CF of the clearing member
251	1134	50,00	19,18	50,00	37,50	0,00	50,00
252	1130	50,00	-3,44	46,56	37,50	0,00	0,00
253	1111	50,00	-18,85	27,71	37,50	22,29	0,00
254	1110	50,00	-1,03	48,97	37,50	0,00	22,29
255	1121	50,00	10,60	59,57	37,50	0,00	0,00
256	1105	50,00	-16,27	43,30	37,50	0,00	0,00
257	1088	50,00	-16,41	26,89	37,50	23,11	0,00
258	1096	50,00	7,82	57,82	37,50	0,00	23,11
259	1117	50,00	21,41	79,23	37,50	0,00	0,00
260	1104	50,00	-13,37	65,86	37,50	0,00	0,00

Table 2: Simulated margin account balances in case of the CCP margining

CCP margin					
date	S	initial margin	variation margin	Margin account balance	Daily CF of the clearing member
251	1134	50,00	19,18		
252	1130	50,00	-3,44	50,00	50,00
253	1111	50,00	-18,85	53,44	3,44
254	1110	50,00	-1,03	72,29	18,85
255	1121	50,00	10,60	73,32	1,03
256	1105	50,00	-16,27	62,72	-10,60
257	1088	50,00	-16,41	78,99	16,27
258	1096	50,00	7,82	95,40	16,41
259	1117	50,00	21,41	87,59	-7,82
260	1104	50,00	-13,37	66,18	-21,41

RESULTS

We have run the simulations 11 000 times. One of the realizations can be seen in Figure 2 and Figure 3. In Figure 2, the cash flows can be seen in both of the margining methodologies.

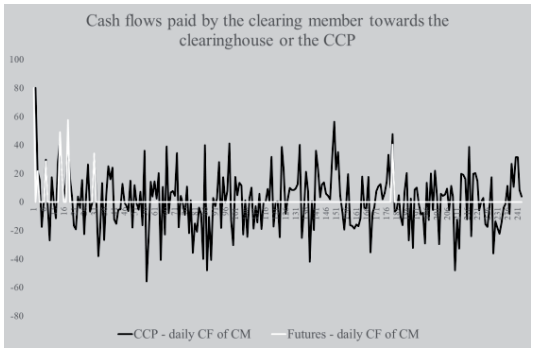


Figure 2: Cash-flows towards the clearing member from the CCP or clearinghouse point of view

It is important to see that there was less cash-flow in the futures margining since the losses did not have a cash-flow effect unless the trader has received a margin call (white line). In contrast, in the CCP margining case, there was a cash-flow every day because of the variation margin or because the initial margin changes.

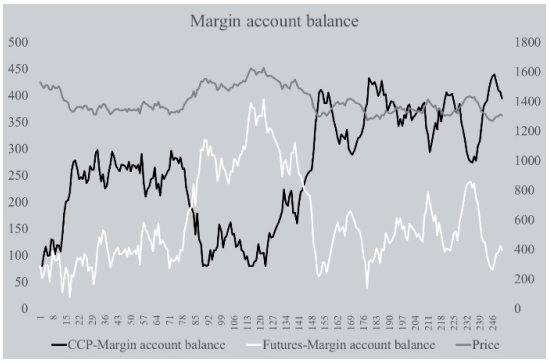


Figure 3: Margin account balances

In Figure 3, the margin account balances can be seen in both cases. The most interesting about the simulation result is that the margin account balance moves exactly the opposite direction in the futures- and CCP margining case. The two methods are capturing the same risk during the calculation of the initial margin with the same method – same VaR model, with the same parameters – but how the marked-to-market valuation is being handled is much different, which causes the balances to change differently. Comparing the margin account balance to the stock's price evolution, we can see that the futures margin account balance decreases/increases when the prices do, while CCP margining behaves the opposite way.

It is important from a procyclicality point of view since when prices are falling, it usually happens when a shock hits the market. If the margin requirements increase when there is a shock, and prices are falling, it is not as efficient from the traders' point of view and can easily cause liquidity problems to finance the increasing collateral requirements. Interestingly, in the case of the CCP margin, the notable point is to handle procyclicality throughout defining the IM in the EMIR framework; meanwhile, futures margining does not focus on this phenomenon. At the same time, futures margin moves together with the cycle and asks for less collateral when prices are falling and require more when the prices increase, so handle procyclicality much better.

To confirm this relationship, we have calculated from the 11 000 simulations the average correlation between the logchanges of the prices, CCP margins, and futures margins. There is a strong correlation between the prices and futures margin with a value of 0,784, and a very low correlation between the prices and the CCP margin, -0,004, and also very low between the futures- and CCP margin, with a value of 0,064.

Besides procyclicality, there are other essential characteristics of the margin balance, which are more critical than procyclicality, namely how good and prudent the model is. This can be quantified by the backtest, which compares a certain day's price change to the margin account balance: whether the margin was enough to cover the price change.

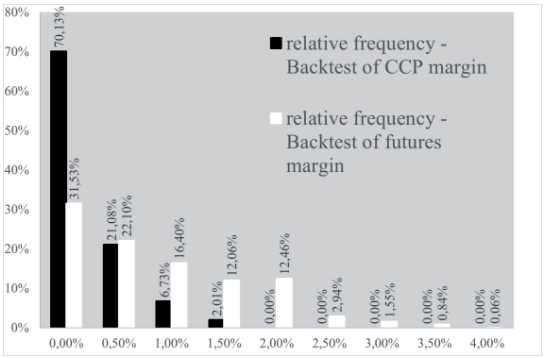


Figure 4: Backtest results

Figure 4 and Figure 5 show the result of the backtest. Figure 4 shows the relative frequency of the backtest result, which does not contain the results of the 1% of the simulated data, which we handled as outliers.

Results show that in case of the CCP margin in more than 70% of the cases the margin was enough to cover losses every day throughout the 250 days, for which the margin was simulated.
Moreover, none of the simulated time series resulted in a worse backtest outcome than 1,5%. In contrast, the futures margins' backtest did not perform as well. Only 31,53% of the cases were the margin adequate to cover every day's price change, and for comparison, only 82,14% of the cases were the backtest's result 1,5% or better. This difference is significant and notable from a risk management point of view.

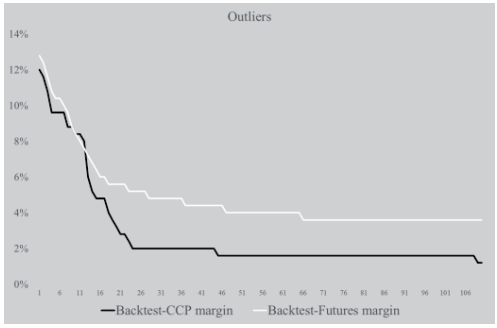


Figure 5: Backtest result in case of the outlier data

Regarding the outlier data of Figure 5, we can conclude the same result as the previous. In this figure, we see the exact values of the worst 110 backtest result in the case of both margining methods. We can see that there were some extreme values in both of the cases, but the CCP margin's most of the "outlier" data was around 2%, while the futures margin's worst values were around 4-5%. This difference is also notable.

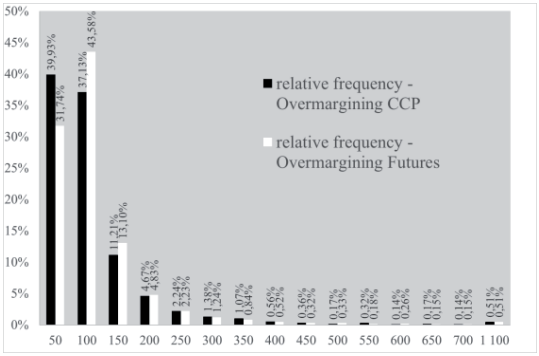


Figure 6: Overmargining

Although CCP margin performed much better than the futures margin from a backtesting point of view, it is also important that which margining method is stricter from the overmargining point of view. If the margin is always unreasonably too high, it is easy to have a good backtest result, which is good from a risk management point of view, but not necessarily good from the clearing members' since it takes away too much liquidity from them. Also, the CCP may have a competitive disadvantage if it is too expensive, requires too much collateral. We define overmargining as the ratio of the margin account a certain day and the actual price change, so how many times did the margin cover the possible losses. We calculated this ratio for all of the 250 days, and we took the average of these values in every simulation. Figure 6 contains the relative frequency of these average overmargining values. Here also we analyzed the worst 1% separately. On the x-axis, we can see how many times the margin value exceeded the price changes, while on the y-axis, we see the relative frequency of this possible outcomes. There is no notable difference above the level of the 200-times overmargining. On lower levels – till 50-times overmargining – it can be seen that the CCP margin was more frequent, while between 50- and 150-times overmargining, the futures margin was.
Figure 7 shows the outlier data results, although it contains only the ten most extreme values. In the remaining 100 outlier cases, the same can be observed as in Figure 6 in the interval of the 200-1000x.

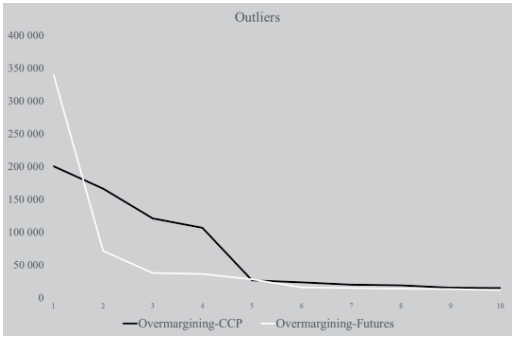


Figure 7: Overmargining in case of the outlier data

The most extreme value was in the case of the futures margining method, although the 2-4th largest values were greater in the CCP margining case.

CONCLUSION

Our results show that calculating the margin balance with the futures margining or the CCP margining can lead to a much different margin account balance, although the initial margin requirements were calculated the same way. We run our simulation 11 000 times. The main result was that the futures margin followed the stock price movements much better, which is more efficient from a trader's liquidity position point of view and from a procyclicality aspect, but the CCP margin was better from the model adequacy perspective. Namely, it performed much better on the backtest, while also it was less strict from an overmargining point of view.
Potential future research can be to change the assumptions regarding the handling of losses/gains throughout the marked-to-market process. Namely, to allow the futures margining case to take away the gains if the margin account balance is above the initial margin or the other way around, prohibiting the CCP margining from taking away the gain.

REFERENCES

Amini, H., Cont, R., Minca, A. (2016). Resilience to Contagion in Financial Networks. *Mathematical Finance*, 26(2) pp. 329–365.

BCBS-IOSCO (2015). Margin requirements on non-centrally cleared derivative. Bank for International Settlements, March 2015. https://www.bis.org/bcbs/publ/d317.pdf

Béli, M, Váradi, K. (2017). Alapletét meghatározásának lehetséges módszertana. *Financial and Economic Review* 16 (2) pp. 117-145.

Berlinger E., Dömötör B., Illés F., Váradi K. (2016). A tőzsdei elszámolóházak vesztesége. *Közgazdasági Szemle,* 63(9) pp. 993–1010.

Biais, B., Heider, F., Horeova, M. (2016). Risk-Sharing or Risk-Taking? Counterparty Risk, Incentives, and Margins. *Journal of Finance*, 71(4) pp. 1669–1698.

BIS (2012). Bank for International Settlements: Payment, clearing and settlement systems in the United States – CPSS Red book. https://www.bis.org/cpmi/publ/d105_us.pdf

CFA (2017). Equity and Fixed Income – CFA Program Curriculum, Level I, Volume 5. CFA Institute, 2017.

CFI (2021). Corporate Finance Institute: National Securities Clearing Corporations. https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/national-securities-clearing-corporation-nscc/

CME Group (2021a). Margin: Know What's Needed. https://www.cmegroup.com/education/courses/introduction-to-futures/margin-know-what-is-needed.html

CME Group (2021b). SPAN Methodology. https://www.cmegroup.com/clearing/risk-management/span-overview.html

DNB (2013). De Nederlandsche Bank – All the Ins and Outs of CCPs. https://www.dnb.nl/en/binaries/711869_All_Ins_Outs_CCPs_EN_web_v3_tcm47-288116.pdf

DFA (2010). Dodd-Frank Wall Street Reform and Consumer Protection Act. 2010, Public Law 111-203. Available: https://www.cftc.gov/sites/default/files/idc/groups/public/@swaps/documents/file/hr4173_enrolledbill.pdf

Doyle, J. Lewis, S., Dillon D., Merlini, K. Hudd, D., Koster, E.M., May, B., Wright, I. (2016). Summary of key EU and US regulatory developments relating to derivatives. Hogan Lovells report, June 2016.

DTCC, (2021). Depository Trust & Clearing Corporation homepage – CCP Resiliency and Resources. https://www.dtcc.com/news/2015/february/02/ccp-resiliency-and-resources

Duffie, D., Zhu, H. (2011). Does a central clearing counterparty reduce counterparty risk? *Review of Asset Pricing Studies*, 1(1) pp. 74-95.

EACH, (2021). European Association of CCP Clearing Houses. https://www.eachccp.eu/members/

EMIR (2012). 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. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32012R0648

ESMA (2021). European Securities Market Authority – list of non-EU CCPs with EMIR recognition. https://www.esma.europa.eu/sites/default/files/library/list_of_applicants_tc-ccps.pdf

Eurex (2021). Eurex AG homepage. https://www.eurex.com/ec-en/

Gregory, J. (2014). *Central Counterparties, Mandatory Clearing and Bilateral Margin Requirements for OTC Derivatives*. John Wiley & Sons Ltd. United Kingdom

Hull, J. C. (2018). *Options, Futures, and Other Derivatives*, 10th Edition. Pearson.

ICE Clear Credit (2021). Homepage of ICE Clear Credit https://www.theice.com/clear-credit

ICE (2021). Intercontinental Exchange homepage – Manage your risk – How clearing works? https://www.theice.com/publicdocs/How_Clearing_Works.pdf

KELER CCP (2021). Homepage of KELER CCP Ltd. https://english.kelerkszf.hu/Risk%20Management/Initial%20margin%20calculator/

King, T., Nesmith, T. D., Paulson, A., Prono, T. (2020). Central Clearing and Systemic Liquidity Risk. Finance and Economics Discussion Series, 2020(009).

Lopez, C., Saeidinezhad, E. (2017). Central Counterparties Help, But Do Not Assure Financial Stability. Munich Personal RePEc Archive.

Murphy, D. (2013). *OTC Derivatives: Bilateral Trading & Central Clearing, An Introduction to Regulatory Policy, Market Impact and Systemic Risk*. Global Financial Markets series. Palgrave-Macmillan. NY

Norman, P. (2011). *The Risk Controllers – Central Counterparty Clearing in Globalised Financial Markets*. John Wiley & Sons Ltd. United Kingdom

OCC (2021). Options Clearing Corporations homepage – What is OCC? https://www.theocc.com/Company-Information/What-Is-OCC

PFMI (2012). Principles for Financial Market Infrastructures. CPSS-IOSCO, Bank for International Settlements, April 2012. https://www.bis.org/cpmi/publ/d101a.pdf

Reg T (2021): Regulation T https://www.ecfr.gov/cgi-bin/text-idx?tpl=/ecfrbrowse/Title12/12cfr220_main_02.tpl

RTS (2013). 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. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32013R0153

SEC (2021a) Capital, Margin, and Segregation Requirements for Security-Based Swap Dealers and Major Security-Based Swap Participants and Capital and Segregation Requirements for Broker-Dealers https://www.sec.gov/rules/final/2019/34-86175.pdf

SEC (2021b). Securities and Exchange Commission Homepage. https://www.sec.gov/tm/clearing-agencies

Turing, D. (2012). *Clearing and Settlement in Europe.* Bloomsbury Professional, Haywards Heath

AUTHOR BIOGRAPHIES

MELINDA FRIESZ is a risk analyst at KELER Ltd. Her primary responsibilities are operational risk management. She is also a Ph.D. student at the Corvinus University of Budapest. Her main research areas are stress tests and market infrastructures.

KATA VÁRADI is an Associate Professor at the Corvinus University of Budapest, at the Department of Finance. She also graduated from the CUB in 2009 and after which she obtained a Ph.D. in 2012. Her main research areas are market liquidity, central counterparties, capital structure, and risk management.

MACROECONOMETRIC INPUT-OUTPUT MODEL FOR TRANSPORT SECTOR ANALYSIS

Velga Ozolina and Astra Auzina-Emsina
Faculty of Engineering Economics and Managment
Riga Technical University
Kalnciema 6, Riga, LV-, Latvia
E-mail:velga.ozolina@rtu.lv

KEYWORDS

Econometric model, input-output model, transport sector, transport policy development

ABSTRACT

Effective government transport policy can be based only on realistic data, sophisticated and detailed transport sector analysis, and productive modelling. The aim of the paper is to demonstrate the main elements used to develop a relatively small macro-economic input-output model with the emphasis on transport for one European Union (EU) country. Transport sector faces similar problems in various countries linked with emissions, transport flows, road accidents and other issues hence appropriate modelling tool should be selected. The model presented in this article consists of econometric and input-output relations. The research analyses and examines three scenarios and stresses the importance of the transport investment not only for development of the transport sector, but also for the economic development in general. The scenarios imply zero, 9 million and 6.7 million additional investment in transport sector eligible to the EU funding. As the result of additional investment, GDP recovers faster leading to 0.3-1.7%points faster growth rates as compared to the base scenario with no additional investment leading to faster cohesion with the average EU level, as well as higher number and turnover of passengers in the public and commercial transport, while the number of passenger cars is lower. The model can also be applied to study regional development, if it is possible to distinguish, which regions will benefit from the investment, as well as influence on fuel consumption and CO2 emissions, if the investments are targeted to specific means of transport.

INTRODUCTION

Policy making in the transport sector is a sophisticated task managed by the Ministry of Transport or other institution with the respective functions. It relies on information provided by the interested parties like road administrations, railway administrators, local municipalities and many others. Also for the EU countries, different EU policies and available funding has to be considered (Jankova, Jurgelane, and Auzina 2016). Moreover, nowadays emphasis has to be made also on the environmental issues, for example, CO_2 emissions (Joint Transport Research Centre 2008). Therefore, the use of different models is not only useful, but it is necessary. Such models are developed not only on a single country bases, but also as global models (Van der Zwaan, Keppo, and Johnsson 2013).

Input-output analysis is widely applied in many countries regrading transport and transportation, for example, to evaluate French maritime transport impact on air pollution (Bagoulla and Guillotreau 2020), investigate the structural emission reduction of transportation in China (Yu et al. 2021), examine economic effect of a port in Italy (Danielis and Gregori 2013).

Depending on the scope and use of the models, general framework of these models also differs. For general analysis of the impact of economic development on transport system, input-output models (Auzina-Emsina, Ozolina, and Jurgelane-Kaldava 2020; Bagoulla and Guillotreau 2020; Danielis and Gregori 2013; Yu et al. 2021) and econometric models (Auzina-Emsina, Ozolina, and Pocs 2018) can be used. In case specific aspects of the transport system are analysed using data on several countries, panel data approach can be used (Lin and Omoju 2017). Computable General Equilibrium models can be used for broader research, including global and regional aspects (Charalampidis, Karkatsoulis, and Capros 2019).

The use of models can sometimes be hindered by the lack of knowledge, as the clerks involved in the policy-making process may not have a strong background in modelling. Also time spent for development of models should be considered, as complex models tend to be very time-consuming (Skribans and Balodis 2016). Therefore, in the policy development process not only highly detailed and sophisticated models, but also relatively simple ones should be used.

The aim of the paper is to demonstrate the main elements used to develop a relatively small macro-economic input-output model with the emphasis on transport for one EU country – Latvia. Such a model can be applied in other countries, there are no limitations or specific characteristics in the modelled country (Latvia) that limit application.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

METHODOLOGY AND DATA

In order to estimate the model, publicly available data, mostly data of the national statistical office - Central Statistical Bureau (CSB) of the Republic of Latvia - were used. For indicators related to the EU, Eurostat data were used. The data period used for the model was mostly 1995-2019, however, for some indicators shorter time-period was used due to the unavailability of older data. The use of the selected databases ensures data comparability and reliability.

It should be noted that when re-estimating the model, careful analysis has to be made regarding the COVID crisis period, which began in 2020. At first, this period should be avoided, because it involves a lot of restrictions and not the conscious choice of passengers to change their mobility preferences. However, later this period can be treated similarly as the previous crises using the dummy variables in econometric equations, if necessary.

The model approach implies the use of input-output linkages and econometric equations. At present, the input-output part is static implying that it cannot be used for long-term analysis, as technologies tend to change. However, by analysing input-output tables of several years, it is possible to establish certain trends in the values of coefficients thus enhancing the applicability of the model for longer-term analysis, which is essential in case of transport projects.

Econometric equations of the model were estimated using *EViews* software, ensuring that they do not contain autocorrelation or heteroscedasticity problems (serial correlation LM test and White test without cross terms were used as the main tests to confirm that), and that all the coefficients are statistically significant.

The model itself was developed in *MS Excel*, ensuring that it can be widely used without the restrictions of specific software. This model is built to be used by the policy elaborators and clerks that have limited modelling skills. This imposes some limitations to the structure and application of the model. These limitations can at least be partly avoided during the scenario development process.

As the model is intended to be in a sense a sub-model in the context of overall economic modelling system of the government, it relies on the economic forecasts elaborated by the Ministry of Economics and the Ministry of Finance for budget planning purposes. However, it is possible also to consider other assumptions regarding the gross domestic product (GDP) growth and other important general economic indicators thus ensuring more flexible use of the model.

The current version of the model consists of the following blocks:

1. Population block, which includes the main indicators characterizing population in general and economically active population in particular, as well as the main age groups (before working age, working age and after working age).
2. GDP formation block, which includes estimates of GDP use elements and respective price indexes.
3. Input-output calculations with the Leontief function enable obtaining the values of real output and value added by industries. Currently the model disaggregrates the national economy into 8 industries – the model calculates indicators for agriculture (A according to NACE classification 2nd revision), industry (BDE), manufacturing (C), construction (F), trade and accommodation (GI), transport (H), government services (O, P, Q) and other services.
4. Employment and productivity calculations by industries.
5. Transport indicators block, which includes calculations of different transport flows, road quality and road accidents.
6. Fuel consumption and emissions, which includes calculations of consumption of the main fuels used in transport, as well as the main groups of emissions.
7. Key performance indicators (KPI) block includes calculation of several policy indicators to keep track on the broad aims of the government and their level of fulfilment. The main KPI are nominal labour productivity as % of the EU-27 respective indicator, GDP per capita as % of the EU-27 average and the decrease of the fatal road accidents.

All above-mentioned blocks are linked and ensure that development processes described within the model are coherent. Additional blocks can be added if certain modelling enquiry demands it.

Transport indicators block

The most important group of equations in the transport indicators block is related to the passenger flows therefore it will be discussed in more detail in this article.

There are many factors influencing these flows like GDP, average wages, level of automobilization, domestic income and expenditures, demographical indicators, changes on urban territories, economic changes, crisis and others (Griskeviciene, Griskevicius, and Griskeviciute-Geciene 2011). As it is not possible to even check all of those factors given that only 25 observations are available for the estimation of the equations, the main factors have to be distinguished. Moreover, distinction has to be made among the factors that influence the necessity to travel as such (the number of passengers) and the amount of travel done (passenger kilometres).

Table 1 summarises the results of specified equations for the number of passengers by the main means of transport in Latvia in the form of the elasticity coefficients as the respective equations are estimated in

log-log form. It is evident that the number of passengers in transport means available only in the big cities (trolley busses and trams) is more affected by the number of population than more long-distance solutions as trains and busses. It can also be noted that the number of passengers in busses reacts slower to GDP changes, because busses are essential in intra-city, intra-regional and multi-regional traffic, while other means of the transport cover only specific areas. It can also be noted that in case of continuous decrease in population, it is expected that the investments will be more targeted at the quality improvements in the transport system rather than the quantitative ones as larger number of vehicles per se.

Table 1: Elasticity Coefficients in Equations for the Number of Passengers by the Means of Transport

Means of Transport	Factor	
	Population	GDP
Train	3.6	1.0
Bus	2.9	0.7
Trolley bus	6.5	0.9
Tram	7.1	1.0

Other factors are included in the scenario analysis indirectly via the coefficients of the distance per passenger (DISTANCE) in railway and road transport, which relate the number of passengers (PASSENGERS) and passenger kilometres (PKM) as Equation (1) shows. These coefficients can also be used to show the modal shift between the rail and road transport.

$$PKM = PASSENGERS* DISTANCE \quad (1)$$

Alternative to the public transport is the use of the private vehicles. Therefore the model includes also the calculation of the number of private cars (CARS) as shown in Equation (2).

$$LN(CARS) = 0.8*LN(GDP) - 1.55 \quad (2)$$

The coefficient of elasticity of private transport to GDP is similar as that of busses, only slightly higher. It may indicate that people in rural areas tend to use more private cars as the economy develops in contrast of using public transport more often. Further the structure of vehicles by the motor type is used to model the fuel consumption and analyse the impact of changes in the structure on fuel consumption.

SCENARIOS

Three scenarios were developed to show the influence of the transport projects on the economic development of Latvia. The base scenario implies no extra investments in the transport sector. Other two scenarios include the implementation of particular transport projects attracting additional investments. These projects are selected by the Ministry of Transport and are eligible to the EU financing.

The first scenario includes a set of projects with investments of 9 billion EUR, which are evenly distributed in 2021-2027 as there was no information available on when it is intended to implement each of those projects. The second scenario includes a set of projects with investments of 6.7 million EUR. In each of the sets of projects 1.5 million EUR are the private investments. Both scenarios imply that additional government consumption expenditures are not necessary.

It is assumed that the increase in the quality of transport infrastructure allows to increase the productivity in transport sector as well. In the first scenario it is assumed that the productivity increase will be 1%point higher than in the base scenario, and in the second scenario it is assumed to be 0.8%points higher.

The first scenario implies that there will be a modal shift from road transport to railway transport, especially in international traffic. Also the use of the public transport will increase. This shift is incorporated in the coefficients of distance per passenger.

Both scenarios ensure that passenger kilometres in road transport in 2030 will increase by 8.2% compared to 2019, but the number of passenger cars will decrease by 5.6% in the same period. In the meanwhile the passenger turnover in railways will be by 3.6% higher in the first scenario and by 2.3% higher in the second scenario.

As one of the projects included in the scenarios implies the adjustment of 200 busses for the use of alternative fuel and several more projects are dealing with the infrastructure for electric vehicles, the structure of the cars by the fuel type is changed and the share of busses run on alternative fuels is increased by 7%points, but the share of private cars run on electricity is increased by 0.5%.

RESUTS AND DISCUSSION

When analysing the results, we have to consider the COVID crisis, which is not yet thoroughly investigated and it is not clear, how exactly it will influence all the relationships included in the model. Therefore, comparison of the scenarios is more important here rather than the forecasts as such.

According to the assumptions regarding the general trends of economic development in Latvia, real GDP in the base scenario is estimated to decrease by 7.3% in 2020 followed by slight slowdown by 0.1% in 2021 and afterwards it would increase by 2.0-2.4% annually (see Figure 1). Due to additional investments in transport sector, the recovery of Latvia after the COVID-19 crisis would be much faster at around 2.3-4.1%.

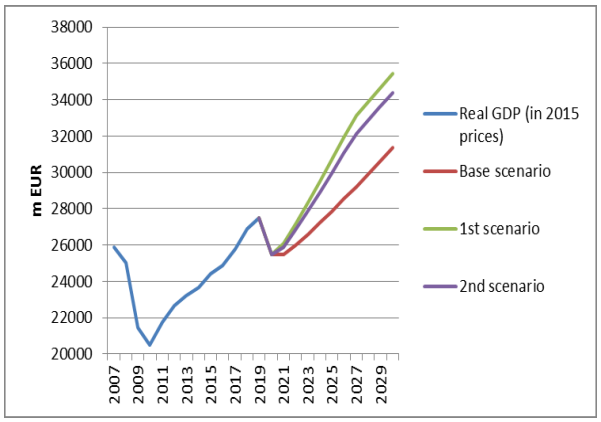


Figure 1: Real GDP of Latvia, m EUR

As it was already mentioned, transport projects included in the 1st and 2nd scenario would help to increase the productivity in transport industry. Figure 2 shows that it would help to sustain the level of productivity in transport sector against the EU-27 level, while in the base scenario the productivity in Latvia would become comparatively lower.

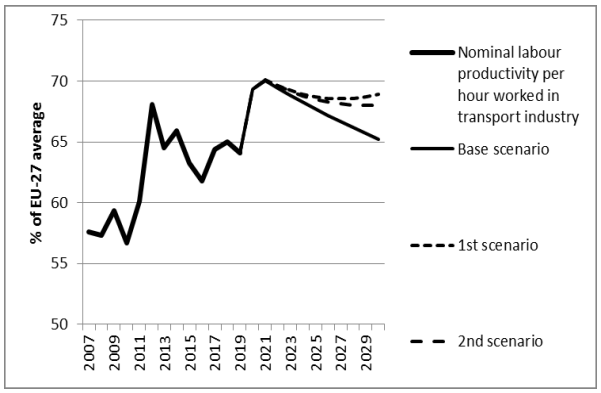


Figure 2: Nominal Labour Productivity per Hour Worked in Transport Industry in Latvia, % of the EU-27 average

In line with the assumptions made regarding the economic development, price levels and demographic situation, GDP per capita (PPP) relative to the EU-27 average would continue to increase. The increase would be even faster, if any of the additional investment packages would be implemented as it is evident from Figure 3. By year 2029 GDP per capita would reach only 75% of the average EU level, while in case of the 1st and 2nd scenario these values would be 85% and 83% respectively.

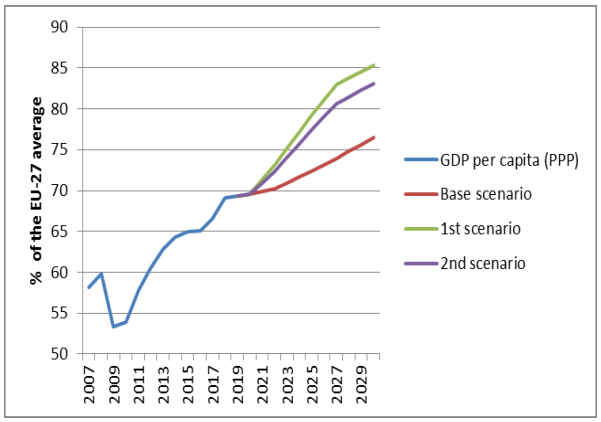


Figure 3: GDP per Capita (PPP) in Latvia, % of the EU-27 average

In case of more rapid economic development, also the number of passengers would increase, as Figure 4 shows. However, after the rapid decrease in 2020, the number of passenger would not recover, in comparison with 2019 during the next 10 years.

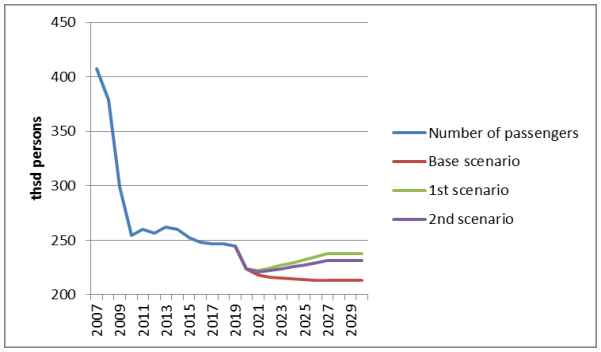


Figure 4: Number of Passengers in Public Transport in Latvia, thsd persons

Although other forecasts seem quite plausible, the forecasted passenger turnover seems a bit too optimistic (see Figure 5). Thus this aspect has to be analysed further more thoroughly also in the context of changes in people behaviour after the COVID-19 restrictions will end.

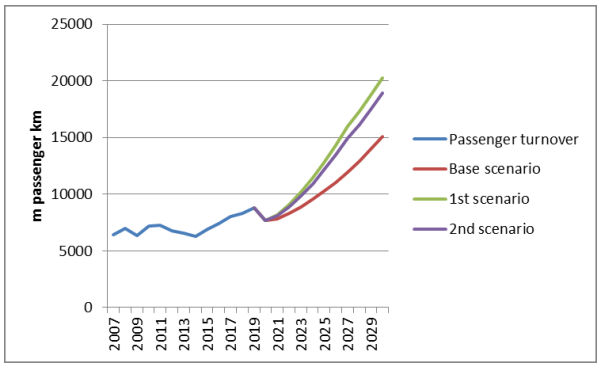


Figure 5: Passenger Turnover in Latvia, m passenger km

Future prospects of the number of registered passenger cars are similar in all the scenarios (see Figure 6). Only at the end of the forecast period the number of cars decreases in the 1st and the 2nd scenario as compared to the base scenario due to the implemented projects.

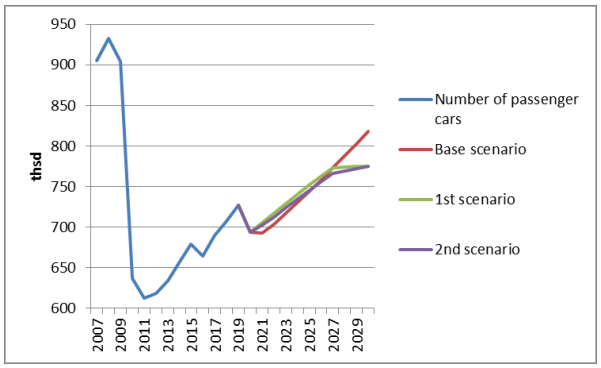


Figure 6: Number of Registered Passenger Cars, thsd.

The elaborated model has proved to be relatively easy used for policy analysis by providing additional information on the impact of transport projects on the economic development.

In case it is possible to distinguish, which regions of the country will benefit from the investment, it is possible to evaluate differences of the regional GDP and the number of population as well as other indicators available in regional detail.

Other application of the model is related to the fuel consumption. If some investment projects imply replacement of fosil fues vehicles by alternative energy vehicles, it is possible to evaluate the effects on the vehicle structure by fuel types and CO2 emissions.

CONCLUSIONS

The main findings show that indeed transport processes are closely related to the economic and demographic development and these relations can be modelled in a relatively simple manner. However, it also implies thorough analysis during the scenario elaboration process as some of the variables included in the model can be highly sensitive to the future estimates of their values.

Future research in this field should be done both by updating and improving the equations already included in the model, but also by introducing the regional aspect to the model as transport projects are in many cases associated to the particular regions of the country and not the country in general.

In order to use the model for long-term analysis, it is necessary to use several input-output tables and analyse trends in the input-output coefficients in order to model changes in the technologies.

ACKNOWLEDGEMENTS

Participation in the conference was funded by European Regional Development Fund (ERDF), Measure 1.1.1.5 "Support to international cooperation projects in research and innovation of RTU". Project No. 1.1.1.5/18/I/008.

REFERENCES

Auzina-Emsina, A.; V. Ozolina; and I. Jurgelane-Kaldava. 2020. "Modeling Global Consumptions Trends Impact on Transport and Logistics: Scenario Analysis." In *ICTE in Transportation and Logistics 2019. ICTE ToL 2019. Lecture Notes in Intelligent Transportation and Infrastructure*, 1–8.

Auzina-Emsina, A.; V. Ozolina; and R. Pocs. 2018. "Competitiveness and Economic Development Scenarios of Latvia." *Business, Management and Education* 16(0), 40–53.

Bagoulla, C. and P. Guillotreau. 2020. "Maritime Transport in the French Economy and Its Impact on Air Pollution: An Input-Output Analysis." *Marine Policy*, 116.

Charalampidis, I.; P. Karkatsoulis; and P. Capros. 2019. "A Regional Economy-Energy-Transport Model of the EU for Assessing Decarbonization in Transport." *Energies* 12(16), 3128.

Danielis, R. and T. Gregori. 2013. "An Input-Output-Based Methodology to Estimate the Economic Role of a Port: The Case of the Port System of the Friuli Venezia Giulia Region, Italy." *Maritime Economics and Logistics* 15(2), 222–55.

Griskeviciene, D.; A. Griskevicius; and A. Griskeviciute-Geciene. 2011. "Peculiarities of Passenger Flows Forecast on the Reduced Market Conditions." In *8th International Conference on Environmental Engineering, ICEE 2011*, 898–904.

Yu, Y.; Sh. Li, H. Sun; and F. Taghizadeh-Hesary. 2021. "Energy Carbon Emission Reduction of China's Transportation Sector: An Input–output Approach." *Economic Analysis and Policy* 69, 378–93.

Jankova, L.; I. Jurgelane; and A. Auzina. 2016. "European Union Cohesion Policy." In *Economic Science for Rural Development: Integrated and Sustainable Regional Development, Production and Co-Operation in Agriculture*, 79–85. http://llufb.llu.lv/conference/economic_science_rural/2016/Latvia_ESRD_42_2016-79-85.pdf\nwos:000389983700010.

Joint Transport Research Centre. 2008. *Discussion paper Transport Outlook 2008, Focusing on CO2 Emissions from Road Vehicles.*

Lin, B. and O. E. Omoju. 2017. "Does Private Investment in the Transport Sector Mitigate the Environmental Impact of Urbanisation? Evidence from Asia." *Journal of Cleaner Production* 153, 331–41.

Skribans, V.; and M. Balodis. 2016. "Development of the Latvian Energy Sector Competitiveness System Dynamic Model." In *Proceedings of the 9th International Scientific Conference "Business and Management 2016"*, May 12–13, 2016, Vilnius, Lithuania, Vilnius Gediminas Technical University. doi:10.3846/bm.2016.12.

Van der Zwaan, B.; I. Keppo; and F. Johnsson. 2013. "How to Decarbonize the Transport Sector?" *Energy Policy* 61, 562–73.

AUTHOR BIOGRAPHIES

VELGA OZOLINA holds a PhD Degree in economics (Dr.oec.), Riga Technical University, Latvia, 2009, where she works since 2004. She is an Assistant Professor since 2017 and Leading Researcher since 2016 at the Faculty of Engineering Economics and Management of Riga Technical University. Research interests include different issues of economic analysis, macroeconomic modelling and external trade of Latvia.
She is a Member of the Association of Latvian Econometrists and INFORUM group (Interindustry Forecasting Project at the University of Maryland). The author's Doctoral Thesis is awarded as the Research of Young Scientist of 2010 in the group of Doctoral Theses by the Association of Latvian Econometrists.
Her e-mail address is velga.ozolina@rtu.lv

ASTRA AUZINA-EMSINA holds the PhD degree in economics (Dr.oec.), Riga Technical University, 2008. Since 2004 she has been a Researcher and since 2009 an Assistant Professor at the Faculty of Engineering Economics and Management of Riga Technical University. She has been involved in research devoted to economic modelling and sectoral interlinkages since 2004 and has developed several multi-sectoral macroeconomic models. She is a member and co-founder of the Association of Latvian Young Scientists, member of the International Input-Output Association and the Association of Latvian Econometrists (also board member since 2012), member of INFORUM group.
Her e-mail: astra.auzina-emsina@rtu.lv

ESTABLISHING A BASIS FOR DECISION SUPPORT MODELLING OF FUTURE ZERO EMISSIONS SEA BASED TOURISM MOBILITY IN THE GEIRANGER FJORD AREA

Børge Heggen Johansen
Department of Ocean Operations and Civil Engineering
Norwegian University of Science and Technology
Larsgårdsvegen 2, 6009 Ålesund
E-mail: borge.a.h.johansen@ntnu.no

KEYWORDS

Cruise tourism, shipping emissions, tourism mobility

ABSTRACT

Destinations for cruise tourism have to manage both opportunities and challenges of hosting cruise ships. Governing bodies in Norway are proposing new environmental regulations to abate environmental impacts, but some stakeholders worry that stringent regulations will cause less value generation for local business. The purpose of this paper is to establish a basis for decision support modelling on future zero emission sea-based tourism mobility for the Geiranger fjord area. The tourism mobility system is mapped through a systems engineering lens. The analysis systematizes the tourism mobility system, prior studies on air pollution and emissions, existing- and upcoming regulations. The study concludes by proposing an objectives hierarchy and measure of effectiveness for use in future works.

INTRODUCTION

The Geiranger Fjord is one of the hallmarks of Norwegian tourist attractions known for its pristine nature. Recently the area has been under much debate due to rising air pollution from traffic and seaborne activity. The Norwegian Maritime Authority has been assigned by the Norwegian Environmental Protection Agency to consider new environmental regulations regarding emissions to air for large ships operating on the Norwegian World Heritage Fjords. This is anticipated to reduce emissions from ships, but business actors in Geiranger fear that this will reduce the number of port calls in Geiranger and divert tourists to buses (increased road transport) from other cruise ports nearby (nrk.no). This is one example of many conflicting stakeholder interest needed to be balanced when making decisions on which new environmental policies that should be implemented. Therefore, the current study aims to formalize the system of tourism mobility and establish an objectives hierarchy for the case of reducing emissions to air from tourism-based mobility in the Geiranger fjord based on systems engineering methodology to further be used in future studies on decision support modelling.

GEIRANGER

Geiranger is a remotely located and scarcely populated village on the west coast of Norway with only 230 permanent residents. The village is situated in a landscape recognized from its steep fjords and rural landscape, listed as a UNESCO World Heritage site.

Figure 1: Geiranger, Photo © Stranda Port Authority

Geiranger is one of Norways' most visited cruise ports, receiving 346.327 passengers in 2018 (Yttredal et al., 2019). The majority of the remaining tourists arrive by private cars or chartered busses. This mobility brings many challenges in Geiranger where the infrastructure is beginning to reach its capacity due to the increasing influx of tourists (Tallaksen and Holm, 2007). This is in line with the findings of Dickinson and Robbins (2008) examining the traffic problems related to tourism at rural destinations.

A study on mobility in Geiranger was made by Shlopak et al. (2014). The second part of this study, Svendsen et al. (2014), assessed the emissions from sea- and land-based transport in Geiranger. The topic was further studied by Weggeberg et al. (2017) conserning emissions from ships in Geiranger and Nærøyfjord on

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

behalf of the Norwegian Maritime Authority (NMA). The Norwegian Ministry of Climate and Environment has initiated NMA to consider options for imposing regulations reducing the environmental impacts from shipping in Nærøyfjord, Aurlandsfjord, Geirangerfjord, Synnulvsfjord and the inner part of Tafjord (regjeringen.no). The current study seeks to use a systems engineering lens to systematize previous studies together with existing and upcoming regulations to establish a basis for decision support modelling on future zero emission sea-based tourism mobility for the Geiranger fjord area.

MOBILITY IN THE GEIRANGER AREA

The tourism mobility system in the Geiranger area entails the characteristics of a system as defined by NASA (2007) being "*...a construct or collection of different elements that together produce results not obtainable by the elements alone.*" The two elements at the top-level in the taxonomy describing the transport system in Geiranger are road transport and seaborne transport. The next level in the taxonomy divides between the main categories of transport within the elements road and seaborne traffic. These are light- and heavy vehicles for road transport and ferries, cruise ships, guding boats, tender boats and RIB boats for seaborne transport.

Geiranger cruise port comprises a cruise terminal, seawalk and up to four anchor positions depending on the size of the ships. There are on average 1-2 cruise ships anchored up in Geiranger in the tourist season. The number of curise passengers in Geiranger is limited to 6000 PAX. Most of the passengers landing from the Cruise ships enter by the cruise ships' own tender boats. There is an own terminal for the tender boats at Geiranger port. The tender boats have a capacity around 150 PAX and are driven by small marine diesel engines, typically around 200 kW. Cruise ships normally use 2-4 tender boats on each port call (Svendsen et al., 2014). It is possible to travel to Geiranger by ferry from the village of Hellesylt with daily departures in the period May through October. In the peak season between June and August, the ferry takes eight roundtrips. The ferry service is run by the vessels "Bolsøy" and "Veøy", built 1971/74, with capacity of 36 cars and 345 passengers (Svendsen et al., 2014). Geiranger Fjordservice operate sightseeing boats on the Geiranger Fjord. In the peak season between June and August there are 11 daily roundtrips. Local tour operators also arrange daily excursions on high-speed RIB boats on the Geiranger Fjord. The vessels carry around 15 passengers per trip and are powered by marine diesel engines around 250 kW. For the 2014 season, about 700 trips were arranged (Svendsen et al., 2014).

Although much of the debate on environmental impacts from tourism in Geiranger focuses on emissions from cruise ships, the bulk of tourists coming to Geiranger are arriving with cars and busses. There are three points of entry in to Geiranger (see figure 2 for reference). From the north through Ørnesvingen from Eidsdal, from the south through Flydalsjuvet from either Grotli or Stryn. The third is by ferry directly to Geiranger. Both the northern and southern route in and out of Geiranger are serpentine mountain roads, characteristic for the area and an experience by themselves, but a challenge with regards to traffic. Calculations of traffic capacity shows that the road network outside the center of Geiranger, can withstand 1100 vehicles per hour, which is well above the actual daily traffic level in the peak of the season (Svendsen et al., 2014). Problems arise when multiple vehicles and especially multiple busses are trying to navigate through the narrow streets of Geiranger center or the tightest bends on the serpentine mountain roads. Queues are also expected to occur as drivers of private cars seek parking in Geiranger.

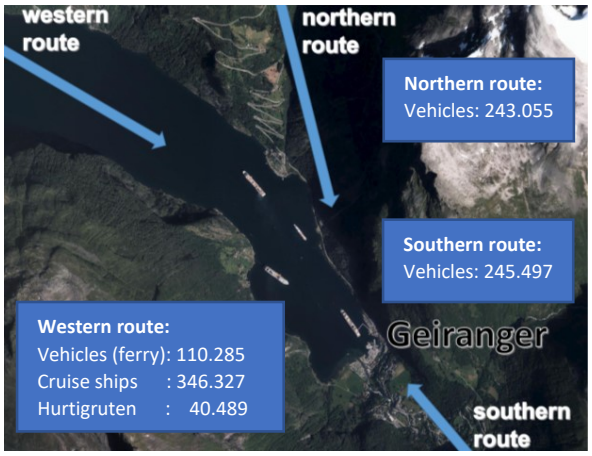


Figure 2: Entry points to Geiranger with respecive number of visitors in 2018 (Yttredal, et al., 2019) Photo © GuleSider

In 2013, the annual mean value of vehicles through Geiranger center per day was measured to be 562 vehicles per day. The corresponding value for the summer months (June, July, August) (SDT) was 1375 and for July alone (JDT) it was 1779 vehicles per day (Svendsen et al., 2014). The Norwegian Public Roads Administration divide between heavy and light vehicles saying that all vehicles above 5,6 meters are considered as heavy vehicles and correspondingly the opposite for light vehicles. In prior studies by Svendsen et al. (2014) it was estimated that heavy vehicles represent between 14% (Svendsen et al., 2014) and 8% (Diez-Gutierrez and Babri, 2020) of the road-traffic in peak season. In addition to higher traffic in general, the activity by heavy vehicles like buses that are more emissions intensive is higher in the peak of the season. One concern expressed by some stakeholders in Geiranger is that they worry that impending strict regulations for emissions to ships operating on the Geiranger Fjord would redirect vessels

to other ports nearby and that the tourists would be transported in to Geiranger with buses instead.

AIR POLLUTION IN GEIRANGER

Prior studies show that emissions to air in Geiranger mainly come from combustion of fossil fuels in ships and vehicles as well as abrasion and wear of roads due to traffic. Emissions from combustion of solid biofuels for heating of residential buildings is outside the scope of this study. It is also not relevant for the tourist season as it is in the middle of summer where heating demand for housing is at a minimum.

Focus is put on (sulfur dioxide) SO_2 , nitrous oxide NO_x and particulate matter (PM). An important constituency of the $PM_{2.5}$ emissions fraction is black carbon. Black carbon originates from incomplete combustion of hydrocarbons. Black carbon serves as a carrier for heavy metals, inorganic salts and organics such as PAHs that are known for adverse health effects. Due to this, the exhaust from diesel engines is classified as carcinogenic to humans. Exposure to NO_x may cause impaired lung function, increased susceptibility to respiratory infections and development or worsening of asthma and bronchitis (WHO, 2016). According to the study "Air Quality in Europe – 2016 report" approximately 1600 Norwegians died prematurely in 2013 due to exposure to $PM_{2.5}$ emissions. The corresponding number for NO_2 was 170 (EEA, 2016; NILU, 2016).

MEASURED VALUES OF AIR POLLUTION

Haugsbakk and Tønnesen (2010) did a measurement of PM_{10} and NO_2 concentration in Geiranger Centre by the ferry dock. Measurements were made consciously from July to September 2010. Results from the study show two occasions of PM_{10} exceeding the threshold level of 50 $\mu g/m^3$. There were no exceedances of NO_2. Löffler (2017) did measurements of SO_2, PM_{10} and $PM_{2.5}$, concentrations in the air in Geiranger from June 2015 to September 2016. SO_2 concentrations were measured with hourly maxima below 10 $\mu g/m^3$ for the entire duration of measurement. Löffler and his team found relatively high concentration of PM_1 and deemed this as the major pollutants in the Geiranger area. Six exceedances of $PM_{2.5}$ over the threshold level of 25 $\mu g/m^3$ were observed. There was also a correlation of high PM_1 concentrations and relatively high SO_2 concentration, making it plausible that the PM_1 emission originate from combustion of petroleum products, although no definitive conclusions were made. It was also observed that the PM fraction is suspended in the air over many weeks and is transported by circulating air along the entire valley, apparently being trapped by a combination of the local topography and special weather conditions.

ESTIMATED VALUES OF AIR POLLUTION

Weggeberg et al., (2017) chose a theoretic approach to estimate the concentration of NO_x, $PM_{2.5}$, PM_{10} and SO_2 in the Geiranger area. Based on vessel information from AIS system on movement and IHS Fairplay data for machinery the emissions from shipping were calculated. Together with statistical data from road traffic the emissions from sea and land transport were fed in to a CALPUFF model. The CALPUFF model is a modeling system for the simulation of atmospheric pollution dispersion. Results from the analysis show only one occasion of NO_2 emissions exceeding the one hour threshold level of 200 $\mu g/m^3$. The report suggest that one hour threshold levels should be the benchmark for Geiranger due to the fact that the activity and emissions are condensed within a few months in the summer making annual mean values less relevant. Some occurrences of emissions of $PM_{2.5}$ in the area around 20 $\mu g/m^3$ were also observed in Geiranger Centre (ibid.).

Emission of NO_2, PM_{10} and $PM_{2.5}$ from road traffic in Geiranger was calculated on the basis traffic statistics and emission factors for road vehicles. Combustion of diesel and gasoline together with abrasion of the road from the tires and brakepads contribute to the emissions. The results of the calculations of two studies made in recent years are given in table 1:

Table 1: Emission of NO_x, PM_{10} and $PM_{2.5}$ from road traffic in Geiranger

Study	Area of focus	Emissions from road traffic [tonnes]		
		NOx	PM_{10}	$PM_{2.5}$
Weggeberg et al., 2017	Estimated NOx $PM_{2.5}$, and PM_{10} emissions from road transport in Geiranger from June to August 2016.	1.75	0.066	0.052
Shlopak et al., 2014	Estimated NOx and CO_2 emissions from road transport in Geiranger for the year 2013.	2.93		

Cruise ships provide accommodation and leisure services throughout their journeys in addition to provide transport for up to 6000 guests and 2000 crewmembers. The ships have many engines on board that allow for flexible operation and electricity generation at varying power requirements. A survey to cruise operators in the Geiranger fjord in 2016 show that 36% of cruise ships operating in Geiranger have direct mechanical propulsion and 64% have diesel electric propulsion (Stenersen, 2017). Nominal engine speeds of the engines are in the area 400-600 rpm and total installed power

range from a few MW on the smallest ships to 120 MW on the biggest ship. There is a linear correlation between ship size in gross tonnage (GT), total installed power and passenger capacity (PAX). Ships maneuvering or at berth/anchor in Geiranger report a normal distribution around a mean of around 50% engine load. The mean time at berth in Geiranger is 6-8 hours (ibid).

Marine diesel engines are able to use a variety of fuel oil qualities. The quality of the fuel affect emissions ranging from high sulphur containing heavy fuel oil (HFO) to low sulphur fuels like marine gas oil (MGO). The cost of fuel is higher in low-sulphur oils than residual high sulphur oils. High sulphur content in fuel give high emissions of sulphur oxides and particulate matter. 70% of cruise ships operating in Geiranger in 2016 used MGO (<0.1% Sulphur content) for main- and auxiliary engines (Stenersen, 2017). After the exhaust gases are emitted into the air from the ship stacks they are diluted in the ambient air. During the dilution process they are partly chemically transformed or removed (Eyring et al., 2005). Nitrogen oxides (NO_x) are formed when combusting fossil fuels at high temperatures. Emissions of NO_x and PM have local effects like air pollution and smog formation, having an effect on human health. It is estimated that, globally, 30% of smog comes from ships (Miola et al., 2010).

In addition to replacing high-sulphur fuels with low-sulphur fuels there are mainly three main emissions abatement technologies available for marine diesel engines; exhaust gas scrubbers for SO_2 and PM reduction and Selective Catalyst Reduction (SCR) and Exhaust Gas Recirculation (EGR) for NO_x reduction. Although 20-25% of respondents to the survey by (Stenersen, 2017) state they have NO_x abating technology installed, only two of the ships responding declared they were compliant to IMO Tier III levels. Sulfur emission requirements can be met either by using low sulfur fuel or by cleaning exhaust for sulphur. About 25% of ships operating in Geiranger state that they have installed scrubber systems to reduce SO_x emissions (Stenersen, 2017).

Due to regulations being implemented in recent years an increasing share of cruise ships are built with emission abatement technologies where 49% of new global passenger capacity is based on LNG for primary propulsion (CLIA, 2020). The challenge for Geiranger is that the mean age of ships operating in the area not renewing at a rate which is fast enough to meet many of the coming regulations (see figure 3). Any regulations capping emissions from ships will have accompanying costs to the ship-owners if re-builds of the marine power generating systems are needed in order to comply.

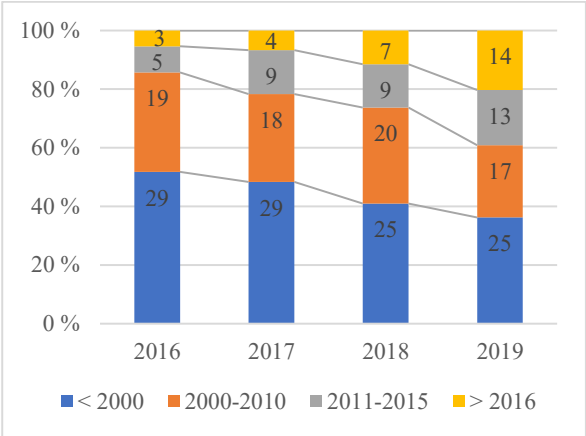


Figure 3: Share of cruise ships calling to Geiranger built according to MARPOL Appendix VI Tier levels. (Based on itinerary data from Stranda Port Authority)

In recent years two substantial assessments on the environmental impacts from seaborne transport in Geiranger has been conducted; Svendsen et al., (2015) and Weggeberg et al., (2017). Both follow a bottom-up approach where data on movement and number of ships were collected. Operation characterization for the ships were set up based on collected technical information like ship size, main engine type, auxiliary engine type, type of fuel etc. Emissions were then calculated combining activity data and technical data for the ships. The discrepancy between estimated values of NO_x between the two studies is due to different systems boundaries both in time and geography. The large relative contribution from cruise ships make it reasonable to consider subjecting further regulations to cruise ships as an effective action to reduce emissions in Geiranger (NMA, 2017).

Table 2: Estimated NO_x, $PM_{2.5}$, and PM_{10} emissions from sea transport in Geiranger

Study and area of focus	**Emissions from sea traffic [tonnes]**		
	NOx	PM_{10}	$PM_{2.5}$
Weggeberg et al., 2017; Estimated NO_x $PM_{2.5}$, and PM_{10} emissions from sea transport in Geiranger from June to August 2016.	67.9	2.15	1.97
Share of total emissions from mobility in Geiranger attributed to cruise ships	81%		
Svendsen et al., 2015; Estimated NO_x $PM_{2.5}$, and PM_{10} emissions from sea transport in Geiranger from June to August 2016.	203	2,85	
Share of total emissions from transport in Geiranger attributed to cruise ships	94%	94%	

REGULATIONS

Air emissions from ships is regulated by MARPOL Annex VI. The convention contains provisions controlling emissions of NO_x, SO_2, PM, VOC and ozone depleting substances as well as waste incineration on board, repairs and the quality of fuel. MARPOL also defines specific emission control areas (ECA). The International Maritime Organization (IMO) has put a cap on sulphur content in fuels used on board ships to 3.5% m/m. In emission control areas (ECA) the sulphur content is capped at 0.1% m/m (imo.org). Reducing sulphur considerably reduce emissions of SO_2 and PM, but also NO_x by around 5% (Cooper and Gustavsson, 2004). Air emissions are of particular interest to EU countries and therefore the EU Directive 2005/33/EC requires that all ships in European ports use fuel with low Sulphur content of 0.1%, this is much more stringent than the MARPOL standard of 3.5%. The implementation of directive 2005/33/EC has had a noticeable effect in reducing SO_2 in some European ports (Schembari et al., 2012). The NO_x-emission limits introduced by IMO apply to marine diesel engines and depend on an engine's operating speed. Requirements in Tier I apply to ships built after 2000, while the stricter Tier II limits apply to ships built after 2011. In emission control zones, Tier III requirements apply to ships built after 2016. A Tier III compliant ship will have about 74% less NOx emissions than a Tier II ship (Martinsen and Torvanger, 2013). Tier I and Tier II are global requirements, whereas Tier III standards only apply to current existing ECAs in North America and the North Sea and Baltic sea in 2021 (imo.org).

Table 3: Proposed regulatory actions for the Norwegian World Heritage Fjords (NMA, 2017)

	Proposed reguatory action
R1	Requirement for ships to have NO_x emissions not exceeding values given in MARPOL Appendix VI, 13.4 (Tier II) within 2018 and 13.5 (Tier III) by 2020
R2	Only allow use of low-sulphur fuel oil within the World-Heritage fjords
R3	Introduce requirement for ship smoke opacity. 50% at start-up and 10% for steaming
R4	Mandatory environmental reporting for all ships operating in the World-Heritage Fjords
R5	Limit number of port-calls and consider long-term operation licenses for Cruise lines with strong environmental limits
R6	Introduce speed limits for given parts of the fjords.

Maybe the most ambitious measure suggested by NMA is to demand ships operating on the World Heritage Fjords to comply to MARPOL Anenx VI, part 13.4 (Tier II) by 2018 and part 13.5 (Tier III) by 2020 (R1)(NMA, 2017). The study made by Stenersen (2017) found that 18% of ships operating in Geiranger in 2016 achieve at least Tier II level while only 6% are Tier III. Emissions of NO_x and PM could be reduced by using low Sulphur fuel oil, but mitigating NOx requires costly alterations in the engine room either by re-building engines, installing new engines or installing SCR systems (NMA, 2017). In addition to R1 entering in to force, the Norwegian parliament also set requirements for government to introduce regulations requiring zero-emission shipping in the World Heritage Fjords by 2026 (stortinget.no).

IMPLICATIONS OF NEW REGULATIONS

There is a large body of literature discussing stakeholders. An often-cited definition of stakeholder is given by Freeman (1984). Stakeholders to tourism mobility system in Geiranger and their perceptions are in the process of being mapped in the research project SUSTRANS (www.sustrans.no). Stakeholders' valuations of different criteria diverge and therefore the task of making the right decisions, suiting the needs and expectations of stakeholders becomes challenging. Measure of effectiveness (MoE) is a term describing the realization of a system (NASA, 2007). The obvious logistic aspects of the tourism mobility system in Geiranger needs to be realized, but there is an available solution space when "designing" the transport system that would give different solutions suiting different demands. Decision makers need to design the transport system, within the scope of the upcoming environmental regulations for ships operating on the World Heritage Fjords. The following measure of effectiveness (MoE) is suggested based on the "*Big picture*"-project by the local municipality (stranda.kommune.no), focusing on the future of tourism within the scope low- and zero emission seaborne tourist mobility: "*Economic, social and environmentally sustainable tourism industry in the Geiranger fjord and adjacent area beyond 2026.* " The assessment of several competing aspects in context of decision making is often referred to as Multi Criteria Decision Analysis (MCDA) (Belton and Steward, 2002). One of the simpler methods within MCDA is by using decision matrixes. In order to get the full view of the possibilities decision makers have, an extensive survey should be made exploring infrastructure epochs in combination with new regulatory regimes where quantifying the effects and perceived stakeholder valuation of initiatives for reducing emissions is one of them. The evaluation of regulatory actions with assessment of emissions abatement and perceived stakeholder valuations discussed in the earlier sections could be summarized in an objectives hierarchy. The objectives hierarchy which aims to rank the different decisions in a risk management setting (NASA, 2007). A proposed simplified objectives hierarchy structure for the case of reducing emissions to air and improving air quality in the Geiranger Fjord is given in figure 4.

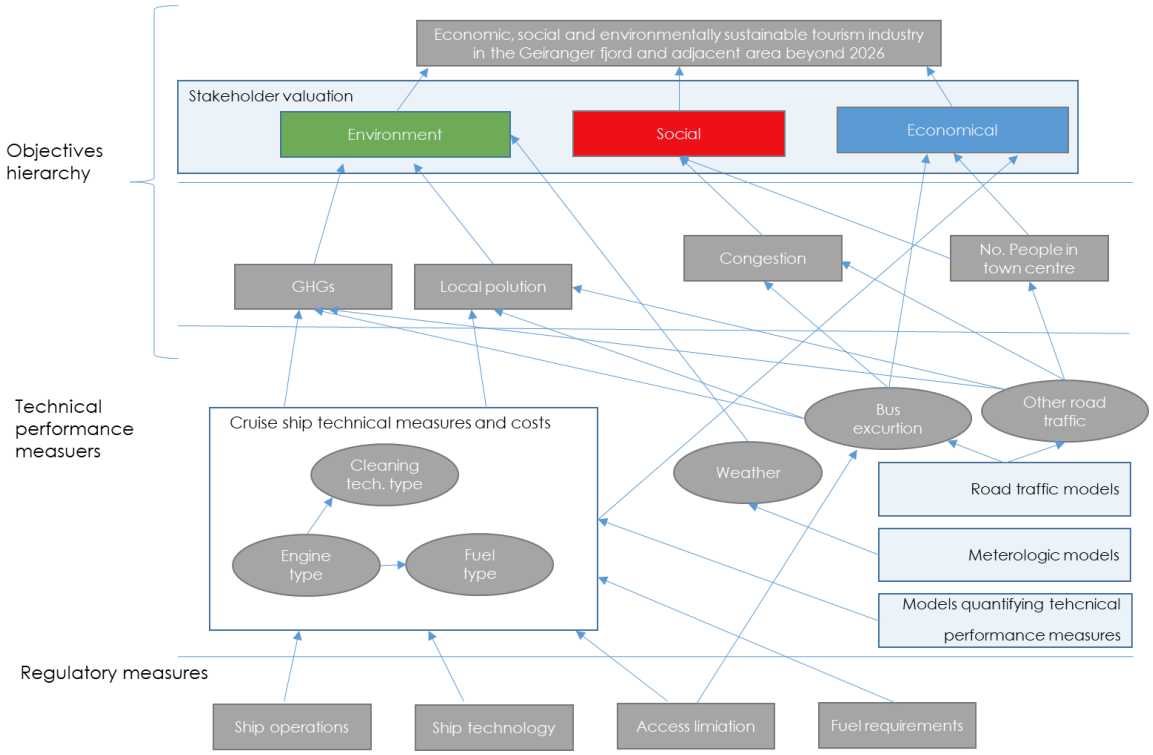


Figure 4: Example of Objectives Hierarchy, Technical Performance Measures and Regulatory Measures for Future Zero Emissions Sea Based Tourism Mobility in the Geiranger Fjord Area

The elements in the system are: a) sensitive to changes in regulation, b) to a varying extent dependent on each other, c) with different perception of utility among stakeholders, d) with varying preference among stakeholders. Regulatory measures will require technical performance measures in order to achieve compliance. With regards to cruise ships the technical performance measures could either be made through investment in cleaning technology as SCR/EGR or scrubbers or through change in operations as slow speed, limitations of port calls and choice of fuel. These measures are the least cumbersome to make reliable estimates of with regards to the measures' direct effect on the emissions. The trouble is to find the indirect effects that imposing these measure will have on the rest of the transport system and the different stakeholders' appraisal for the outcome for the environmental, social and economic criteria. One example which is discussed in relation to the Geiranger Fjord area is that if regulations on ships are too strict, then this will reduce the number of cruise calls and more tourists will arrive by bus instead and increase traffic congestion. Another example is if access limitation is introduced, this might reduce emissions and traffic, but would potentially reduce economic output.

DISCUSSION AND CONCLUSION

Both MCDM and systems thinking methodology require decision making to put into the proper context. "Who is the decision maker?" is according to MCDM methodology one of the first questions to ask when analyzing a decision process while in systems thinking one might ask who are the ones influencing and being affected by the system where decision are made? Both views are important to follow and in a systems thinking manner it is important to firstly understand the dynamics behind the elements of the systems being managed or considered for some act of change making. How will management or change influence the different stakeholders of the systems? Will the decision have the desired effects on the impacts of sustainability? Will the decision provide a balanced solution within all three dimensions of sustainability? These are important questions to guide the decision process. One formal decision maker in relation to cruise tourism destinations that both has a formal mandate and responsibility for its surroundings and stakeholders is the local municipality and subordinate port operation organization. The alternatives in future assessments should be analyzed and expressed with sustainability criteria and performances and the decision makers should provide with a weighting of the different criteria.

REFERENCES

Belton V., Stewart T. 2002. "Multiple Criteria Decision Analysis", Kluwer Academic Publishers 2002

Cooper D., Gustavsson T. 2004. "Methodology for Calculating Emissions from Ships: 1 Update of Emission Factors", Swedish Environmental Research Institute ordered by Swedish Environmental Protection Agency, Stockholm, February 2004

Cruise Lines International Association (CLIA). 2020. "Environmental Commitment, Innovation and Results of the Cruise Industry" September 2020

Dickinson J. E., Robbins D., 2008. "Representations of tourism transport problems in a rural destination". Tourism Management 29 (2008) 1110–1121

Diez-Gutierrez M., Babri S. 2020. „Explanatory variables underlying the route choice decisions of tourists: The case of Geiranger Fjord in Norway". Transportation Research Part A 141 (2020) 398-409

European Environment Agency (EEA), 2016. "Air quality in Europe — 2016 report", EEA Report No 28/2016 ISSN 1977-8449

Eyring V., Isaksen I.S.A., Berntsen T., Collins W.J., Corbett J.J., Endresen O., Grainger R.G., Moldanova J., Schlager H., Stevenson D.S., 2010. "Transport impacts on atmosphere and climate: shipping". Atmospheric Environment. 44, 4735-4771.

Freeman R. E. 1984. "Strategic Management: A Stakeholder Approach". Pitman Publishing 1984, Cambridge University Press 2010

Haugsbakk I., Tønnesen D. 2010. "Luftkvalitet i Geiranger – Sommeren 2010", NILU report O-110108, May 2011

International Maritime Organization (IMO), 2017. "Sulphur oxides (SOx) and Particulate Matter (PM) – Regulation 14"

International Maritime Organization (IMO), 2017. "Nitrogen Oxides (NOx) – Regulation 13",

Löffler J., 2016. "Annual Scientific Report 2016 - Long-Term Air Quality Monitoring Program UNESCO World Natural Heritage Geiranger Fjord", Department of Geography, University of Bonn, September 2016

Martinsen K., Torvanger A. 2013. "Control mechanisms for Nordic ship emissions", Nordic Council of Ministers, ISBN 978-92-893-2516-5

Miola A., Ciuffo B. 2011. "Estimating air emissions from ships: meta-analysis of modelling approaches and available data sources". Atmospheric. Environment. 45, 2242-2251.

National Aeronautics and Space Administration (NASA), 2007. Systems Engineering Handbook, NASA/SP-2007-6105 Rev1, December 2007

Norwegian Institute for Air Research (NILU), 2016. "Harmful air pollution requires innovative changes", http://www.nilu.no/Nyhetsarkiv/tabid/74/language/en-GB/NewsId/817/Harmful-air-pollution-requires-innovative-changes.aspx, accessed 18.10.2017

Norwegian Government, 2017. "Skal få ned forurensningen fra cruiseskip i verdensarvfjorder», https://www.regjeringen.no/no/aktuelt/skal-fa-ned-forurensningen-fra-cruiseskip-i-verdensarvfjorder/id2564118/, accessed 09.11.2017

Norwegian Maritime Authority (NMA). 2017. «Utslipp til luft og sjø fra skipsfart i fjordområder med stor cruisetrafikk», Report ver. 01 May 5th

Norwegian Parliament 2018, «Klimastrategi for 2030 - norsk omstilling i europeisk samarbeid» Meld. St. 41 (2016-2017), Innst. 253 S (2017-2018)

NRK. 2019. «Fryktar at nullutslepp for cruiseskip fører til busskaos». Published June 21st. https://www.nrk.no/mr/fryktar-at-cruiseskipa-i-geiranger-blir-erstatta-med-bussar-1.14575872

C. Schembari, F. Cavalli, E. Cuccia, J. Hjorth, G. Calzolai, N. Perez, J. Pey, P. Prati, F. Raes, 2012. Impact of a European directive on ship emissions on air quality in Mediterranean harbours. Atmospheric Environment. 61, 661-669.

Shlopak M., Bråthen S., Svendsen H. J., Oterhals O., 2014. «Grønn fjord Bind I. Beregning av klimagassutslipp i Geiranger», Rapport nr. 1413 (2014), Møreforskning Molde AS

Stenersen, D. 2017. «Operasjonsdata fra skipsfart i Geiranger, Nærøy- og Aurlandsfjorden - Datainnsamling fra cruiseskip og lokal trafikk», Norsk Marinteknisk Forskningsinstitutt AS, Report No 302002020-1, April 2017

Stranda Municipality. 2020. „Det store bildet». https://www.stranda.kommune.no/tenester/naring-og-skatt/anna-naringsliv/det-store-bildet/

Svendsen H. J., Bråthen S., Oterhals O. 2014. "Grønn fjord Bind I. Analyse av metningspunkt for trafikk i Geiranger", Rapport nr. 1412 (2014), Møreforskning Molde AS

Tallaksen H., Holm R. S., 2007. «Stedsanalyse Geiranger», Asplan Viak, november 2007

Weggeberg H., Stenersen D., Keskitalo T., Järvinen E., Strutz T. M., Polley D. A., Brashers B. 2017. „Utslipp til luft og sjø fra skipsfart i fjordområder med stor cruisetrafikk - Kartlegging og forslag til tiltak", M-rap-001-1350003037-002_Utslipp til luft og sjø fra skipsfart i norske fjorder_2017-05-02

World Health Organization (WHO), 2016. "Ambient (outdoor) air quality and health Fact sheet, Updated September 2016", http://www.euro.who.int/__data/assets/pdf_file/0006/189051/Health-effects-of-particulate-matter-final-Eng.pdf, accessed 18.10.2017

Yttredal E. R., Babri S., and Diez M. 2019. «Antall besøkende og kjøretøy i Geirangerområdet 2018». ISSN 1891-5973. Volda University College, Volda, Norway.

MODELLING ECONOMIC CRISES IN HUA HE FRAMEWORK

Nóra Felföldi-Szűcs
Corvinus University of Budapest
H-1093, Fővám tér 8, Budapest Hungary
nora.felfoldi-szucs @uni-corvinus.hu

Péter Juhász
Corvinus University of Budapest
H-1093, Fővám tér 8, Budapest Hungary
peter.juhasz @uni-corvinus.hu

Gábor Kürthy
Corvinus University of Budapest
H-1093, Fővám tér 8, Budapest Hungary
gabor.kurthy @uni-corvinus.hu

János Száz
Corvinus University of Budapest
H-1093, Fővám tér 8, Budapest Hungary
janos.szaz @uni-corvinus.hu

Ágnes Vidovics-Dancs
Corvinus University of Budapest
H-1093, Fővám tér 8, Budapest Hungary
agnes.dancs @uni-corvinus.hu

KEYWORDS
Hua-He methodology, corporate liquidity, agent based model, state intervention

ABSTRACT

In our paper we model firms' liquidity using the Hua He methodology. We investigate how cooperation of firms improve the possibilities of liquidity management. During a crisis, various effects identified in the literature hurt firms' liquidity position and lead to increased bankruptcy risk. We may counterbalance these adverse effects by providing immediate cash transfers and granting periodic cash flow transfers or additional credit lines. Cooperating with peers pays off. The importance of liquidity transfer between agents is higher during a crisis than in normal economic environment. It contributes to a lower default rate the losses are more moderate as well. The promotion of holdings of conglomerate-type or cross-ownership across local firms but with a wide variety of sectors may relieve some of the state burdens during a crisis. Several consequences can be drawn for policy makers how ameliorate resilience of agents.

INTRODUCTION

In our paper, we model how partnership and cooperation along the supply chain can contribute to better corporate liquidity and decrease the default frequency of participant firms. In ISCRM (Integrated supply chain risk management) literature, operational performance is the most frequently covered (Bredell, R. and Walters, J. 2017). Supply chain disruptions – a special topic within ISCRM - are partially as an unforeseen triggering event, like supplier bankrutcy (Bugert and Lasch, 2018). In our paper we model a less sever event than bankruptcy, we focus on liquidity shocks of interrelated agents. Specially, our scope of research within the large topic of financial flow of supply chain is solely whether common liquidity management like a cash pool agreement can improve the surviving ability of agents during a crisis. Since we are interested in the changes of default frequency depending on the different liquidity management practices, we treat liquidity as an exogen variable, and we disregard the specific reasons that change a liquidity position because they may be varied and complex. We describe liquidity developments as a random process for certain types of analysis, and we focus on the possible strategies of agents. The correlation of liquidity changes is the only representant of connection between the agents: A positive correlation can describe the case of agents in the same supply chain, a negative correlation can characterise competing agents. The scope of our paper is how the liquidity of correlated agents emerges with and without their cooperative liquidity management.

At that point, it is important to precise the terminology of agents' liquidity. In this paper, we assume that agents target a level of liquidity reserves or a level of cash reserves that has to be reached to maintain business continuity. We are not modelling the entire cash flow of the agents: we focus on that part of it, which is kept within the agent as operative cash to assure liquidity. Excess liquidity, the remaining part of created cash flow can be used for investments or paid out to creditors or shareholders. Therefore we assume that liquidity shocks have an expected value equaling zero. The growth of the cash flow of a prospering agent is not modelled here; it is part of the excess liquidity withdrawn from the scope of liquidity management. Rather than a sole figure, the targeted liquidity is an interval acceptable for the agent.

THE FRAMEWORK OF OUR MODEL

Literature suggests that a crisis may affect firms in at least four different ways. We may see (1) our sales and profit rate falling, and thus, expected cash flow might descend. During the Covid-19 crisis, restaurants, cinemas, and pubs had to stay closed, radically cutting expected returns. Simultaneously, (2) the uncertainty increases in the economy, so the cash flows' standard deviation may climb. The pandemic caused customers to stock up some food and detergents at the beginning that they consumed, later on, creating waves in otherwise steady demand.

Moreover, (3) activities following separate trends earlier may show the same development pattern. In other words,

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

the correlation between business units and firms may increase. Companies like travel agencies, shopping malls, and car-sharing services had seen little connections in their performance earlier, but because of the lockdown, their performance became more correlated, removing most of the diversification opportunities offered earlier. Finally, (4) banks may cut back on credit lines available and reduce outstanding loan amounts to limit their exposure to increased bankruptcy risk in the economy. Besides these effects, even the cost of financing (interest rates) may arise, but we will keep those stable in our runs.

Our simulation focuses on identifying the effects of these changes on a simple system consisting of three agents. These agents may be viewed as three business units of the same firm, three companies owned by the same investor, and three sectors of the same economy.

To model cooperation among the agents, we may allow them to offer their additional liquidity to each other during hard times. This option could represent a co-owned bank account, a cash pool system, or even rearranging state spending and tax incomes. Since there is an empirical evidence that trade credit positions tend to increase in a crisis situation throughout the supply chain. (e.g. as per central bank of Hungary's data trade credit volume increase by 4% between end of 2019 and end of Q3 2020 (MNB, 2020) We can interpret the increased level of trade credit among firms as a form of cooperation in managing liquidity even for not commonly owned firms as well.

Our research aims at identifying crisis consequences and the effectiveness of possible countermeasures targeting to lower the chance of bankruptcy. Our model offers three ways to improve the situation of the agents. We may (A) increase the expected cash flow by reducing fees, dues, and taxes to be paid or offering state subsidies over a while. Also, (B) an additional on-time monetary aid could be provided to raise the available amount of cash initially. Besides, (C) we may also offer additional outside financing sources like low-cost bank loans or credit lines.

MODELLING OF LIQUIDITY SHOCKS

Liquidity shocks are described by the Hua He methodology. (Hua He 1990).

It can be easily seen, that the X_1, X_2, X_3 variables have

a) zero expected value,
b) unit standard deviation, and
c) they are independent,

if their possible values are the ones in the table below, provided that each row is selected randomly with a q = 1/4 probability. Once a row is selected, all the 3 variables take their value from the selected row simultaneously.

a) *zero expected value*: the sum in each column is zero,
b) *unit standard deviation*: the sumproduct of each column with itself is 4 (this should be multiplied by q = ¼ to get the unit variance)
c) *Independence:* the sumproducts of any two different columns are zero.

Table 1. : X_1, X_2, X_3 variables

X_1	X_2	X_3
0	0	$\sqrt{3}$
0	$\sqrt{\frac{8}{3}}$	$-\sqrt{\frac{1}{3}}$
$\sqrt{2}$	$-\sqrt{\frac{2}{3}}$	$-\sqrt{\frac{1}{3}}$
$-\sqrt{2}$	$-\sqrt{\frac{2}{3}}$	$-\sqrt{\frac{1}{3}}$

Having 3 independent standardised random variables is its basic statistical procedure to create 3 variables with given covariance and expected value.

This way we have a non recombining tree with 4 branches in each step, and the 3 liquity values of the firms in each node. L(i, j, k) is the current size of the liquidity of company k in step i, on node j. We depart from the original Hua He type of tree, cutting it on both side at 0 and 200, not allowing a negative or unnecessary high level of liquidity.

Illustration: Two agents, with given correlation of the change of their liquidity

In this scenario, we need only 2 variables and 3 possible outcomes:

X_1	X_2
0	$\sqrt{2}$
$\sqrt{\frac{3}{2}}$	$-\sqrt{\frac{1}{2}}$
$-\sqrt{\frac{3}{2}}$	$-\sqrt{\frac{1}{2}}$

Since X_1 and X_2 are independent, the Y_1 and Y_2 variables will have a correlation ρ, if

$$Y_1 = X_1 \text{ és}$$

$$Y_2 = \rho X_1 + \sqrt{1-\rho^2}\, X_2$$

The next step is to adjust to the given variances of the changes in liquidity. We assume that the drifts of the liquidity changes are zero.

MODELLING OF AGENTS' LIQUIDITY MANAGEMENT

After having modelled the liquidity shocks, the liquidity policy has to be defined according to which the three agents are acting. The starting L0 level of liquidity reserves will be equally 100 for each of the agents. As liquidity shocks occur, agents have to respect some simple rules of liquidity management.

Assuming individually managed liquidity, the following rules are applied to each of the three agents:

1. The desired level of liquidity is 100.
2. If an agent's liquidity is less than 100, the agent has to apply for a bank loan.
3. Once the liquidity of the agents has reached 0, default occurs.

The commercial bank offers the following construction to the agents:

1. Agents under the liquidity of 100 can be financed by a loan.
2. The bank limits its exposure toward the individual agents: the maximum level of total outstandings for the same agent is limited to 50.
3. The bank collects an interest rate of ib=0.50% on the outstanding amount of the loan.
4. Repayment takes place in each of the periods where the agent's liquidity is above 100.

Allowing cooperation in liquidity management like a cash pool, the following rules are applied to each of the three agents:

1. The desired level of liquidity is 100.
2. If an agent's liquidity is less than 100, the agent has to apply for the cash reserves of related partner agents.
3. Above the desired level of liquidity, agents can provide their cash reserves to distressed partners.

The commercial bank offers the same construction to the agents as in the case of individually managed liquidity.

The characteristics of partner loans:

1. Partner loans are provided for one period (month),
2. at a rate of ip=0.25%.
3. Partner loans can be renewed if the issuer still has liquidity reserves above the level of 100.

As Diamond and Rajan suggest (Diamond - Rajan, 2001) the lender can face a liquidity shock as well. In our model we desregard from the illiquidity of commercial banks, we focus solely on the liquidity of the three firms/agents.

Order of financing and repayment:

1. If there are two distressed agents in the given period, the third will offer its liquidity surplus to the one facing a higher liquidity shortage.
2. In the case of two potential financing partners, the agent with the higher liquidity surplus will first provide partner loan to the distressed agent.
3. Agents first repay the partner loan of a higher volume.
4. Banks can lend to all the three agents at the same time.
5. Agents have to redeem first their bank loan.

After the occurrence of liquidity shock, agents assess their modified liquidity position and apply the above-listed elements of the model.

SIMULATION OF NORMAL AND STRESSED ECONOMIC ENVIRONMENT

Base case

First, we define the state of the world without any crises. Let us suppose all three agents have the same parameters. We keep the initial cash balance for all our runs for all agents at 100, which is equal to the cash need of the operation that the agents aim to maintain. (If falling below that level, companies try to attract additional cash.) The cost of borrowing from the peers (cash pool) remains at 0.25% per period (month) while that of the bank loan stagnates at 0.50%, implying a 12-period (yearly) rate of 3.04% and 6.16%, respectively.

All simulations last for 120 periods (10 years), and each Monte Carlo simulation covers 10 thousand individual runs. Those runs could represent alternative paths for a chosen group of firms and the development of a different set of agents in the same economy. Thus, the sum of the outcomes may be interpreted as a country-wide performance.

For the base case, the agents face an expected cash flow of 1 with a standard deviation of 10 for each period. The correlation among the cash flows of the agents is 0. The maximum credit line available with our bank is 100, as there is no cooperation among the agents (cash pool not available). (Table 2)

Table 2: No-crisis outcome matrix

Outcomes	A	B	C
1	1.0000	1.0000	18.3205
2	1.0000	17.3299	-4.7735
3	15.1421	-7.1650	-4.7735
4	-13.1421	-7.1650	-4.7735

Our results show that in 2.49% of the cases, at least one period existed at the end of which at least one firm had a negative cash balance. When considering the total of firm periods, only 0.22% ended with a bankruptcy even in the worst case. (In each period, we may count 0 to 3 firm bankruptcy periods.) Closing cash balance ranged from -205 to 636, with an average of 224 for the three agents. (Table 3)

Table 3: No-crisis case results

	Average	Min	Max
Bankruptcy firm-periods	0.00%	0.00%	0.22%
Closing Cash	224.02	-204.74	635.67
Closing Pool Debt	0.00	0.00	0.00
Closing Bank Loan	9.07	0.00	100.00

As the firms operate independently, any change in the correlation of cash flows remains without effect. When we allowed for cooperation (Table 4), the bankruptcy rate fell to 0.01%, while the maximum ratio of bankruptcy firm periods was 0.01%. The cash pool's existence raised

the minimum closing cash level but let the agents accumulate a considerable debt and deposit towards their peers.

Table 4: No-crisis with cooperation case results

	Average	Min	Max
Bankruptcy firm-periods	0.00%	0.00%	0.01%
Closing Cash	227.74	26.90	635.71
Closing Pool Debt	0.00	-311.62	241.44
Closing Bank Loan	0.79	0.00	100.00

As the pool added liquidity to the system, it is no wonder that bankruptcy became less frequent. At the same time, we may very well imagine that there is not too much pressure for the agents to cooperate once there are also transaction costs associated with teaming up as the expected advantages are moderate when just focusing on averages instead of considering the extreme values.

Crisis cases

As a next step, four crisis effects were simulated separately and in one joint case. The modified parameters were (1) expected cash flow cut back to 0, (2) standard deviation increased to 20, (3) correlation climbed to 0.4, and (4) maximum bank loan available decreased to 50.
When cutting back expected cash flow to 0, the bankruptcy rate jumped to 29.7% (Table 5). It is worth noting that while the average and the maximum cash balance has declined by almost the total of 120 (1 unit for each period) compared to the base case, the minimum closing cash had a slighter decline.

Table 5: Lower expected cash flow crisis

	Average	Min	Max
Bankruptcy firm-periods	0.02%	0.00%	0.48%
Closing Cash	127.82	-293.90	561.79
Closing Pool Debt	0.00	0.00	0.00
Closing Bank Loan	38.98	0.00	100.00
With cooperation			
Bankruptcy firm-periods	0.00%	0.00%	0.41%
Closing Cash	121.84	-130.76	566.06
Closing Pool Debt	0.00	-457.14	390.76
Closing Bank Loan	31.03	0.00	100.00

With cooperation allowed, the results were less extreme, and the bankruptcy ratio fell to 4.34%. We may see how cross-agent transfers enhance the surviving ability of the system. Thanks to the cheaper help received from peers, minimum closing cash also climbed radically. Here we see the advantages that we might seriously underestimate if considering non-crisis average performance only.
When the crisis increases the standard deviation of the cash flows, the extreme values may change radically. (Table 6) While the bankruptcy ratio boomed to 53.88, minimum decreased and maximum increase radically boosted inequality across firms without any fundamental differences.

Table 6: Higher standard deviation crisis

	Average	Min	Max
Bankruptcy firm-periods	0,07%	0,00%	0,76%
Closing Cash	240,53	-562,40	1 097,40
Closing Pool Debt	0,00	0,00	0,00
Closing Bank Loan	26,51	0,00	100,00
With cooperation			
Bankruptcy firm-periods	0,01%	0,00%	0,75%
Closing Cash	234,30	-323,69	1 098,75
Closing Pool Debt	0,00	-642,54	836,55
Closing Bank Loan	12,89	0,00	100,00

The cash pool cut back the bankruptcy rate to 7.89%, increased the minimum closing cash, but could not reduce the maximum of bankruptcy firm-periods and bank loan usage.
As explained earlier, without allowing for cooperation, the change in correlation has no mathematical effect in the model, as, e.g., the bank loan available for the agents does not depend on the amount taken by their peers, like it would in real life. So, an increased correlation level only limits the positive effects of cooperation.

Table 7: Higher correlation crisis

	Average	Min	Max
Bankruptcy firm-periods	0.00%	0.00%	0.48%
Closing Cash	206.89	-179.35	671.27
Closing Pool Debt	0.00	0.00	0.00
Closing Bank Loan	13.58	0.00	100.00
With cooperation			
Bankruptcy firm-periods	0.00%	0.00%	0.44%
Closing Cash	216.36	-120.78	607.16
Closing Pool Debt	0.00	-220.66	199.22
Closing Bank Loan	5.23	0.00	100.00

Table 8: Lower bank loan limit crisis

	Average	Min	Max
Bankruptcy firm-periods	0,01%	0,00%	0,33%
Closing Cash	218,88	-168,97	660,87
Closing Pool Debt	0,00	0,00	0,00
Closing Bank Loan	5,76	0,00	50,00
With cooperation			
Bankruptcy firm-periods	0,00%	0,00%	0,20%
Closing Cash	221,71	-69,50	681,97
Closing Pool Debt	0,00	-286,24	366,43
Closing Bank Loan	1,13	0,00	50,00

While the bankruptcy rate was 5.65% in our MC, no major differences could be identified then the initial results in Table 3. (Table 7) When cooperation was allowed, the bankruptcy rate declined to 0.85%, far higher than the 0.22% we received in the no-crises scenario. The lowest closing cash is well below the level estimated (Table 4) with the initial parameters. These results call for the policymakers to aim at a well-diversified economy and promote holdings interested in less interlinked business fields. Cooperation of suppliers-and buyers or competitors could be less fruitful. Thus, creating interconnected supply chains in the same country may not be optimal from the liquidity risk point.
Finally, limiting bank loans particularly hit the firms when no cash pool system was available. The bankruptcy rate climbed to 9.25% (base case: 2.49%) while closing cash and bank loan data were not affected.
With the cash pool system, the bankruptcy ratio was reduced to 0.29% that was still considerably higher than the 0.01% in the base case. Also, bankruptcy firm-periods increased in proportion.

Table 9: Complex crisis

	Average	Min	Max
Bankruptcy firm-periods	0,20%	0,00%	1,12%
Closing Cash	110,96	-727,10	976,81
Closing Pool Debt	0,00	0,00	0,00
Closing Bank Loan	24,06	0,00	50,00
With cooperation			
Bankruptcy firm-periods	0,12%	0,00%	1,12%
Closing Cash	129,74	-594,35	997,99
Closing Pool Debt	0,00	-542,18	565,05
Closing Bank Loan	21,77	0,00	50,00

When all crisis effects appeared at once in our model, consequences became radical. The bankruptcy rate reached 76.87%, average closing cash fall, while the distribution range doubled.
Adding the possibility of cooperation to the model reduced the proportion of the bankruptcy cases to "only" 44.75%. The difference between the extreme values got smaller but was still dramatically boosted. (Table 9)

CRISIS MANAGEMENT OPPORTUNITIES

Seeing the majority of the agents failing is usually unacceptable for policymakers. Our model allows for three types of anti-crisis actions. Offering a one-time monetary help would boost initial cash reserve, reducing taxes, and providing transfers would push up expected cash flow, while offering additional loans will extend our bank credit lines.
Next, we review what measures would be necessary to counterbalance the complex crisis. The aim is to reduce the bankruptcy ratio to a similar level we experienced in the base scenario.

Table 10: Crisis management with initial cash aid

Startup cash	Bankruptcy rate	
	Without cash pool	With cash pool
100	76.87%	44.75%
200	49.24%	18.99%
300	21.16%	7.75%
400	9.54%	1.94%
500	4.26%	0.34%
600	2.67%	0.05%
700	0.01%	0.00%
w/o crisis	2.49%	0.01%

Results show that offering a startup remedy of 100 cash units (doubling the liquidity) would push down the non-cooperative case's bankruptcy rate to the cooperative scenario level. An additional 100 units would be still only enough to reach a 21.16% level. Altogether, we need to add almost 500 extra cash units (total available: 600) to keep the agents as safe as before the crisis. (Table 10)
When focusing on periodic transfers, cooperation lowers the needed support. As increased correlation has hit only the cooperative case, lowering the bankruptcy rate to the standard (before-crisis) level requires far more state support when allowing for the cash pool. (Table 11)

Table 11: Crisis management with periodic cash aid

Expected period cash flow	Bankruptcy rate	
	Without cash pool	With cash pool
0	76.87%	44.75%
1	57.64%	30.16%
2	35.15%	11.48%
3	19.02%	7.87%
4	11.09%	1.86%
5	4.03%	1.58%
6	1.56%	0.20%
7	0.45%	0.16%
8	0.17%	0.15%
9	0.12%	0.01%
w/o crisis	2.49%	0.01%

Last, we also estimated the needed increase in bank loans available. Here again, the state has to offer more aid in the cooperative case to return to the before-crisis level to counterbalance the effect of higher correlation that does not affect the non-cooperative case. (Table 12)
Our results show that completely rebalancing the crisis effect would cost us (600-100)=500 units of cash, or (6-0)=6 units of periodic transfers, or (600-50)=550 units of credit line for each of the agents in the non-cooperative case. For the cooperative alternative, shadow costs equal (700-100)=600 cash, (9-0)=9 periodic transfer, or (800-50)=750 units of surplus in the credit line. In other words, in a cooperative system, the relative cost of using initial cash transfer is lower than in a non-cooperative economy.

Also, additional credit lines are relatively better than periodic transfers.
Our results allow us to express the value of cooperation among agents in startup cash (100), periodic subvention (1.5), and additional credit line (150) by contrasting the cooperative and non-cooperative cases for bankruptcy rates. As the agents' cross-financing always adds value, it is advisable for policymakers to introduce or enhance cross-sector and cross-company transfers, e.g., using additional income due to increased corporate tax rates to support companies in serious need.

Table 12: Crisis management with an extended credit line

Credit line available	Bankruptcy rate	
	Without cash pool	With cash pool
50	76.87%	44.75%
100	61.63%	31.96%
200	43.63%	20.16%
300	27.00%	8.43%
400	17.31%	3.59%
500	7.02%	1.14%
600	2.93%	0.28%
700	1.06%	0.20%
800	0.13%	0.05%
w/o crisis	2.49%	0.01%

Nevertheless, would a combination of the possible intervention measures offer a better solution than single-measure solutions? We identified some complex crisis management packages leading to a bankruptcy rate similar to the before-crisis status. First, we set back two of the three parameters to their initial level and checked how the remaining effect of increased standard deviation and correlation could be compensated with the single leftover parameter. (Table 13)

Table 13 Crisis management with an increase in the credit lines available

	Initial cash	Expected periodic cash flow	Credit line available	Bankruptcy rate
w/o crisis	100.00	1.00	100.00	2.49%
A	100.00	1.00	**500.00**	1.95%
B	**400.00**	1.00	100.00	2.51%
C	100.00	**5.00**	100.00	1.51%
With cash pool				
w/o crisis	100.00	1.00	100.00	0.01%
A	100.00	1.00	**700.00**	0.01%
B	**600.00**	1.00	100.00	0.00%
C	100.00	**7.00**	100.00	0.02%

When contrasting results in Table 13 with the no-crisis case, we may estimate the cost of added standard deviation and correlation. Accepting that bankruptcy rates are almost the same, the increased standard deviation has a shadow price of (400-100)=300 initial cash, (5-1)=4 additional cash flow over all the simulated periods, or (500-100)=400 units of additional credit line. Based on this, assuming additivity, the increased correlation that hits only the cooperative case would cost (600-400)=200 initial cash units, (7-5)=2 units of additional periodic cash flow, or (700-500)=200 units of surplus credit line.
When contrasting the total price of the crisis calculated earlier, we may see that 60-70% of the crisis-management costs is linked in our example to the higher standard deviation. This result may be interpreted as stabilising the markets is more important than regaining profitability or improving liquidity immediately.
As we may have various alternatives to reset economic stability, we should also compare costs and other consequences of using those alternatives. Our simulation covered 120 periods, so when assuming a positive cost of capital for the financing, 1 unit of periodic transfers has a maximum present value of 120. Our results showed that 500 units of initial cash subvention have similar effects as 6 units of additional periodic transfers in the non-cooperative case. When rates are low, immediate cash aid might be preferred, while higher rates may make periodic transfers cheaper. During a crisis, we usually see inflation and risk premium climbing. Thus, periodic transfers could offer a cheaper solution. Another argument for choosing periodic transfers would be that we can quickly stop those if the crisis ends earlier than initially assumed.
When focusing on providing additional credit lines, those seem to be even more attractive. The main reasons for that include (1) we do not have to provide the total of the credit line in the form of loans to all firms immediately and for all the periods, and (2) loans are repaid sooner or later and earn us interest during their lifetime. As credit lines needed to manage the crisis are only slightly larger than initial cash transfers, we may consider them the cheapest alternative. Simultaneously, this method may call for very different conditions than applied outside of crisis periods as we should provide the estimated amount of loans even without adequate collaterals, probably for the total length of the crisis without any forced repayment.

SUMMARY AND CONCLUSIONS

Our paper analysed the crises effects on our three agents' model. As the literature suggests, during a crisis, not only expected cashflow might decrease, and the standard deviation of those may climb, but the correlation among various actors' performance could increase if used the Hua He model to simulate the random payoffs of the agents. Our results build on Monte Carlo simulations with 120 periods and 10 thousand runs. Our most important results are as follows.
1. Various effects of crisis identified in the literature hurt firms' liquidity position and lead to increased bankruptcy risk.

2. We may counterbalance these adverse effects by providing immediate cash transfers and granting periodic cash flow transfers or additional credit lines.
3. Cooperating with peers pays off. Our results illustrate why it is dangerous to consider the advantages of cooperation based only on records from non-crisis periods. It is during crisis periods that we may see how vital help from fellow firms may be. Thus, to mitigate risks, the state should promote such cooperation, and it might be justified to use cross-sector transfers to stabilise the economy.
4. Voluntary cooperation or forced reallocation across firms helps the economy to perform better during crises. That is why policymakers should promote the establishment of holdings or cross-ownership across local firms. Crises more jeopardise standalone firms and put more jobs at risk there.
5. Cooperation among firms with less correlated business performance is more advantageous than for other companies. However, it is an already widespread consequence that cooperation or integration within the supply chain (among collated firms) can reduce transaction costs through the reduction of uncertainties in normal market circumstances as well (see for example: Zhao, Huo, Sun, and Zhao, 2013), we find that cooperation like holdings created by conglomerate-type mergers and economies with a wide variety of sectors can survive crises with less loss. An economic policy giving a unique preference to investments or FDI in a few interlinked sectors (car manufacturing, tourism) aiming to boost the GDP may cut back on its crisis-resistance.
6. Should the before-crisis status involve cooperation, the state must provide more aid to make the economy return to the initial level during the crisis as the increased correlation needs counterbalancing. In non-cooperative economies, creating holdings and introducing cross-sector or cross-firm transfers during a crisis may relieve some of the state burdens.
7. A crisis enhancing standard deviation would increase inequality across firms with no fundamental differences. These random effects may annul any competitive advantages leading to a higher-than-normal survival rate for the less efficient and lower-than-normal survival rate for the more efficient companies. Thus, we may see an efficiency loss across the whole economy. To evade those losses, the resilience of lending and transfer rules has to be boosted when facing hard times. Often, we see the contrary in commercial banks so that regulatory interventions could be justified.
8. During the crisis-management, a particular focus should be given to reduce fluctuations in the economy by upkeeping laws and evading panic. A less hectic environment may dramatically cut back on the loss that the crisis may cause.
9. Providing an increased credit line with very lax conditions (no collaterals, extreme duration) may offer the cheapest risk-management alternative.

REFERENCES

Bredell, R. and Walters, J. 2017. “Integrated supply chain risk management”. Journal of Transport and Supply Chain Management Vol.1, No.1 (Nov), 1-17.

Bugert, N. and Lasch, R. 2018. “Supply Chain Disruption Models: A Critical Review”, Logistics Research, Vol.11, No.5 (Jun), 1-35.

Diamond, D. W. - Rajan, G. R 2001. “Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking.” Journal of Political Economy, Vol. 109, No. 2: 287-327

Hua He 1990. “Convergence from Discrete- to Continuous-Time Contingent Claims Prices.” *The Review of Financial Studies* 3(4), 523-546.

Zhao, L.; Huo, B.; Sun, L.; Zhao, X. 2013. “The impact of supply chain risk on supply chain integration and company performance: a global investigation”, Supply Chain Management: An International Journal, Vol.18, No.2, 115-131.

AUTHOR BIOGRAPHIES

PÉTER JUHÁSZ works as an Associate Professor for the Department of Finance at Corvinus University of Budapest (CUB). He is a CFA charterholder and holds a PhD from CUB. His research topics include business valuation, financial modelling, and performance analysis.

ÁGNES VIDOVICS-DANCS, PhD, CIIA is an associate professor at the Department of Public Finance and Banking 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. Since 2015, she is the chief risk officer of a Hungarian asset management company.

JÁNOS SZÁZ, CSc is a full professor at the Department of Public Finance and Banking at Corvinus University of Budapest. He was the first academic director and then president of the International Training Center for Bankers in Budapest. Formerly he was the dean of the Faculty of Economics at Corvinus University of Budapest and President of the Budapest Stock Exchange. Currently, his main field of research is financing corporate growth when interest rates are stochastic.

GÁBOR KÜRTHY, PhD, is an associate professor and head of the Department of Public Finance and Banking at Corvinus University of Budapest. Once he applied for the job of village idiot of South-Buda. He got down to the last two, but he failed the final interview, because he turned up. The other bloke was such an idiot he forgot to. And if you’ve read this, you may have a clue what a (liquidity) shock is.

NÓRA FELFÖLDI-SZŰCS, she has served as a lecturer at Corvinus University since 2006. She obtained her PhD at CUB in 2013. From 2015 to 2020, she has been a researcher at John von Neumann University. Her primary field of interest covers microfinance, credit risk, and contract theory.

DISCRETE EVENT SIMULATION OF THE COVID-19 SAMPLE COLLECTION POINT OPERATION

Martina Kuncová, Kateřina Svitková, Alena Vacková and Milena Vaňková
Department of Econometrics
University of Economics in Prague
W.Churchill Sq. 4, 13067 Prague 3, Czech Republic
E-mail: martina.kuncova@vse.cz; svitule10@seznam.cz; vackovalena@gmail.com; vankova.mila@email.cz

KEYWORDS
Discrete event simulation, healthcare, COVID-19, sample collection point, SIMUL8..

ABSTRACT

The year 2020 was very challenging for everyone due to the COVID-19 pandemic. Many people turn their lives upside down from day to day. Politicians had to impose completely unprecedented measures, and doctors immediately had to adapt to the huge influx of patients and the massive demand for testing. Of course, not all processes could be planned completely efficiently, given that the situation literally changes from minute to minute, but sometimes better planning could improve the real processes. This contribution deals with the application of simulation software SIMUL8 to the analysis of the COVID-19 sample collection process in a drive-in point in a hospital. The main aim is to create a model based on the real data and then to find out the suitable number of other staff (medics) helping a doctor during the process to decrease the number of unattended patients and their waiting times.

INTRODUCTION

Although simulation modeling is a relatively widespread tool and the use of ICT in developed countries is a common part of everyday life, in reality we still face problems in which simulation modeling could help avoid unpleasant effects - and yet no similar analysis has been made. Even at a time when we face many restrictions related to the COVID-19 pandemic and when the implementation of various measures and changes is often very rapid, we should not forget the benefits of simulation modeling and discrete event simulation before launching new projects or processes. Due to its probabilistic and dynamic aspects, a realization of experiments with the simulation model helps the decision-maker set the processes in a better way than without the model, especially when no previous similar situation was not tested.

In the Czech Republic, unlike other countries, simulation modeling in the healthcare environment is not a common part of introducing new processes and changes. According to the review of approximately 250 high-quality journal papers published between 1970 and 2007 on healthcare related simulation made by Katsaliaki and Mustafee (2011), no paper was connected with the Czech Republic healthcare system. Most of the papers used Monte Carlo simulation models, but discrete event simulation was also mentioned in 20 % of the papers analysed. Brailsford, Carter and Jacobson (2017) commented the situation of 50 years of simulation modelling in healthcare context – they agreed that healthcare was a prolific application area for simulation modeling ever since the very early days of computer simulation and during the 5 decades a lot of models using Monte Carlo simulation and discrete event simulation was done, mainly in USA and UK. Walsh et al. (2018) showed that among the 100 most cited articles on simulation in healthcare, 88% of articles are from the USA, UK and Canada. So there is still much space for improvement in other countries, including the Czech Republic.

Most of the papers used Monte Carlo simulation models, but discrete event simulation was also mentioned in 20% of the papers analysed by Katsaliaki and Mustafee (2011). Van Buuren et al. (2015) presented a detailed discrete event simulation model for emergency medical services call centers. As Hamrock et al. (2013) stated, discrete event simulation in healthcare commonly focuses on improving patient flow, managing bed capacity, scheduling staff, managing patient admission and scheduling procedures, and using ancillary resources (e.g., labs, pharmacies).

This paper focuses on improving patient flow and better staff scheduling in the drive-in COVID-19 sample collection point. Our goal is to analyze the situation and verify whether certain aspects of the process could not be better addressed or whether it was not possible to propose an alternative approach. The main aim is to create a simulation model and analyze the impact of the number of resources' changes on queuing time. For the model, SIMUL8 software is used.

SIMUL8

SIMUL8 is a software package designed for Discrete Event Simulation or Process Simulation and developed by the American firm SIMUL8 Corporation (www.simul8.com). The software started to be used in 1994, and every year a new release has come into being.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

A visual 2D model of an analyzed system can be created by placing objects directly on the screen. This software is suitable for discrete event simulation (Shalliker and Ricketts 2002). SIMUL8 uses 2D animation only to visualize the processes, but for the given problem, this view is sufficient, especially because queue analysis is important.

SIMUL8 belongs to the simulation software systems widely used, especially in industry (Greasley 2003), but several case studies were aimed at analysing the queues in the healthcare processes. Pisaniello et al. (2018) used SIMUL8 to develop the simulation model of the call center in the children's hospital, they demonstrated the meaning of the application of validation and verification techniques as the most critical aspects of the simulation modelling process. Viana et al. (2014) showed the usage of SIMUL8 for discrete event simulation in combination with system dynamics in VENSIM software to analyze how the prevalence of Chlamydia at a community level affects (and is affected by) operational level decisions made in the hospital outpatient department.

SIMUL8 main components

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 Point, Resource (Concannon et al. 2007).

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), an image of the item (showed during the animation of the simulation on the screen) and advanced properties (multiple Work Item Types).

Work Entry Point: an 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 the Work Items Label or setting the following discipline (Routing Out).

Storage Bin: queues or buffers where the Work Items wait before the next processes. It is possible to define the capacity of the queue or the shelf life as time units for the expiration.

Work Center: main object serving for the activity description with the 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: an object that describes the end of the modeled system in which all the Work Items finish their movement through the model.

Resource: objects that serve to model limited capacities of the workers, material or means of production used during the activities.

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.

After the system is modelled, the 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, several trials can be run and 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 also be seen on the website www.simul8.com.

Our experience shows that SIMUL8 is easy to learn when only 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 but also for the simulation of various production processes (Fousek et al. 2017).

PROBLEM DESCRIPTION

The impetus for the creation of the model presented in this paper was the experience of one of the authors in the test for COVID-19 as there were long queues at a selected drive-in center in one of Prague's hospitals. Therefore, we decided to analyze the problem and use a simulation model to assess the number of service personnel changes to reduce customer waiting time.

HOSPITAL DATA

The model is focused on the analysis of the drive-in COVID-19 sampling point in one of the Prague's hospitals. The drive-in collection point is used for patients arriving for the test by car. The hospital allows people to order a test using the online form (e-request) only. Examination for coronavirus infection is performed either on the basis of an indication by a general practitioner or the Regional Hygiene Station or without any indication as a self-payer. A patient who orders a drive-in test at their own expense must pay for the test online, and a patient who has a test request from a general practitioner does not have to pay. Therefore, payment is not included in the model at all, because it takes place when ordering (online), or it does not take place at all.

Each patient must be booked, the collection point does not accept unordered patients. Unfortunately, there are still cases where an unordered patient appears in the

queue. The hospital also offers a walk-in collection point, and occasionally a patient who goes for a walk-in test also appears at the drive-in collection point. All these situations must be included in the model.

According to the hospital's information, tests are performed from 8:00 to 11:55 and from 13:00 to 16:55 at intervals of five minutes, or from 18:00 to 21:56 at four minute intervals (FNKV.cz 2021). Based on the experience of the hospital's doctor (obtained from the authors' interview with the doctor), patients arrive at the collection point approximately at the time of order, but the order of the cars may not match the order of ordering, as some patients arrive in advance, some on time and some even with a delay. However, it is not possible to change the order in the car queue, so patients are admitted in the order in which they arrived. Upon arrival at the collection site itself, the doctor first finds out whether the patient has been ordered, whether all necessary data is in the patient form and whether a request has been sent from a general practitioner or a patient is a self-payer. If everything is in order, the doctor will take a sample and store it. Then the patient leaves the drive-in. Occasionally, there may be complications when the patient is ordered for a walk-in and not for a drive-in, or he/she is not ordered at all, a request is not sent from a general practitioner. The differences are also when the patient is a child and not an adult. It takes longer to take samples with a child, as it is more difficult to take samples.

Based of our knowledge of the process and according to the discussion with a doctor and patients' interviews (interviews with several patients were conducted by the authors at the sampling site) we collected these kind of data:

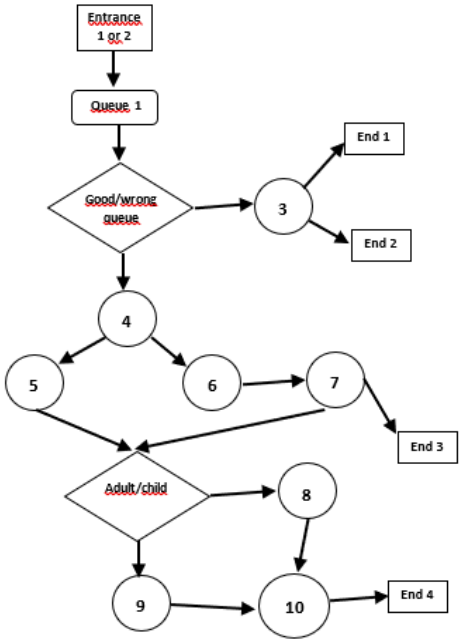
- No request for a doctor or a payment from the self-payer can be found in 5% of the patients ordered. If a practitioner has to be called, the problem will be solved in 80% of patients, so the request will be sent immediately. In 20% of cases, the patient must go either to the end of the queue or home again and order for another day.
- 4% of patients are in the wrong queue, half of them should be in the walk-in and half is not ordered.
- 85% of patients are adults and 15% of patients are children.
- Doctors and medics also work during the lunch and evening breaks plus overtime (50 minutes of service) if patients are still waiting in line.
- Although patients are booked on a specific time, they do not always arrive on time, so an exponential distribution can be used for the intervals between arrivals.

The model should include only activities related to the process at the drive-in collection point. The list of activities, including probability distributions for durations, is given in Table 1. Symbol EXP is used for exponential distribution with the mean value for interarrival times. T is used for the triangular distribution with lower, mode and upper time limits. U is for uniform distribution with minimum and maximum duration, N for normal distribution with mean value and standard deviation of the duration.

Table 1: List fo Activities and Probability Distributions

Activity No.	Probability distribution and times in minutes
1.arrivals during the day	EXP (5)
2.arrivals during night	EXP (4)
3. wrong queue activities	T (1;2;6)
4. database search	T (0,5;1;3)
5. well entered form	T (1;3;5)
6. badly entered form	T (3;5;7)
7. calling a practitioner	U (1.5;3)
8. adult sampling	T (0.5;1;2)
9. child sampling	T (2;2.5;4)
10. sample storage	N (1;0.25)

The scheme of the process with all 10 activities is on the Figure 1.



Figures 1: Scheme of the Process

The patient comes to the drive-in center and stands in Queue 1. Afterward, the doctor or medic controls the ordering form in the database. If there is some problem (activity 3; 4% of patients) - a patient is in the wrong queue (he/she should be in the Walk-in instead) or if he/she is not ordered – the doctor or medic sends the patient to the Walk-in center (50%; end1) or home to order electronically (50%; end2). For those correctly ordered for the drive-in (96%), the request/order form is

checked (activity 4). Patients, who have everything correctly filled in, are prepared for sampling. (activity 5) For those who do not have the correct request from a general practitioner, it is necessary to try to supplement the information by calling a general practitioner (activities 6 and 7). When successful, the patient is ready and continues testing (activities 8 or 9). If the information cannot be completed, the patient is sent re-ordered for another day (end3). Finally, after testing, the sample is stored (activity 10) and the patient can go home (end4).

MODEL IN SIMUL8

The simulation model was developed in SIMUL8 software. The aim of the simulation is to find a suitable number of doctors and medics to serve patients so that the queue length does not exceed 30 cars and the waiting time is acceptable. We have tested 3 models:

- Model 1 – 1 doctor, no medic – this model corresponds to the real situation in the selected hospital
- Model 2 – 1 doctor, 1 medic
- Model 3 – 1 doctor, 2 medics

Entities in all 3 models are patients generated via the exponential distribution with 5 minutes or 4 minutes interarrival times during the working hours. According to Table 1, activities are modeled as work centers (as an example of activity 5 see Figure 2). In Model 1 only 1 resource (doctor) is used (see Figure 3), in Model 2 one medic is added. The so-called pooled resource is created – it means doctor and medic together when part of the activities can be done by any of them (see Figure 4). Only the sample collection and storage activities are made by a doctor. Model 3 is similar to Model 2, only 2 medics instead of 1 are used.

Figure 2: Activity Settings

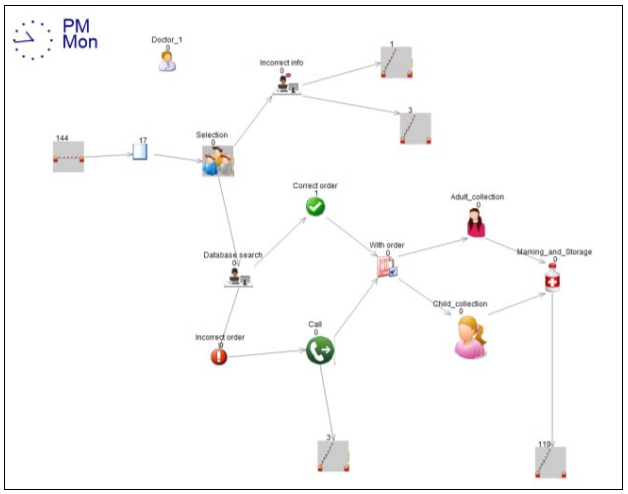


Figure 3: Model 1

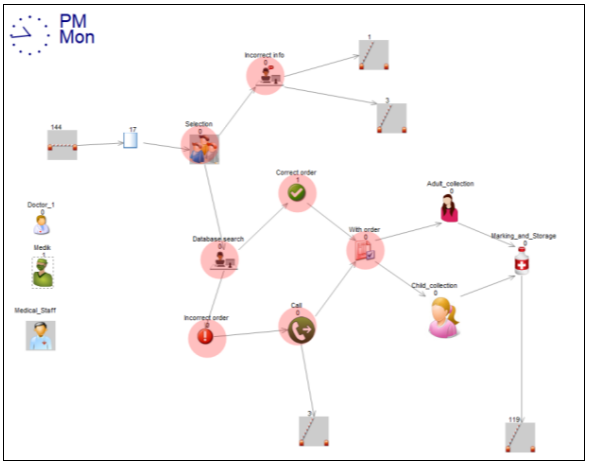


Figure 4: Model 2

RESULTS

After 1000 experiments with Model 1 the average of 30.7 unattended patients at the end of the day (see Figure 5) was identified, with a 95% confidence interval (29.96; 31.40). This is an alarming result, as this system would not be sustainable - if about 30 patients were not served every day, they would have to book another day and the demand for consumption would increase very quickly.

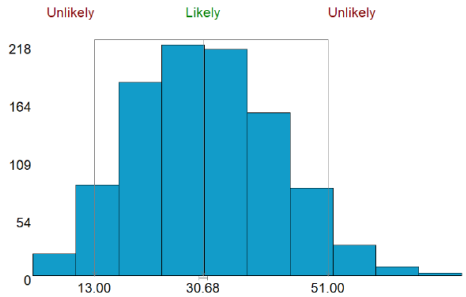


Figure 5: Model 1 - Curent Content of the Queue 1 (1000 runs)

The average queue length is 15.7 cars. Although this means a relatively long line, it would not cause serious problems in traffic. However, when we focus on the maximum queue length, we find that the 95% confidence interval is (37.04; 38.46). This means that there is a high chance that a queue could be longer than 30 cars, which will significantly complicate traffic.

The average time spent in the queue is about 83 minutes (see Figure 6) which is not acceptable.

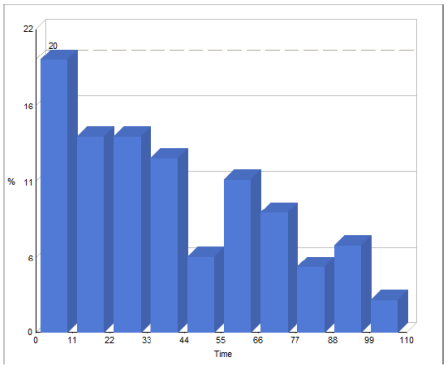


Figure 6: Time in the Queue 1

A doctor occupancy is on average up to 98%. This unfavorable result means that a doctor should work up to 14: 45h a day without any breaks, and yet patients remain unattended after the closing time. Definitely, this result is not satisfactory even if the doctors take turns. We recommend including the medic in the model to help with the administration.

In Model 2 a medic was added. He/she, however, cannot take samples and then label and store samples, these activities will remain with the doctor. Other activities can be performed by both a medic and a doctor. This allows two patients to be served at the same time (eg, the doctor takes a sample from the first patient while the medic checks the second patient's request).

The average number of unattended patients dropped rapidly to just 0.12 per day. Also, the estimate of the average queue length (see Figure 7) was reduced to only 1.05.

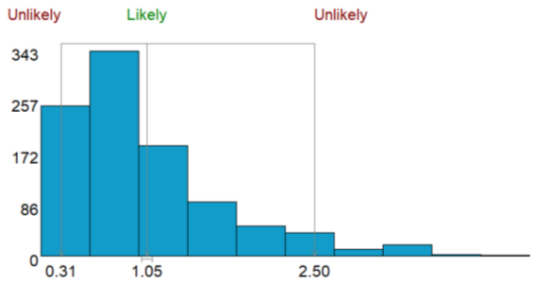


Figure 7: Model 2 - Curent Content of the Queue 1 (1000 runs)

The average workload of the doctor decreased to 68.44%, the confidence interval is (68.11; 68.76). We consider this workload reasonable, as the doctor is not busy all day and has time for lunch/dinner or a short rest. The workload of the medic can also be considered reasonable (see Figure 8). The model estimates the average workload of the medic to be 58.87%.

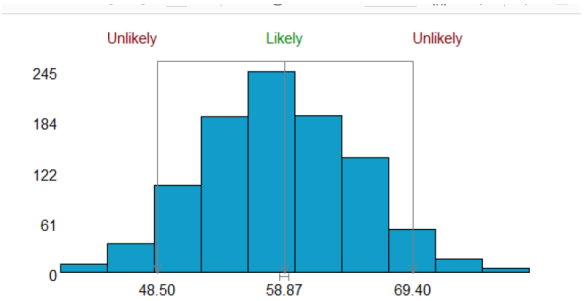


Figure 8: Model 2 – Medic Usage (1000 Runs)

The average time in the system was significantly reduced to 17.76 minutes. At the same time, the average time in the queue was shortened to 5.78 minutes. The estimate of the 95% confidence interval of the average time in the queue is then (5.57; 6.00). The average waiting time in the queue can thus be considered acceptable.

Model 3 included two medics. In this case, the doctor only take samples and then mark and store the sample. Other activities are in charge of two medics. This allows up to three patients to be served at the same time, which of course shortens the queue and speeds up the system. Other settings remain the same as in Model 1.

With this solution, in no simulation experiment did an unattended patient remain in the queue for 50 minutes after the end of working hours. The average queue length was about 0.194 cars. Of course, the average time in the queue was also reduced to 1.08 minutes. The average time in the system of the patient who was sampled is 11.02 minutes. The workload of the doctor is the lowest in this model, about 34.16% on average. The workload of medics has also decreased compared to the previous model, as there are now two administrative activities, so the workload was 43.16% on average.

Finally, we compare the results of the individual models. Estimates of all significant average characteristics are worst for Model 1, while best for Model 3 (see Table 2). The basic model (Model 1) is very bad in all respects – but unfortunately it corresponds to the reality that one of the authors experience. The differences between the average time spent in the system and in the queue, the average queue length, or the number of unattended patients are not significant between the remaining models. Both Model 2 and Model 3 are acceptable with regard to queue length and waiting time. If the hospital wants to speed up the process with patients not spending

more than a quarter of an hour on the tests, Model 3 is more appropriate, provided that the hospital has 2 medics available to involve in the process. When 1 medic is involved, his/her workload and the doctor's workload will be higher but still acceptable. Patients would spend a little longer in the process, but still, that time would be much shorter than in real practice shown in Model 1.

Table 2: Comparison of the Average Results of All Models

Average results	Model 1	Model 2	Model 3
No. of doctors	1	1	1
No. of medics	0	1	2
Unattended patients	30.78	0.12	0.00
Queue length	15.70	1.05	0.19
Busy doctor	97.91	68.44	34.16
Busy medics	x	58.87	43.16
Patient time in system (minutes)	91.57	17.76	11.02
Queuing time	83.53	5.78	1.08

CONCLUSION

The aim of the contribution was to demonstrate the applicability of SIMUL8 on the drive-in COVID-19 sample collection point process in a hospital. The model was based on available information given by patients and doctors of the hospital and one author's experience. The simulation model shows the real situation with very long queues, a lot of unattended patients and a long time spent in the hospital. Only a small change – adding 1 medic to help the doctor – could rapidly improve both hospital time and queue length. Unfortunately, the deployment of doctors and medics is done on an ad hoc basis.

Our results show that only a small change in the system can significantly benefit the situation. In this case, it is a more efficient division of labor into the administrative part and the sample collection itself. We demonstrate that the use of simulations has a real use even in crisis situations where there is not enough time to analyze the impacts of the selected system. Thanks to the simulation, it is possible to see whether the proposed change would have a large or negligible impact on the overall societal benefit, whether in terms of doctor's workload, time spent in the queue, or a negative impact on traffic in adjacent streets.

As Brailsford, Carter and Jacobson (2017) mentioned: active stakeholder engagement in the modeling process is a critical success factor for a healthcare simulation model to be useful in practice and despite major advances in both software and hardware, there is still a general lack of implementation of simulation in healthcare, compared with other sectors such as manufacturing industry or defense. As this paper describes, a relatively simple simulation model can very quickly show the effects of changes and its use would be beneficial to both doctors and patients, and thus for the hospital as a whole.

It would certainly be interesting to analyze the situation in other hospitals, but this would require access to data, which is not always easy or possible. However, the presented analysis can also help raise awareness of the possibilities of using simulation models in healthcare.

ACKNOWLEDGEMENTS

This work was supported by the grant No. F4/42/2021 of the Faculty of Informatics and Statistics, Prague University of Economics and Business.

REFERENCES

Asgary, A.; S.Z. Valtchev.; M. Chen; M.M. Najafabadi.and J. Wu 2021. "Artificial Intelligence Model of Drive-Through Vaccination Simulation." *International Journal of Environmental Research and Public Health,* Vol. 18, No. 1, 268. https://doi.org/10.3390/ijerph18010268

Brailsford, S.C.; M.W. Carter and S.H. Jacobson 2017. "Five Decades of Healthcare Simulation". In Proceedings of the 2017 Winter Simulation Conference, IEEE Press, Piscataway, N.J., 365-384.

Concannon, K. et al. 2007. *Simulation Modeling with SIMUL8*. Visual Thinking International, Canada.

Greasley, A. 2003. *Simulation modelling for business*. Innovative Business Textbooks, Ashgate, London.

Fousek, J., M. Kuncová and J. Fábry 2017. "Discrete Event Simulation – Production Model in SIMUL8." In *Proceedings of the 31st European Conference on Modelling and Simulation ECMS 2017* (Budapest, May). Dudweiler: Digitaldruck Pirrot, 229–234.

FNKV.cz 2021. FNKV-Aktuality [online], available https://www.fnkv.cz/zprava-odberova-mista-covid-19-ve-fnkv [cit. 2021-01-30]

Hamrock, E; K. Paige; J. Parks; J. Scheulen and S. Levin 2013. "Discrete Event Simulation for Healthcare Organizations: A Tool for Decision Making." *Journal of Healthcare Management*, Vol. 58, No.2, 110-124.

Katsaliaki, K. and N. Mustafee 2011. "Applications of simulation within the healthcare context." *Journal of the Operational Research Society*. Vol. 62, 1431—1451.

Pisaniello, A.; W.B. da Silva; L. Chwif and W.I. Pereira 2018. "Discrete Event Simulation of Appointments Handling at a Children's Hospital Call Center: Lessons Learned from V&V Process." In: *Proceedings of the 2018 Winter Simulation*

Conference, IEEE Press, Piscataway, N.J., 3861–3872.

Shalliker, J. and C. Ricketts. 2002. *An Introduction to SIMUL8, Release nine*. School of Mathematics and Statistics, University of Plymouth.

Simul8.com – SIMUL8 software. [online], [cit. 2020-02-20]. Available: https://www.simul8.com/

van Buuren, M.R., G.J. Kommer, R. van der Mei, and S. Bhulai. 2015. A simulation model for emergency medical services call centers. In *Proceedings of the 2015 Winter Simulation Conference*, IEEE Press, Piscataway, N.J., 844-855.

Viana, J.; S.C. Brailsford; V. Harindra and P.R. Harper 2014. "Combining discrete-event simulation and system dynamics in a healthcare setting: A composite model for Chlamydia infection." European Journal of Operational Research, Vol. 237, No. 1, 196-206. https://doi.org/10.1016/j.ejor.2014.02.052

AUTHOR BIOGRAPHIES

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 (in 2020 renamed as Prague University of Economics and Business), 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 five books and the author of many scientific papers and contributions at conferences. She is interested in the usage of the operational research, simulation models and methods of multi-criteria decision-making in reality. Her email address is: martina.kuncova@vse.cz

KATEŘINA SVITKOVÁ was born in Pilsen, Czech Republic. She studies at Prague University of Economics and Business, study programme Quantitative Methods in Economics, study field Econometrics and Operational Research. She has Bachalor´s degree Econometrics and Operational Research. Her email address is svitule10@seznam.cz

ALENA VACKOVÁ was born in Prague, Czech Republic. She studied at Prague University of Economics and Business, majoring in Econometrics and Operation Research. Her secondary field of study is financial engineering. She also worked with Jan Evangelista Purkyně University in Ústí nad Labem on various projects regarding environmental protection as a statistician/econometrician. Her email address is vackovalena@gmail.com

MILENA VAŇKOVÁ was born in Jaroměř, Czech Republic. She is studying at Prague University of Economics and Business, study programme Quantitative Methods in Economics, study field Econometrics and Operational Research. She has Bachelor's degree Mathematical Methods in Economics. Her email address is vankova.mila@email.cz

Open and Collaborative Models and Simulation Methods

Pedestrian Simulation in SUMO Through Externally Modelled Agents

Daniel Garrido, João Jacob, Daniel Castro Silva, Rosaldo J. F. Rossetti
Artificial Intelligence and Computer Science Lab
Department of Informatics Engineering
Faculty of Engineering of the University of Porto
Rua Dr. Roberto Frias, S/N, 4200-465 Porto, PORTUGAL
Email: {dlgg, joajac, dcs, rossetti}@fe.up.pt

KEYWORDS

SUMO; Pedestrian Safety; Pedestrian Simulation; Unity3D; Virtual Reality; Distributed Simulation; Social Forces Model

ABSTRACT

Pedestrian simulation is often forgotten or implemented poorly in most high-profile traffic simulators. This is the case of SUMO, where pedestrian models are very simple and not based in real human behaviour, making it impossible to study pedestrian safety with it. With this in mind, the ability to externally control pedestrians in SUMO was explored. Using Unity3D to create an external 3D representation of a running SUMO simulation, we were able to create and control pedestrians through the TraCI API. This also opened the possibility to use virtual reality immersed subjects to participate in the simulation, opening the door to study real pedestrian behaviour to create more elaborate models. It also allowed us to completely offload the pedestrian simulation from SUMO to Unity3D, which was tested with the external implementation of the social forces model, without losing SUMO's interactions between pedestrians and motorized vehicles.

INTRODUCTION

Pedestrians are the smallest, lightest, slowest and least protected road users. Also being the ones with the most movement freedom makes them the most vulnerable (Yannis et al. 2011). Despite pedestrian mortality being on decline, 5320 fatalities were registered in 2016 in the European Union alone, making up 21% of all road fatalities (European Commission 2018).

Traffic simulators have been used to study the flow of traffic at least since 1955 (Pursula 1999). Through the years they have evolved to use better simulation paradigms, better models and eventually cover whole traffic networks instead of specific locations. A direct application of these simulators is to study road safety conditions (Young et al. 2014).

Although pedestrians are the most vulnerable road users, their inclusion in traffic simulators is not always guaranteed. Even in cases where they are represented, they are not always modelled with the care and finesse of the motorized cohabitants. This is the case of the traffic simulator SUMO, an open-source microscopic traffic simulation package first released in 2002 (Behrisch et al. 2011). It is one of the most popular traffic simulator for research, second only to VISSIM in terms of number of published papers that make use of it (Mubasher and Jaffry 2015). Despite its popularity, it was only in 2014 that SUMO included pedestrian modelling in its package, 12 years after its initial release (Krajzewicz et al. 2014; Erdmann and Krajzewicz 2015). This model, called stripping model, was developed to fit SUMO developers requirements, which focused on enabling vehicle-pedestrian interactions and how the flow of pedestrians affects traffic flow (Erdmann and Krajzewicz 2015).

In the current model, the sidewalks and crosswalks are split in several lanes, akin to lanes on a multi-lane road. Pedestrians can move to adjacent lanes to prevent being stuck behind a slower pedestrian or to avoid a collision with another pedestrian head on (Krajzewicz et al. 2014; Erdmann and Krajzewicz 2015). This model, while efficient, is not the best at representing pedestrian behavior. In the real world, pedestrians don't move in lanes like cars do, don't always walk in their designated areas and might attempt to cross the road in places other than the crosswalk, or when the crosswalk signal indicates not to cross. These oversights limit the impact of pedestrian safety research made in SUMO.

Currently, the simulation information generated by SUMO can be used by an external application through the use of TraCI (Traffic Control Interface). This API (Application Programming Interface) also allows real-time manipulation of simulation states and variables, enabling the control of vehicles through the interface. Currently, TraCI supports the free movement of vehicles, while for pedestrians it is not specified in SUMO's documentation whether this is possible or not. This means that it currently might not be possible to implement new pedestrian models, without direct manipulation of the source code.

The main goal of this work is to improve SUMO's ability to integrate improved pedestrian models, without the need of tampering with its source code. To achieve this, a communication interface between SUMO

This work was financially supported by Base Funding - UIDB/00027/2020 of the Artificial Intelligence and Computer Science Laboratory – LIACC - funded by national funds through the FCT/MCTES (PIDDAC).

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

and Unity3D was developed to create a 1:1 representation of the simulated world and its inhabitant in Unity3D, where models can be developed and applied to pedestrians. Another goal is to allow people to control a pedestrian in the simulation with the objective of studying their behaviour to further improve the used models. The contributions of this work are:

- Facilitate creating pedestrian models in SUMO.
- Simplify the integration of these external pedestrian models in SUMO.
- Allow real people to control a pedestrian in SUMO to study pedestrian safety in a danger-free environment.

The remainder of this document is structured as follows. In Section *Traffic Simulators*, a review of popular traffic and pedestrian simulators is made, with special attention to SUMO. This is followed by Section *Related Work* which presents similar research. Section *Methodology* documents the approach of the solution, including its architecture along with the tools and models used. The details for the implementation following the proposed solution can be found in Section *Implementation - Platform and Systems* and *Implementation - Pedestrian Model*. Section *Results and Analysis* presents the results of the experiments and their discussion. Finally, Section *Conclusion and Future Work* ends the article with a conclusion of the developed work and ideas for future improvements.

TRAFFIC SIMULATORS

Currently, several traffic simulators with different features are readily available, some being available commercially, and others being free to use and open source. Some surveys about the state of the art of traffic simulators were published in the past decade, providing useful information on the capabilities of different simulators and in some cases, performance comparisons.

In 2009, a review of traffic simulation software compared six of the top platforms of the time (Kotusevski and Hawick 2009). The authors created a list of criteria that would be used as the comparison basis, containing: Licensing model; Platform portability; Documentation/UI; Creation of traffic networks; Outputs; Large network simulation; and computational performance.

More recently, an update version of the previous study was published in 2017(Ejercito et al. 2017). It followed the same comparison criteria, but with a reduced number of platforms, of which two were present in the 2009 review, and the other two were new.

A systematic literature review on traffic flow simulators was also conducted and published in 2015 (Mubasher and Jaffry 2015). A list of 15 traffic simulators used in all types of research is presented, of which 5 were isolated for being the ones more frequently used in the studied literature.

Table I summarizes the investigated traffic simulators in the mentioned surveys. It becomes evident that there is a lot of overlap and that some of the platforms used 10 years ago are still relevant today, one of which is SUMO, the platform that is the target of this project.

TABLE I: Summary of traffic simulators discussed in each literature review: Kotusevski and Hawick 2009 (1); Mubasher and Jaffry 2015 (2); and Ejercito et al. 2017 (3). Column "Peds?" indicates the presence of pedestrian simulation, and "OSS" whether or not it is open-source software.

Simulator	Peds?	OSS?	(1)	(2)	(3)
SUMO	✓	✓	✓	✓	✓
PARAMICS	✓	✗	✓	✓	✗
Treiber's Microsimulator	✗	✓	✓	✗	✗
AIMSUN	✓	✗	✓	✓	✓
Trafficware SimTraff	✓	✗	✓	✗	✗
CORSIM	✓	✗	✓	✓	✗
VISSIM	✓	✗	✗	✓	✓
MatSim	✓	✓	✗	✗	✓

SUMO

SUMO is an open-source, free to use, microscopic traffic simulation package with several articles published documenting its development and progress. In (Behrisch et al. 2011), a general overview of the package is found. Despite being an old publication, it still reflects the current architecture and basic functionality of current SUMO versions.

More recently, pedestrian modeling has been implemented in SUMO, a feature that it had been lacking when compared to some of its alternatives like VISSIM and AIMSUN. Two papers report the development of this feature, an earlier one from 2014 overviewing the implementation of pedestrians and bicycle traffic (Krajzewicz et al. 2014), and a more detailed one published in 2015 about the process of modelling pedestrians in SUMO (Erdmann and Krajzewicz 2015).

Currently SUMO offers two simulation options for pedestrians, the "nonInteracting" and the "striping" models. As the name suggests, in the "nonInteracting" model the pedestrians walk at constant speed, don't interact with each other or with vehicles, and teleport across intersections. It is recommended to use this model when pedestrian dynamics are not essential.

The striping model is a big step up from the previous model, as it enables pedestrian-pedestrian and pedestrian-vehicle interactions. For this model to work properly, the layout of road infrastructure used by pedestrians must be modelled. In this implementation, 3 distinct zones exclusive to pedestrians exist, visible in the color-coded Fig. 1: sidewalk (grey), crossings (zebra) and walking areas (blue). When pedestrians move in the sidewalk or crossing, they follow a striping formation, akin to the way vehicles follow lanes on the road. These areas are split in different lanes in which the pedestrians can move along. If they find an obstacle or a slower pedestrian in their lane they can move to another adjacent one and proceed. In walking areas, which serve as a connection patch between sidewalks and crossings, pedestrians follow a predetermined trajectory calculated at the beginning of the simulation.

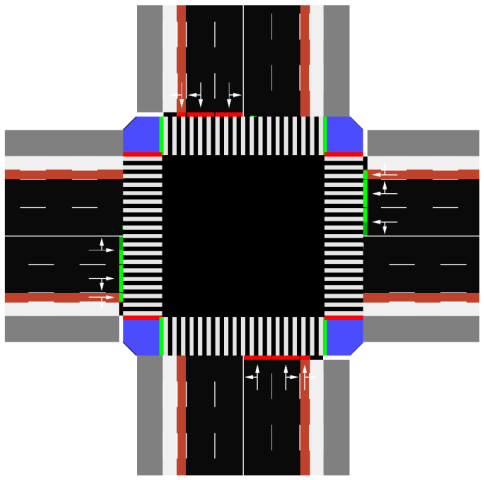


Fig. 1: Pedestrian network model in SUMO (Krajzewicz et al. 2014)

RELATED WORK

As mentioned before, the main objective of this work is to enable external pedestrian models to be connected to SUMO via the TraCi API and incorporate this system in Unity3D to be able to create a 3D version of the network used in SUMO. With this in mind, a literature search was made using the terms 'SUMO', 'TraCi', 'Pedestrian' and 'Unity' in different combinations.

The earliest publication found mentioning a connection between SUMO and Unity3D dates back to 2014, to create a platform for faster, safer and cheaper testing of ADAS (Advanced Driving Assistance Systems) using simulated environments (J. S. V. Gonçalves et al. 2014). This implementation used SUMO as a central server that provides information to remaining modules, including the IC-DEEP simulator (J. Gonçalves et al. 2012). This simulator is based in Unity3D and was developed to test safety aspects of IVIS (In-Vehicle Information Systems).

The work mentioned above then served as the basis to a project funded by the Austrian Ministry for Transport, Innovation and Technology (BMVIT) that focuses on researching vehicle to vehicle and vehicle to pedestrian communications to study systems related to pedestrian safety. To create a platform for this type of communications, SUMO was used in conjunction with Unity3D to create a user interactable microscopic driving simulation. This was first used to create a Traffic Light Assistant system for drivers (Olaverri-Monreal et al. 2018). This work was later expanded to receive data from the pedestrian locations to warn the driver of their position when a danger situation was detected (Artal-Villa and Olaverri-Monreal 2019). This research was later validated by studying the behaviour of 20 subjects driving in the simulated environment with and without the help from the system (Artal-Villa, Hussein, et al. 2019). While some of the components of this research are similar to the work of this paper, it is still lacking in the main goal to introduce external models to pedestrians in SUMO.

For the project mentioned above, the developers linked Unity3D to SUMO through TraCI directly. It was mentioned that usually TraaS (TraCI as a Service) is used to allow multiple types of clients to connect to SUMO, but using it as a web service introduces unacceptable delay for a driver-centric simulation. Ultimately, they implemented the TraaS library locally to avoid that side effect (Biurrun-Quel et al. 2017).

More recently, a new 3D traffic simulator is being developed in Unity3D which, like in this project, receives the vehicle data from an external source. The 3D simulation was then used to collect data through cameras, mimicking the way real autonomous vehicles collect data from its environments, significantly reducing the time and cost of its acquisition (Jin et al. 2018).

In another project, SUMO was used as the microscopic traffic to an autonomous vehicle simulator. The architecture was set to make SUMO handle general network vehicles, while USARSim was used to simulate the autonomous vehicles, using Unreal Engine to create a 3D scene of the simulation (Pereira and Rossetti 2012).

When it comes to the use of Virtual Reality (VR) to aid the creation of pedestrian models some concerns over the realism of the data collect in the virtual world are often raised. This is mainly due to the chance that test subjects may not act the same way while immersed as they do in the real world. A study into this concern have clarified that only minor differences can be found between the two scenarios (Bhagavathula et al. 2018).

This VR approach to pedestrian modelling has been used to better understand the mechanisms of microscopic pedestrian behaviour to develop a disutility minimisation model (Iryo et al. 2013). VR environments have also been used to collect data to create a model for pedestrian behaviour prediction (Costa et al. 2019).

METHODOLOGY

This section introduces the general details of the development of this project, starting with a detailed solution proposal, followed by a description of the overall system model and architecture.

Proposed Solution

As mentioned before, implementing new pedestrian models or adding new features to the existing ones in SUMO can be a cumbersome and difficult process, as it entails adding or modifying the original code base, written in C++. With this in mind, to achieve the first goal of simplifying this process, the modelling and application of the new models would need to be done externally. This would also allow different pedestrian models to coexist in SUMO, something not currently possible and necessary to fulfill the second goal of having a subject control one pedestrian.

Fortunately, SUMO can easily communicate with external application in several programming languages like C++, Python, C#, JAVA and Matlab (DLR 2020). Exploring the different interfaces for different programming languages reveals that the ones for Python and

C++ are the most developed and frequently updated. Currently, these are also the ones that include the *Person.moveToXY* method, which allows the interface to control the position of pedestrians in the network.

To facilitate the integration of new models and real people in the simulation, an interactive 3D representation of the simulation is needed. SUMO works in a 2D network, but it is necessary to transfer it to a 3D network for test subjects to feel immersed and visualize the world as they do in real life. Game engines are a popular choice to create dynamic 3D environments and, as mentioned before, Unity3D has been used in conjunction with SUMO in the past. When also considering its good integration with virtual reality and simple scripting language it became the choice for this project. Being a game engine, Unity3D also includes classes and features for path finding, which can simplify the pedestrian model creation.

In the end, the proposed solution is to connect SUMO and Unity3D to create a traffic and pedestrian simulation using the best of both software packages: the depth of traffic simulation in SUMO and the ease of use and integration of new models from Unity3D. With both working in sync, SUMO's traffic simulation influences the pedestrian simulation in Unity3D and vice-versa, creating a more complete and modular simulation.

System Architecture

The created system for the proposed solution is composed of 3 main modules: SUMO for the microscopic traffic simulation; Unity3D for the immersive visual representation and new pedestrian models; and a Python script responsible for connecting the other 2 modules. Figure 2 shows how these modules interact with each other.

The SUMO simulation module is a regular SUMO simulation with vehicles and pedestrians that follow the internal models for these categories: the default traffic model and the striping model respectively. The distinguishing factor is that it receives through TraCI the position of the pedestrians modeled from the outside in the Unity3D module. Despite being modeled externally, they still interact with SUMO's internal inhabitants, influencing their choices in movement.

On the other end, the Unity3D module functions in reverse. It receives the position from the SUMO modeled inhabitants, and uses that information to determine the best action for the modelled pedestrians, or to influence the decision of a test subject immersed in the simulation. This module is also responsible for recreating the simulation from its 2D nature from SUMO, to a more immersive 3D environment.

Connecting these two modules is a Python script, responsible for several tasks. First, launching the SUMO simulation and establishing the TCP (Transmission Control Protocol) connection to it through TraCI. Then it opens the connection to the Unity3D visualization through a ZeroMQ TCP connection, and finally, it connects both these communication protocols.

The transmitted data varies with the intended usage of the platform. For the case demonstrated in Fig. 2, where a single pedestrian is being externally modelled or controlled by a human, all other pedestrians' and vehicles' data has to be sent from SUMO to Unity3D, with the latter only having to send this externally modelled pedestrian to SUMO. On the other hand, if all pedestrians are being simulated in Unity3D, this sends the positional data of all of them to SUMO, while SUMO only needs to transmit the positions of the vehicles. This last module, sitting between the two main modules is called the middleware.

IMPLEMENTATION - PLATFORM AND SYSTEMS

This section describes the implementation process of the discussed proposed solution. It focuses on how the road network was created and modelled, the communication protocol connecting SUMO and Unity3D, followed by how the pedestrians are controlled remotely, and finishing with the VR component of the simulation.

Network Modelling

The implementation process started with the creation of a simple SUMO scenario to test the system dur-

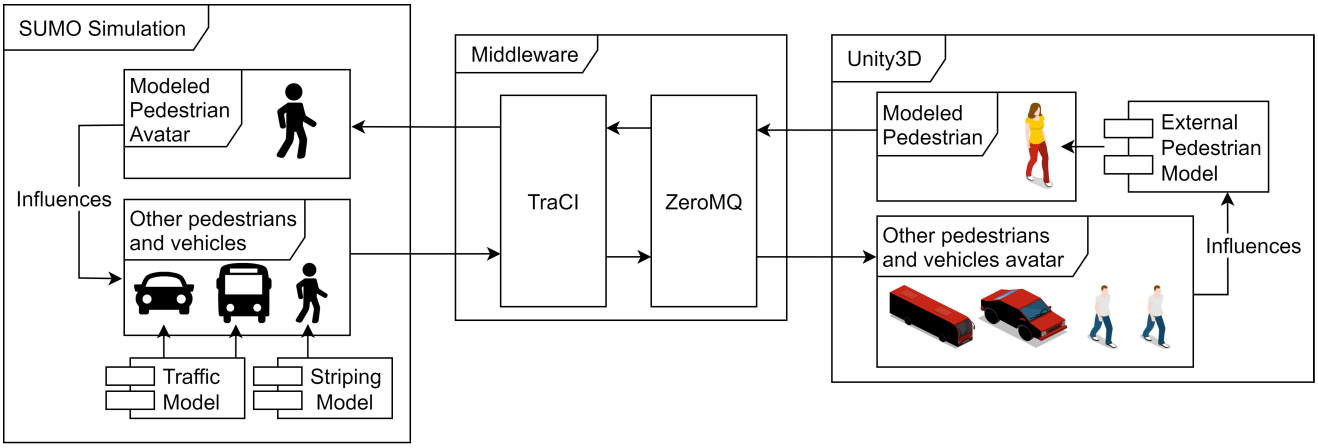


Fig. 2: System architecture for the platform, composed of three main modules: SUMO, the middleware and Unity3D.

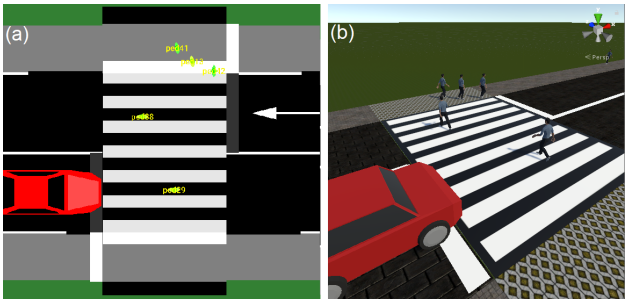


Fig. 3: SUMO simulation (a) and Unity3D reconstruction (b).

ing development. Since the main focus of the system is pedestrian modelling, priority was given to spaces where they usually can be found. As such, a single lane roadway was created with walking areas on both sides and a crosswalk in the middle using SUMO's NETEDIT[1]. To populate the simulation, routes for the cars and pedestrians were created, which at runtime create new inhabitants that move to their destination following said routes.

The created network needed to be replicated in the Unity3D environment, where a scene was created by hand that matched the original scene following the data found in the *net.xml* file created by NETEDIT, which details the nodes and the edges of the network. This process could be automated, but that falls outside the scope of this work. The scale from the original SUMO network was preserved on the Unity3D side to simplify the process of transmitting coordinate and distance data between them.

Communication Protocol

To establish the communication between the two main modules, SUMO's communication protocol TraCI was paired with ZeroMQ to enable fast and reliable message transmission between the two. This implementation of Python-Unity3D C# was based on the work by Chanchana Sornsoontorn[2]. As mentioned before, this is being done in the middleware, since TraCI for Python has more functionalities implemented than its C# counterpart. The communication loop follows a client server model, with the middleware acting as the server and Unity3D as the client. When it starts, it waits for a message from the client. Upon receiving a message, if it includes the position of the modeled pedestrians, it uses TraCI to update their positions in SUMO. Then, it sends the position of all pedestrians and vehicles in the simulation gathered through TraCI and transmits it back to the client.

Upon receiving this message, the client updates the position and angle of the agents in the 3D world and then sends the server the position of the agents it is controlling, completing the cycle. The SUMO simulation step time is kept by the middleware, and can be chosen between as fast as possible for simulation scenarios without humans in the loop or to keep real-time when a human subject is immersed in the simulation. This protocol is depicted in Fig. 4.

To prevent blocking of the main thread in the Unity3D simulation, which would negatively impact VR performance, the data from the server is received in a parallel thread. Due to the limitations of Unity3D, data from the simulation can only be accessed from the main thread. To bypass this, after receiving a message, it is placed in a queue of functions that are executed by the main thread when it can. This implementation was based on the work of Damian Mehers [3].

At this point, it is already possible to recreate the SUMO simulation in Unity3D, as pictured in Fig. 3.

Controlling Pedestrians Through TraCI

With the communication protocol established, it is necessary to make the externally modeled pedestrians move in SUMO as well.

The *Person.moveToXY* function available in TraCI would suggest a simple solution to this, but at the time of implementation, while it can in fact move a pedestrian to any spot on the map, other pedestrians and vehicles will not become aware of its presence. This is due to the method not updating the pedestrian position in terms of edge, but only in terms of coordinates, preventing vehicles and pedestrians outside that edge to be aware of its presence.

To bypass this limitation, the message the server receives also contains the edge the pedestrians are currently at. This is calculated in Unity3D by casting a ray directly down from its position and listening to which object it intersects with. With this information, the server can detect if the pedestrian has changed edge, in which case it removes the pedestrian from the simulation and creates a copy on the current edge. While this

[1]More details at: https://sumo.dlr.de/docs/NETEDIT.html
[2]Original code and information at:https://github.com/off99555/Unity3D-Python-Communication
[3]Original code and information at:https://github.com/PimDeWitte/UnityMainThreadDispatcher

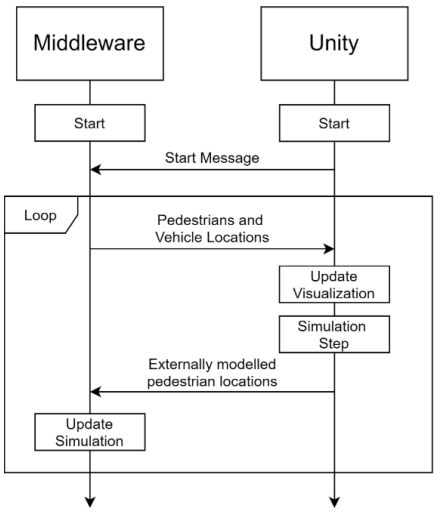


Fig. 4: Sequence diagram of communication loop between Unity3D and the middleware.

solution can sometimes be prone to crashes, it was the only way found to circumvent SUMO's limitations.

Another adjustment necessary for free pedestrian movement in SUMO is to make roadways available to pedestrians. Otherwise, when external pedestrians walk on the road surface outside the crosswalk the vehicles would not be aware of them and run them over.

Controlling Pedestrian in VR

The developments reported above enable achieving the goal of allowing externally controlled pedestrians in SUMO in a simple manner. In this section, we tackle the second goal of allowing real people to control one of these pedestrians, in order to study potential dangerous behaviour in a safe way.

One of the reasons that led to choosing Unity3D was due to the easy integration of virtual reality. Several Unity3D software packages exist to simplify the integration of virtual reality, with VRTK[4] being one of the most popular. It was used to integrate SteamVR to the simulation, making it compatible with most mainstream VR systems in the market.

VRTK also includes several VR-specific locomotion methods. The typical teleport technique was not implemented as it creates unnatural motion. To maintain realism and immersion, joystick activated smooth locomotion was used, with it also being possible to rotate the camera with buttons as a fallback for when a 360º VR setup is not available. When VR is not available, it is also possible to use VRTK's VRSimulator to control the pedestrian with keyboard and mouse.

IMPLEMENTATION - PEDESTRIAN MODEL

After implementing the general structure and functionality of this distributed simulation platform, it is still necessary to demonstrate it could indeed be used to offload pedestrian simulation from SUMO to Unity3D, to ease the implementation of new models and allow more complex models to be used. This section details the implementation of the simple and established Social Force Model for pedestrian simulation and the required adjustments to the platform to make it possible.

Multi-Agent Pedestrian Models

Many pedestrians models have been developed over the years. In the case of microscopic pedestrian simulation, it is usually achieved through the use of multi-agent systems. Two main approaches can be found in the literature, those being force-based methods in the beginning, and the more recent velocity-based methods, following different techniques such as time-to-collision or computer vision (Karamouzas et al. 2017).

While velocity-based methods more accurately simulate life-like pedestrian behaviour, they are more complex to implement and require more computational resources to run in real-time. Hence, the simpler force-based models were chosen to be implemented.

[4]More details at: https://vrtoolkit.readme.io/

Originally, the social force model was developed by Craig Reynolds in 1987 to simulate the behaviour of flocks of birds. Each bird in the flock was simulated as a point in space where several forces (separation, alignment, and cohesion) are applied, creating a velocity vector which is then applied to the bird to update the position (Reynolds 1987).

Years later, in 1995, Helbing and Molnár took this approach to the simulation of crowds of pedestrians. For this, they based their model in 3 distinct social forces. The first, desire, emulates the want of the pedestrian to reach his destination as fast as possible, being translated as a force with the orientation and direction of the goal position. The second, repulsion, models the want of the pedestrian to keep a certain distance from other pedestrians, being calculated as the sum of the repulsion vectors in relation to other pedestrians nearby. The third one, attraction, is used to simulate the cases in which the pedestrian might want to get closer to other things, such as friendly pedestrians, a shop window or a street performer for example (Helbing and Molnár 1995).

Social Force Model Implementation in Unity3D

Before implementing the social forces model, the base for its integration in the platform was prepared.

First, the pedestrian spawner was transferred from SUMO to Unity3D. This spawner takes as input the initial number of pedestrians to spawn and the number of pedestrians to spawn per minute passed, to keep the simulation filled with pedestrians, as they are destroyed once they reach their destination.

When pedestrian agents are created, they are given a random initial location and destination on top of a

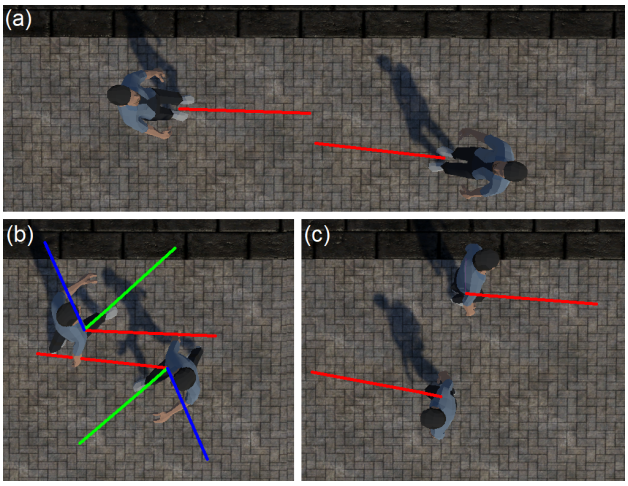


Fig. 5: Social Forces Model in action to avoid pedestrian collision. In image (a) the pedestrians are only influenced by the desire force (red). In image (b), when the pedestrians get close, the repulsive forces (blue) alter the pedestrian behaviour, resulting in the movement direction vector (green) to deviate from the desire force vector. In image (c), after avoiding collision, they return to only being commanded by the desire vector.

sidewalk. At the same time, the pedestrian calculates the best course to reach its destination, in the form of several waypoints. To make SUMO accept the externally simulated pedestrians, the same workarounds detailed in the previous section were used. With more simulated pedestrians, the SUMO server has increased propensity to crash when a pedestrian changes edge. This is a bug that would have to be fixed in SUMO itself, in order for this platform to work reliably.

While the original implementation of the Social Forces Model was not released to the public, there is a NetLogo implementation openly available through GitHub, credited to Antoine Tordeux [5]. This implementation only takes into account the desire and repulsion forces, as the attraction force can be optional, depending on the simulation scenario. This implementation was adapted to C# and Unity3D's game engine. Figure 5 demonstrates how the modelled forces influence the pedestrian movement behaviour.

The first test showed that pedestrians tended to exit the sidewalks to avoid collisions with other pedestrians, ending up either outside the network (in the green area) or on the road, where cars circulate. To combat this issue, metaphorical walls were added to the edges of the surfaces where pedestrians can walk (sidewalks, walking areas, and crosswalks). Every simulation step, the pedestrian agent calculates the closest point on a "wall" to him and the distance to that point. When it gets close, a repulsive force with the opposite direction of the wall is added to the rest. In extreme cases where even with the repulsive force the pedestrian walked into unexpected terrain, it is forced to stay inside the lines.

SIMULATION TESTING

The developed code for this project, along with the necessary instructions, is accessible on Github[6]. All the tests were performed on Windows 10, in a machine with an Intel i7-8750h processor, an RTX 2060 graphics card and 16GB of RAM. For immersion, an Oculus Rift was used with Oculus Touch Controllers through the SteamVR SDK (Software Development Kit).

[5]More details at: https://github.com/chraibi/SocialForceModel
[6]Codebase at: https://github.com/dalugoga/sumo-unity-distributed-pedestrian-simulator

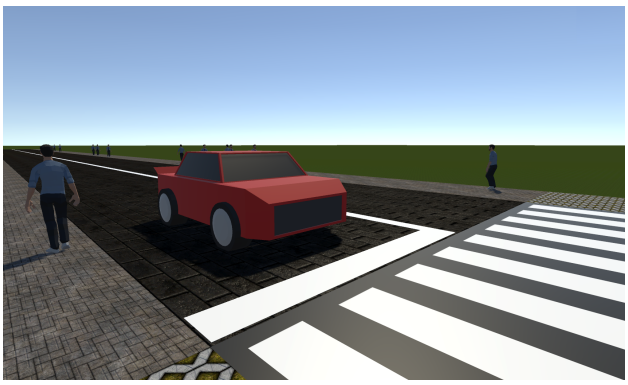


Fig. 6: SUMO vehicle waiting for immersed user to cross.

The first step to start the simulation is to execute the middleware Python Script. The SUMO simulation will remain frozen until the Unity3D scene is played, either from the Unity3D editor or the standalone executable, making the two simulation sides synchronize and run.

Two test scenarios were created and executed, one for each implemented use scenario. The first one focuses on the VR control of a single pedestrian, while the second creates a test scenario for the social forces model.

Virtual Reality Test Scenario

A simple simulation of SUMO controlled pedestrians was created, with each pedestrian having a random start and end point. At the same time, vehicles are spawned randomly. In some cases, the pedestrians have to cross the road through a crosswalk, making the road traffic stop.

A test subject was then immersed in Unity3D with the objective of moving around and behave like a real pedestrian, trying to cross the road. To make the subject movement be represented in SUMO, his camera movement and rotation were tied to a SUMO pedestrian, now being controlled by TraCI with the inputs from Unity3D being fed through ZeroMQ.

The inputs from this pedestrians were correctly transmitted to SUMO, making the other pedestrians and vehicles respond to its presence in the simulation. This environment was very immersive for the test user with no noticeable delays or simulation inconsistencies. Figure 6 shows the subject waiting for the red car to see him and stop, before initiating the crossing safely.

Simulating all Pedestrians Externally

To test the feasibility of completely offloading the simulation of pedestrians to Unity3D, a new scenario was created, where Unity3D itself spawns, removes and controls the actions of dozens of pedestrians simultaneously. In this test, 100 pedestrians are initially created, with 5 more added each minute, to keep the pedestrian population high. In this scenario, SUMO is only handling the simulation of cars, also randomly spawned.

This simulation scenario ran without major problems, with pedestrians moving at a regular pace the majority of the time. In simple collision scenarios between 2 or 3 pedestrians, the social forces model does a good job at keeping the flow in a human-like manner. However, when the pedestrian density increases, mostly around the crosswalk area, the age of the model becomes apparent, as some pedestrians are forced to wait for a chance to move forward, or even backtrack to find a better position to move forward.

Despite these problems, the most important requirements were still met. A large pedestrian crowd was successfully externally simulated in Unity3D to be used by SUMO, while the vehicles being modelled by SUMO were still aware of their presence and would stop to let the pedestrians cross the road. Figure 7 shows the two simulation representations synchronized while running the Social Forces Model.

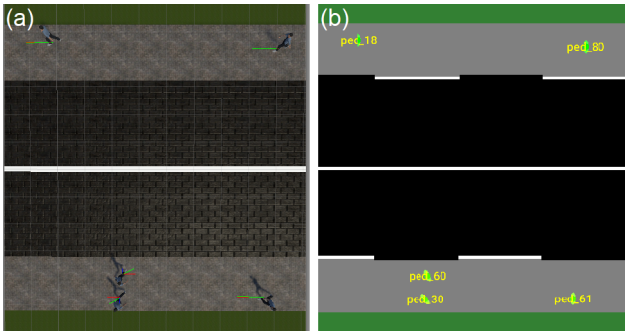


Fig. 7: Unity3D's (a) externally simulated pedestrian agents being replicated in SUMO (b).

CONCLUSIONS AND FUTURE WORK

An extension to the SUMO microscopic traffic simulator was developed that simplifies the process of integrating externally controlled pedestrians, with both objectives being achieved. Using TraCI, SUMO was connected to the Unity3D game engine to create a 3D representation of the simulation and provide information to new pedestrian models. It also allows for real people to be incorporated in the simulation through immersive virtual reality headsets. The end result allows researchers to develop more complex pedestrian models for SUMO and to safely study the behaviour of real people in the role of pedestrians.

For future work, it would be important to implement a modern pedestrian model that simulates pedestrians in a even more realistic way, with it being possible for them to jaywalk, use the crosswalk while it indicates not to cross, etc. By introducing this potential risk behaviours, studies could be performed to help understand them and ways to prevent accidents from happening. This model could be created from data collected from immersed users and verified with other user tests in the platform. It would also be fundamental to fix the problems related with how the external pedestrians are controlled in SUMO to prevent the seemingly random crashes and increase the platform's stability.

REFERENCES

Artal-Villa, L., A. Hussein, and C. Olaverri-Monreal. 2019. "Extension of the 3DCoAutoSim to Simulate Vehicle and Pedestrian Interaction based on SUMO and Unity 3D". In: *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 885–890.

Artal-Villa, L. and C. Olaverri-Monreal. 2019. "Vehicle-Pedestrian Interaction in SUMO and Unity3D". In: *New Knowledge in Information Systems and Technologies*. Ed. by Á. Rocha, H. Adeli, L. P. Reis, and S. Costanzo. Cham: Springer International Publishing, pp. 198–207.

Behrisch, M., L. Bieker, J. Erdmann, and D. Krajzewicz. 2011. "SUMO – Simulation of Urban MObility: An Overview". In: *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*. ThinkMind, pp. 55–60.

Bhagavathula, R., B. Williams, J. Owens, and R. Gibbons. 2018. "The Reality of Virtual Reality: A Comparison of Pedestrian Behavior in Real and Virtual Environments". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62.1, pp. 2056–2060.

Biurrun-Quel, C., L. Serrano-Arriezu, and C. Olaverri-Monreal. 2017. "Microscopic Driver-Centric Simulator: Linking Unity3D and SUMO". In: *Recent Advances in Information Systems and Technologies*. Ed. by Á. Rocha, A. M. Correia, H. Adeli, L. P. Reis, and S. Costanzo. Cham: Springer International Publishing, pp. 851–860.

Costa, J. F., J. Jacob, T. Rúbio, D. Silva, H. L. Cardoso, S. Ferreira, R. Rodrigues, E. Oliveira, and R. J. F. Rossetti. 2019. "Using Virtual Reality Environments to Predict Pedestrian Behaviour". In: *2019 IEEE International Smart Cities Conference (ISC2)*, pp. 508–513.

Ejercito, P. M., K. G. E. Nebrija, R. P. Feria, and L. L. Lara-Figueroa. 2017. "Traffic simulation software review". In: *2017 8th International Conference on Information, Intelligence, Systems Applications (IISA)*, pp. 1–4.

Erdmann, J. and D. Krajzewicz. 2015. "Modelling Pedestrian Dynamics in SUMO". In: *SUMO 2015 – Intermodal Simulation for Intermodal Transport*.

European Commission. 2018. *Traffic Safety Basic Facts on Pedestrians*. Tech. rep. European Commission, Directorate General for Transport.

German Aerospace Center (DLR) et al. 2020. *TraCI - Interfaces by Programming Language*. Accessed: 2021-01-20. URL: https://sumo.dlr.de/docs/TraCI.html#interfaces_by_programming_language.

Gonçalves, J., R. J. F. Rossetti, and C. Olaverri-Monreal. 2012. "IC-DEEP: A serious games based application to assess the ergonomics of in-vehicle information systems". In: *2012 15th International IEEE Conference on Intelligent Transportation Systems*, pp. 1809–1814.

Gonçalves, J. S. V., R. J. F. Rossetti, J. Jacob, J. Gonçalves, C. Olaverri-Monreal, A. Coelho, and R. Rodrigues. 2014. "Testing Advanced Driver Assistance Systems with a serious-game-based human factors analysis suite". In: *2014 IEEE Intelligent Vehicles Symposium Proceedings*, pp. 13–18.

Helbing, D. and P. Molnár. 1995. "Social force model for pedestrian dynamics". In: *Physical Review E* 51.5, pp. 4282–4286.

Iryo, T., M. Asano, S. Odani, and S. Izumi. 2013. "Examining Factors of Walking Disutility for Microscopic Pedestrian Model – A Virtual Reality Approach". In: *Procedia - Social and Behavioral Sciences* 80, pp. 940–959. ISSN: 1877-0428.

Jin, Z., T. Swedish, and R. Raskar. 2018. *3D Traffic Simulation for Autonomous Vehicles in Unity and Python*.

Karamouzas, I., N. Sohre, R. Narain, and S. J. Guy. 2017. "Implicit Crowds: Optimization Integrator for Robust Crowd Simulation". In: *ACM Transactions on Graphics* 36.4.

Kotusevski, G. and K. Hawick. 2009. "A Review of Traffic Simulation Software". In: *Research Letters in the Information and Mathematical Sciences* 13.

Krajzewicz, D., J. Erdmann, J. Härri, and T. Spyropoulos. 2014. "Including Pedestrian and Bicycle Traffic into the Traffic Simulation SUMO". In: *10th ITS European Congress*.

Mubasher, M. M. and S. W. u. Q. Jaffry. 2015. "Systematic literature review of vehicular traffic flow simulators". In: *2015 International Conference on Open Source Software Computing (OSSCOM)*, pp. 1–6.

Olaverri-Monreal, C., J. Errea-Moreno, A. Díaz-Álvarez, C. Biurrun-Quel, L. Serrano-Arriezu, and M. Kuba. 2018. "Connection of the SUMO Microscopic Traffic Simulator and the Unity 3D Game Engine to Evaluate V2X Communication-Based Systems". In: *Sensors* 18.12, p. 4399.

Pereira, J. L. F. and R. J. F. Rossetti. 2012. "An Integrated Architecture for Autonomous Vehicles Simulation". In: *Proceedings of the 27th Annual ACM Symposium on Applied Computing*. SAC '12. Trento, Italy: Association for Computing Machinery, pp. 286–292.

Pursula, M. 1999. "Simulation of Traffic System – An Overview". In: *Journal of Geographic Information and Decision Analysis* 3.1, pp. 1–8.

Reynolds, C. W. 1987. "Flocks, Herds and Schools: A Distributed Behavioral Model". In: *SIGGRAPH Comput. Graph.* 21.4, pp. 25–34.

Yannis, G., E. Papadimitriou, and P. Evgenikos. 2011. "About pedestrian safety in Europe". In: *Advances in Transportation Studies* 24, pp. 5–14.

Young, W., A. Sobhani, M. G. Lenné, and M. Sarvi. 2014. "Simulation of safety: A review of the state of the art in road safety simulation modelling". In: *Accident Analysis & Prevention* 66, pp. 89–103.

MCX — An Open-Source Framework for Digital Twins

*Sajad Shahsavari, Eero Immonen,
and Mohammed Rabah
Computational Engineering and Analysis (COMEA)
Turku University of Applied Sciences
20520 Turku, Finland
Email: *sajad.shahsavari@turkuamk.fi

Mohammad-Hashem Haghbayan, Juha Plosila
Department of Computing
University of Turku (UTU)
20500 Turku, Finland

KEYWORDS

Digital twin, Cyber-physical systems, on-line simulation

ABSTRACT

This article describes *ModelConductor-eXtended* (*MCX*), which is an open-source software architecture for digital twins. The *MCX* framework facilitates co-execution of, and asynchronous data communication between, physical systems and their digital simulation models. *MCX* supports running FMUs (simulation models packaged according to the FMI specification) as well as machine learning models and customized models. We propose extensions to the previously published *ModelConductor* framework for higher performance and better scalability. The extensions include decoupling of the queue and the model computation module, utilization of a standard data transmission protocol and implementation of the facility to run time-consuming simulation models in a time synchronous manner. Additionally, three new validation case studies are presented. A performance evaluation shows that the extensions improve the average response time almost 4 times in three specific experiments.

I INTRODUCTION

I-A Background

The cyber-physical systems of the celebrated fourth industrial revolution — so-called *digital twins* (DT), i.e. accurate numerical simulation models operated alongside their physical counterparts — are projected to constitute the backbone of modern industrial automation (see Colombo, Karnouskos, Kaynak, Shi, and Yin (2017); He, Chen, Dong, Sun, and Shen (2019)). While the promise of numerical simulation and system modeling has traditionally been in saving time and money during *product development*, today, digital twins can extend far into *product operations* through proliferation of ubiquitous wireless communication. There are, however, a number of practical challenges in co-execution of physical and digital assets, which is the concern of the present article.

Many examples of digital twin simulation models interacting in real-time with a physical device have been reported in the literature. For example, in *predictive maintenance*, real-time measurement data from a physical system is processed in a simulation model for predicting potential damage in the physical system, should the prevailing situation or the trend continue (see e.g. de Azevedo, Araújo, and Bouchonneau (2016) for discussion on wind turbine applications). On the other hand, with the advent of 5G technology and cloud, edge, fog and mist computing, digital twins also facilitate *remote control*, whereby only minimal data acquisition and safety circuitry reside onboard the physical device, and model-based control is calculated on the digital twin (see e.g. Lee, Suh, Kwak, and Han (2020) for discussion on remote drone control). In spite of this progress, it seems that practical applications of digital twins are still designed, implemented and operated on a case-by-case basis.

The big promise of digital twins for the future is, ultimately, in them providing *full system autonomy*: That any design and/or adaptation of control mechanism for the physical system would be first optimized and implemented on the digital simulation model and then be transferred verbatim to the physical device. To realize this, a flexible *software framework* for running a physical system and its digital twin simulation model (or a collection thereof) seamlessly alongside each other is required. Perhaps somewhat surprisingly, this is not trivial, and, to the authors' knowledge, there are no off-the-shelf solutions for this purpose. In fact, while there is an abundance of simulation software that facilitate creation of digital twins, and there are many Internet-of-Things (IoT) solutions for data transfer between devices, such solutions that address both aspects at once appear to be few and far between. This is reflected in the case-by-case nature of practical applications mentioned above.

Recently, Aho and Immonen (2020) introduced an open-source software framework, called *ModelConductor* that defines a digital twin design pattern to facilitate on-line asynchronous data interexchange between a physical device and a digital twin simulation model. *ModelConductor*, basically a connector of the physical devices to their digital twins as described in Subsection II-B, is capable of handling multiple asynchronous data streams via a variable-length queue containing objects measured from the physical environment, waiting to be processed by the simulation model. It also supports running different simulation model types through the Functional Mock-up Interface (FMI) (Blochwitz et al., 2011), which standardizes model exchange and co-simulation between different computation environments and is now supported by more than 150 simulation platforms.

While *ModelConductor* provides a proof-of-concept solution addressing the above basic concerns for co-execution of physical systems and digital twin simulation models, it is not fully compatible with many-to-many relationship between data sources and running simulation instances because of direct function calls between the queue structure and the simulation model.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

Additionally, two way communication facility with the purpose of controlling the physical device by its digital twin is not implemented there. Moreover, the structure of the *ModelConductor* framework restricts distribution of computation on several machines and balancing the requests load between multiple instances running the computational models. This could prevent the framework to scale properly with execution of several computational simulation models being fed by big data streams. In this paper, we propose an extension to the *ModelConductor* framework, namely *ModelConductor-eXtend*, or *MCX* for short, that addresses the above concerns. Several use cases are included for validation and illustration.

I-B Contributions and key limitations

In this article, we describe these extensions to the *ModelConductor* framework:

- Increase scalability and performance of the framework under the condition of high load by decoupling the queue which stores the measurements from the actual execution of computational model.
- Follow a more general message passing and data transmission protocol.
- Introduce a solution for executing time-consuming simulation models (with relatively high data income rate).
- Provide new use cases for system validation by real-world applications.

We emphasize that the *MCX* framework is, at present, only tested for transferring data from physical devices to their simulation model replications. Model-based actuation and feedback control are important implementation steps left for future work.

I-C Relation to previous work

The contributions described in Subsection I-B, all are implemented on the *ModelConductor* framework introduced by Aho and Immonen (2020). The present work is its continuation to improve its functionality and performance, and introduce new use cases for validation.

Toward realizing the digital twin, there exist software platforms that are capable of creating virtual representation of a physical object or system to act as their digital replica. Using these simulation modeling tools, such as Matlab/Simulink, ANSYS twin builder, Dymola, STAR CCM+, Unity3D engine, etc., several studies have been carried out to propose implementation of digital twins for a variety of applications. List of different studies can be found in (Lim, Zheng, & Chen, 2019) and (Cimino, Negri, & Fumagalli, 2019). Previously conducted studies, mostly for pre-specific device, process or system, are limited to simulation and modeling, are not connected to physical devices and do not incorporate real time data. On the other hand, open-source and commercial frameworks are developed to provide the necessary interfaces for handling and authorizing real time data streams in order to securely collect and monitor the data from multiple physical devices (such as Eclipse Ditto (Eclipse, 2020) and Microsoft Azure Digital Twins (Microsoft, 2020)), but they do not address co-execution of simulation models.

This work attempts to bridge the gaps between these distinct domains. More specifically, the *MCX* framework, as an open-source software infrastructure enabling seamless data communication and execution of the twin's simulation model, regardless of application and modeling tool is developed to address the above two capabilities both together.

I-D Organization of the article

Next sections of this paper is structured as follow: In section II, we will introduce the *MCX* framework briefly, explain its capabilities and discuss particularly the new development and alterations proposed in this work. Afterward, three implemented digital twin case examples are discussed in section III as validation applications. Finally, we conclude the paper in section IV and discuss the way to continue the work in the future.

II *MCX* FRAMEWORK

II-A Framework overview

MCX is an open-source software program designed as a ready-to-use structure for the basic connections for implementing and running a digital twin packaged in a standardized general form. The framework itself does not include any simulation, measurement or sensory data, but, on the other hand, it includes the placeholder for running the simulation model and asynchronous message handling structure.

The simplest conceptual use case of the framework is illustrated in Figure 1, with one physical device and its corresponding simulation/prediction model (digital twin). The measured data from a physical sensor, typically (but not necessarily) in a constant frequency, is transmitted in a standard format to the queue by the client. Then, the queue stores the measurements until they are fetched by the subscribed simulation/prediction model. Whenever the model is ready to process new data, it will receive the first element in the queue, then the step method of the model will be executed to produce the output. The output (model response) is logged and could be used for monitoring purpose and also sent back to the physical device's actuator. Although the actuator is not implemented in the framework yet, it is used to demonstrate the concept.

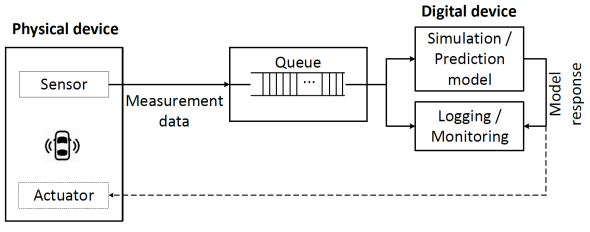


Fig. 1: Basic components of *MCX*

II-B *The starting point for* MCX *development*

The basis for *MCX* is the *ModelConductor* framework introduced by Aho and Immonen (2020). There, the `Experiment` class was used to hold the model, queue and results all together. The sequence diagram of its main data processing loop is reproduced for reference in Figure 2. It shows that, on each iteration of the loop, the code checks whether there is at least one element in the buffer. If so, and also the model is in a `Ready` state (i.e. not preoccupied processing a previous data element), a measurement data point is removed from the buffer and used to make an inference from the associated model by calling `step` method of the `ModelHandler` object. The result — a `ModelResponse` object — is then appended to another list, an attribute of the `Experiment` object.

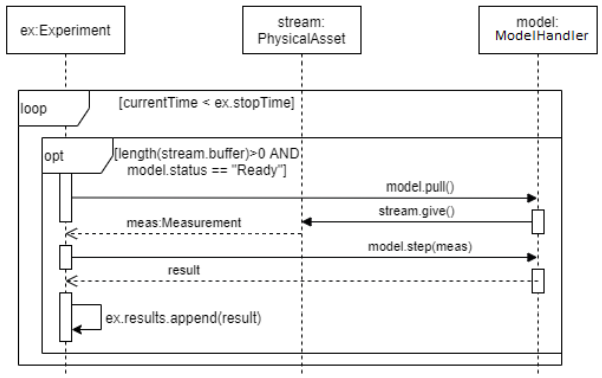


Fig. 2: Sequence diagram in *ModelConductor* (based on Aho and Immonen (2020)).

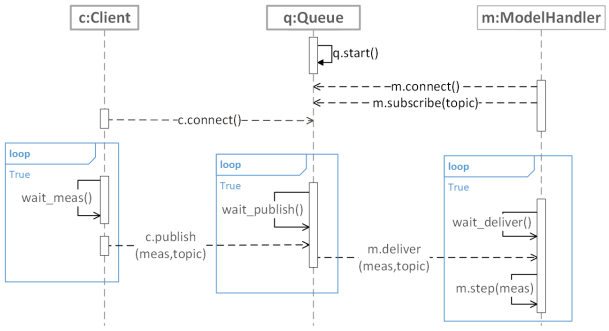


Fig. 3: Sequence diagram in *MCX*

II-C *New developments in* MCX

The sequence diagram of the proposed *MCX* framework is shown in Figure 3. In the remainder of this section, we will compare the two frameworks to each other and point to the shortcomings of the initial version one by one. We will also discuss the proposed modifications and new functionalities in detail.

1) *Decoupling the queue from the model*

The queue data structure in Figure 1 is a FIFO (First In, First Out) data buffer storing measurements which have arrived from the physical device and are waiting to get processed by simulation model. Utilizing this structure is a necessity to accommodate asynchronous data streams. As explained in the (Aho & Immonen, 2020), the queue operations (push & pop) for handling measurement data observations are both executed in one process but in separated threads. As illustrated in Figure 2, the `Experiment` object explicitly waits until a new message is appended to the queue buffer. It means that an instance of the `ModelHandler` class (that is exactly the place where one step of the simulation model is executed) must wait until the arrival of a new measurement data once it has processed current data. The condition for the availability of new data is checking the queue being non-empty (length of `stream` be greater than zero). Hence, it can be observed that the class which holds the simulation model is tightly coupled with the queue structure. In such coupling situation, a change in one module may enforce changes in other modules, affecting code reusability and scalability. Tightly coupled systems are often seen as disadvantage (Beck & Diehl, 2011).

In *MCX*, see Figure 3, the queue structure is decoupled from the `ModelHandler` and each of them are executed in different processes. Then the connectivity between the two modules (Queue and Model) is established via an MQTT (Message Queuing Telemetry Transport) connection (see Section II-C2). This allows for distribution of computational load, as the queue can be processed on a remote computer. This not only increases the scalability and loosens the coupling of the system, but also facilitates many-to-many relationships between multiple data producers and multiple data consumers.

2) *Replacing raw TCP with MQTT*

In *ModelConductor*, measurement data was transmitted using a TCP socket with a pre-defined message format (stringified JSON with fixed header size describing the length of the message). In the *MCX*, we use MQTT instead of TCP. MQTT is a Client Server publish/subscribe messaging transport protocol which has been widely used in data-intensive IoT applications. The motivation for using MQTT include:

- MQTT offers a standard messaging protocol which is supported by IoT community. It also provides integration to the open-source and commercial cloud services (such as Google Cloud, Microsoft Azure and Amazon web services).
- MQTT enables two-way communication between physical device and its digital twin in an effective and scalable way. The structure of MQTT also facilitates many-to-many relationships in data streams.
- MQTT messaging transport is agnostic to the content of the payload. This make the messaging protocol indifferent to the application of the digital twin.
- There are three different qualities of service (QoS) for message delivery status in MQTT protocol. These qualities (including “at most once”, “at least once” and “exactly once”) are practical in digital twin.
- MQTT is considered as an application layer protocol in the well-known Open Systems Inter-

connection model (OSI model)[1] utilizing TCP as the infrastructure for message transportation.

A schematic of the general connectivity facilitated by the proposed MQTT-based communication is shown in Figure 4. The built-in queue in the MQTT broker enables asynchronous data connection and the structure follows the publisher-subscriber paradigm. Each publisher can be seen as an individual sensor, sending the measured data into the system with specific topic which one or more subscribers are listening to its messages. In the context of this paper, subscribers could be seen as computational simulation models. With this generic structure, the framework can fit to the variety of applications while maintaining scalablilty and performance. Additionally, reverse data stream i.e. from model to the physical device, is feasible which facilitates model-based control (not experimented in this work though).

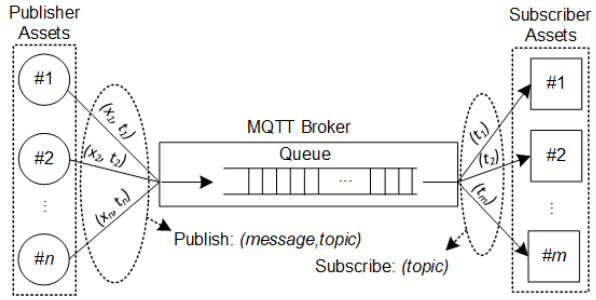


Fig. 4: General overview of the *MCX* framework with multiple publisher assets streaming the sensory data each with specific topic and multiple computational models

3) Memory

ModelConductor was first designed to process one measurement data at a time and produce its response by executing one step of simulation/prediction model. This can be referred to as sample-synchronized memoryless model execution: Each iteration of the model execution is only dependent on the current input data, and each input measurement data has precisely one model response. However, in some practical applications, the model can be computationally expensive, with execution time exceeding the arrival time of new measurements. Thus, following a sample-synchronized memoryless procedure may cause an accumulative delay in responses and also lengthening of the queue over time.

MCX addresses this issue by proposing a time-synchronized memoryful historical modeling procedure by adding a local buffer which stores incoming data while the model is being executed on previous data points. Historical modeling can be further categorized into two types based on the length variability of their local buffer memory: fixed-length and variable-length buffer (see the example in Subsection III-B).

II-D MCX *interfaces to simulation models*

To use the *MCX* framework, the simulation model of the specific application is embedded into the framework acting as the running digital twin. There are three different ways in *MCX* for this purpose:

1) Functional Mock-up Unit (FMU) models: If the simulation model is developed in one of the 150+ tools that support FMU export[2] (such as ANSYS, CATIA, GT-SUITE, MapleSim, MATLAB® Simulink®, SimulationX and etc.), then the exported model could be easily embedded in the framework. The facilities for importing the model, setting input/output variables and running one step of the model are implemented in the `FMUModelHandler` class (using FMPy library[3]).
2) Scikit-learn model: If the digital twin is based on the predictive machine learning model developed in Scikit-learn library (Pedregosa et al., 2011), it could be integrated into *MCX* using `SklearnModelHandler` class.
3) Custom class: If neither of the above options apply, then the user can implement customized behavior in a Python class and embed it as a running simulation model. The class is inherited from `ModelHandler` with specific methods to load, step and shutdown the model.

III CASE STUDIES

III-A Drone simulation and control

In this validation example, an open-source MATLAB/Simulink based dynamic modeling and simulation of a drone (quadcopter) (see Hartman, Landis, Mehrer, Moreno, and Kim (2014)) has been exported into FMU container and then the FMU package has been embedded into the *MCX* to act as the flying drone's digital replica. The simulation is an attitude-command-only model which means the controller only tries to track attitude commands (orientation in terms of angles: ϕ for roll, θ for pitch and ψ for yaw) and altitude command (z) using a PID controller. The idea of drone's digital twin, here, is that a copy of the control command (ϕ, θ, ψ, z), initiated from the drone remote controller, is sent to the drone's digital simulation running on *MCX*. The idea is illustrated in Figure 5, however, dashed arrows were not considered in this example and a client has been placed in the $\times$ sign to mimic the behaviour of the remote controller. With this set up, we can follow the internal dynamic variables of the drone using its simulation while the actual physical device is running.

In this 50 second simulation, the drone is initially stationary at $z = 3.048$m. Then at $t = 25s$, a simple roll command is performed (with ϕ changing as a step signal while keeping the other command variables θ, ψ and z constant). The input command and observed positional values y and z of the drone are illustrated in Figure 6 (x position remains close to zero and is omitted). As shown, the controller tries to keep the altitude (z) constant and due to the roll command drone starts to move forward along the y direction.

[1] https://osi-model.com/

[2] https://fmi-standard.org/tools/

[3] https://github.com/CATIA-Systems/FMPy

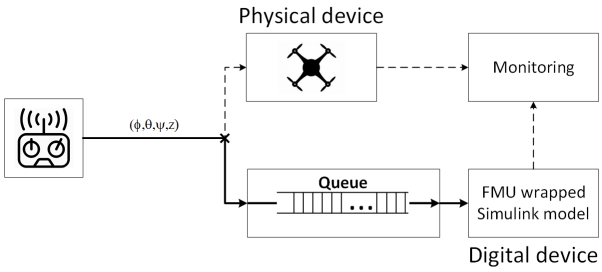


Fig. 5: Overview of the components in the drone experiment.

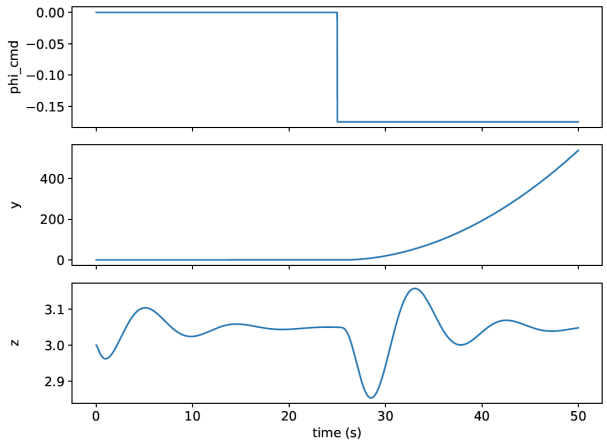


Fig. 6: Input command and positional values of the drone simulation over 50 seconds of the experiment

III-B Historical time series prediction

As a proof-of-concept example of the memoryful historical procedure (explained in II-C3), we built an artificially slow (computationally expensive) model by explicitly "sleeping" during the model response computation. In this demonstrative experiment, we tried to predict the next value of a time series by fitting a linear regression model to the trailing window of the series (last N data points).

We send a numerical value as a synthesized measurement data (with noisy 3rd degree polynomial pattern) every 10 milliseconds while the model computation takes R milliseconds long where R is sampled from a uniform distribution in $[0, 1000)$ interval for every measurement. This setup leads us to the variable-length window time-synchronous historical modeling. Hence, the window (buffer) holds $N \in [0, 100)$ data points each time and those are used to fit a regression model and predict the next value (Figure 7).

III-C NO_x emission prediction

In this example, Scikit-learn machine learning models were used as digital twin simulation models for predicting NO_x exhaust emission from a real 4-stroke offroad diesel engine during run time. The experiment is the same as described Aho and Immonen (2020) (Section III) except that 1) input feature vectors are normalized to have zero mean and unit variance, and 2) NO_x emission output

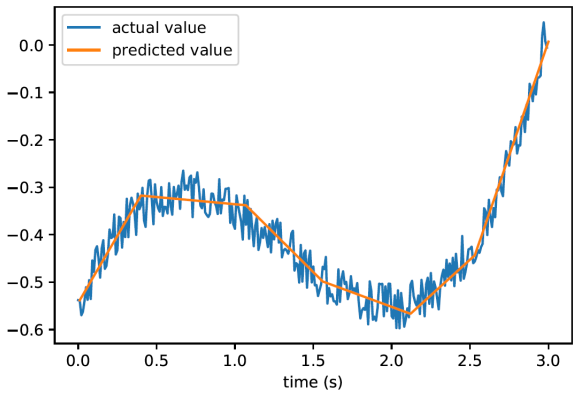


Fig. 7: Demonstrative example of memoryful historical modeling: fit a linear regression model (orange curve) at times to predict next value of time series (blue curve).

value is normalized as relative to initial NO_x output. Three regression models were fitted to the data: Random forest with 100 trees and maximum depth 25, Linear and 2nd order polynomial regression models. The obtained results are shown in Table I as Mean Absolute Error and coefficient of determination (also known as r-squared or R^2) for both training and test data. The normalized data is now publicly available in the project repository on Github.

TABLE I: Training and test results for NO_x emission prediction on different regression models

Model Type	train R^2	test R^2	train MAE	test MAE
Linear regression	0.75	0.73	0.64	0.65
Polynomial regression	0.96	0.96	0.24	0.24
Random forest	0.99	0.99	0.03	0.08

III-D Performance evaluation

A key motivation for the proposed *MCX* development is performance optimization. Table II describes the results of a performance evaluation comparison between *ModelConductor* and *MCX* in different computational experiments. Here, performance is measured in terms of response time, defined as time difference from timestamp client sends the data until the timestamp model response is ready, averaged across a number of samples. The experiments were carried out with the client and model on the same computer to mitigate the effect of random network packet transmission delays. The results show that *MCX* is almost 4 times faster than *ModelConductor*.

TABLE II: Average response time for different experiments: A comparison between *ModelConductor* and *MCX* (in milliseconds)

Experiment	# of samples	ModelConductor	MCX
FMU couple-clutches	4000	52.7	9.9
FMU drone simulation	2500	67.31	18.98
Historical time-series prediction	3000	-	11.54
Sklearn NO_x emission	1422	92.5	22.6

First experiment in the Table II, is a simple open-source simulation for drive train with 3 dynamically

coupled clutches implemented in MapleSim. The other three experiments are explained in Subsections III-A, III-B and III-C respectively. Since *ModelConductor* does not support historical modeling, the average response time regarding historical time-series prediction example in the Table II could not be calculated.

IV CONCLUSIONS

In this work, *MCX* framework was presented as an open-source software platform enabling digital twin implementation by providing asynchronous scalable data transmission facilities as well as online co-execution of different simulation models. Three extensions to the previous version of the framework (namely *ModelConductor*) were described which are: 1) decoupling the queue from the model computation module, 2) usage of MQTT instead of raw TCP, and 3) implementing a solution for time-synchronization in computationally expensive models (memoryful historical models). With these extensions applied, *MCX* performed faster and more scalable compared to *ModelConductor*. In addition, the experimental setup and results of three validation examples of the framework were described.

In spite of the functionality implemented in the *MCX* framework, it has some limitations. First, an explicit waiting routine for the data to be ready is used in measurement handling procedure which consumes processing resources inefficiently. This can be refined with event driven implementation and callback functions to improve the performance. Another limitation is that the framework does not include a standard implementation for a two-way communication, although it is supported by the architecture.

Future research work on the topic should focus on implementing model-based control over *MCX* with the two way communication infrastructure between physical and digital devices. Another interesting direction for future research is identification and adaptation of the simulation model during run time. Here, one begins with a rough model and attempts to refine it as new measurement data becomes available (this feature may not be supported by the current FMU specification). Finally, more real-world validation applications for *MCX* should be considered in the future, also including digital twins for manufacturing processes besides the physical devices considered thus far.

SOURCE CODE AND EXAMPLES

The source code of the framework is available on: https://github.com/COMEA-TUAS/mcx-public

ACKNOWLEDGEMENTS

The authors gratefully acknowledge funding from Business Finland (e3Power project).

REFERENCES

Aho, P., & Immonen, E. (2020, Sep). Modelconductor: An on-line data management architecture for digital twins. In *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)* (Vol. 1, p. 1397-1400).

Beck, F., & Diehl, S. (2011). On the congruence of modularity and code coupling. In *Proceedings of the 19th acm sigsoft symposium and the 13th european conference on foundations of software engineering* (pp. 354–364).

Blochwitz, T., Otter, M., Arnold, M., Bausch, C., Clauß, C., Elmqvist, H., ... others (2011). The functional mockup interface for tool independent exchange of simulation models. In *Proceedings of the 8th international modelica conference* (pp. 105–114).

Cimino, C., Negri, E., & Fumagalli, L. (2019). Review of digital twin applications in manufacturing. *Computers in Industry*, *113*, 103130.

Colombo, A. W., Karnouskos, S., Kaynak, O., Shi, Y., & Yin, S. (2017). Industrial cyberphysical systems: A backbone of the fourth industrial revolution. *IEEE Industrial Electronics Magazine*, *11*(1), 6–16.

de Azevedo, H. D. M., Araújo, A. M., & Bouchonneau, N. (2016). A review of wind turbine bearing condition monitoring: State of the art and challenges. *Renewable and Sustainable Energy Reviews*, *56*, 368–379.

Eclipse, F. (2020). *Ditto: where iot devices and their digital twins get together.* Retrieved 2020-11-01, from `https://www.eclipse.org/ditto/`

Hartman, D., Landis, K., Mehrer, M., Moreno, S., & Kim, J. (2014). *Quadcopter dynamic modeling and simulation (quadsim) v1.00.* Retrieved 2020-11-01, from `https://github.com/dch33/Quad-Sim`

He, R., Chen, G., Dong, C., Sun, S., & Shen, X. (2019). Data-driven digital twin technology for optimized control in process systems. *ISA transactions*, *95*, 221–234.

Lee, W., Suh, E. S., Kwak, W. Y., & Han, H. (2020). Comparative analysis of 5g mobile communication network architectures. *Applied Sciences*, *10*(7), 2478.

Lim, K. Y. H., Zheng, P., & Chen, C.-H. (2019). A state-of-the-art survey of digital twin: techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, 1–25.

Microsoft. (2020). *Azure digital twins: Next-generation iot solutions that model the real world.* Retrieved 2020-11-01, from `https://azure.microsoft.com/en-us/services/digital-twins/`

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

AUTHOR BIOGRAPHIES

SAJAD SHAHSAVARI works as Researcher at Turku University of Applied Sciences, and is a PhD student at University of Turku, Finland.

EERO IMMONEN is an Adjunct Professor at Department of Mathematics at University of Turku, Finland, and works as Principal Lecturer at Turku University of Applied Sciences, Finland.

MOHAMMAD-HASHEM HAGHBAYAN is a post-doctoral researcher at University of Turku, Department of Future Technologies, Finland.

MOHAMMED RABAH (PhD) is working as a Research Engineer at Computational Engineering and Analysis Research Group, Turku University of Applied Sciences, Finland.

JUHA PLOSILA is a Professor in the field of Autonomous Systems and Robotics at the University of Turku, Department of Future Technologies, Finland.

MACHINE LEARNING TECHNOLOGY OVERVIEW IN TERMS OF DIGITAL MARKETING AND PERSONALIZATION

Anna Nikolajeva
Artis Teilans
Faculty of Engineering
Rezekne Academy of Technologies
Atbrivosanas aleja 115, Rezekne LV-4601, Latvia
E-mail: ann@gmz.lv

KEYWORDS
Machine learning, artificial intelligence, digital marketing, personalization.

ABSTRACT

The research is dedicated to artificial intelligence technology usage in digital marketing personalization. The doctoral theses will aim to create a machine learning algorithm that will increase sales by personalized marketing in electronic commerce website. Machine learning algorithms can be used to find the unobservable probability density function in density estimation problems. Learning algorithms learn on their own based on previous experience and generate their sequences of learning experiences, to acquire new skills through self-guided exploration and social interaction with humans. An entirely personalized advertising experience can be a reality in the nearby future using learning algorithms with training data and new behaviour patterns appearance using unsupervised learning algorithms. Artificial intelligence technology will create website specific adverts in all sales funnels individually.

INTRODUCTION

Personalization is the process of adjusting the website to individual users characteristics or preferences. Use to strengthen customer service and e-commerce sales. The website is customized to target each consumer. Personalization means meeting the customers needs more effectively and efficiently, making interactions faster and easier and, consequently, increasing customer satisfaction and the probability of repeat visits (Rouse 2007). There are several personalization software products available, including “Broadvision”, “OptinMonster”, “Monetate” and others.

Retargeting campaigns are advertising that can change in real-time based on user behaviour on the insights gathered from the data. Personalization improves clicks to the top position by 3.5% and reduces the average error in the rank of a click by 9.43% over the baseline (Yoganarasimhan 2019). A survey of 200 marketing leaders by “Forbes Insights” and “Arm Treasure Data” reveals that personalization is giving positive results. Two in five executives 40% report that their customer personalization efforts have had a direct impact on maximizing sales, basket size, and profits indirect channels. Another 37% pointed to increased sales and customer lifetime value through product or content recommendations. More than one-third of respondents have seen an increase in their transaction frequency as result of personalization strategy (Forbes Insights and Arm Treasure Data 2020).

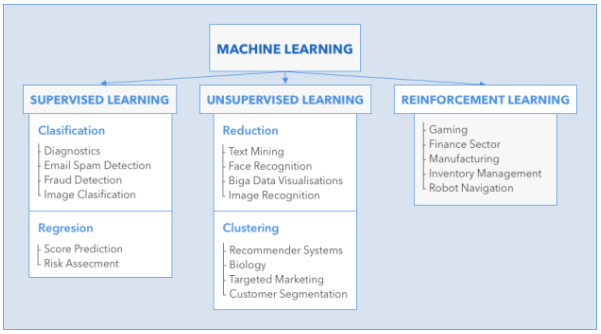
Digital marketing is a form of marketing that focuses on marketing activities in the digital environment, meaning several key activity platforms - email marketing, web browser marketing, social network marketing, smartphone marketing. Each of these platforms uses different channels and technologies to reach its target market. At the heart of digital marketing is the classic marketing need to segment and reach potential or existing customers with a marketing message to drive sales of products or services (Yannopoulos 2011).

Machine learning is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference. It is a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data", to make predictions or decisions without being explicitly programmed to perform the task (Bishop 2006).

Machine learning tasks classified into several broad categories. In supervised learning, the algorithm builds a mathematical model from a set of data that contains both the inputs and the desired outputs. Classification algorithms and regression algorithms are types of supervised learning (Cheeseman et al. 1996). Classification algorithms use when the outputs are restricted to a limited set of values. For a classification algorithm that filters emails, the input would be an incoming email, and the output would be the name of the folder in which to file the email. For an algorithm that identifies spam emails, the output would be the prediction of either "spam" or "not spam", represented by the Boolean values true and false. Regression algorithms are named for their continuous outputs, meaning they may have any value within a range. Examples of a continuous value are the temperature, length, or price of an object (Ryan et al. 2015). Semi-supervised learning algorithms develop mathematical models from incomplete training data, where a portion of the sample input doesn't have labels (Zander et al.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

2005). In unsupervised learning, the algorithm builds a mathematical model from a set of data that contains only inputs and no desired output labels. Unsupervised learning algorithms are used to find structure in the data, like grouping or clustering of data points. Unsupervised learning can discover patterns in the data, and can group the inputs into categories, as in feature learning. Dimensionality reduction is the process of reducing the number of "features", or inputs, in a set of data (Sugiyama 2016).



Figures 1: Machine learning classification

Active learning algorithms access the desired outputs for a limited set of inputs based on a budget and optimize the choice of inputs for which it will acquire training labels. When used interactively, these can be presented to a human user for labelling. Reinforcement learning algorithms are given feedback in the form of positive or negative reinforcement in a dynamic environment and are used in autonomous vehicles or in learning to play a game against a human opponent (Cohnet al. 1996). Other specialized algorithms in machine learning include topic modelling, where the computer program is given a set of natural language documents and finds other documents that cover similar topics. Machine learning algorithms can be used to find the unobservable probability density function in density estimation problems. Meta-learning algorithms learn their own inductive bias based on previous experience. In developmental robotics, robot learning algorithms generate their sequences of learning experiences, also known as a curriculum, to cumulatively acquire new skills through self-guided exploration and social interaction with humans (Hochreiter and Schmidhuber 1997).

Research tasks:

1. Explore machine learning technologies
2. What machine learning technologies can be used to personalize adverts

Research object: machine learning technologies.

The subject of research: machine learning technologies

Research methods: synthesis method, analysis method.

Theoretical background: Theoretical literature of foreign authors, materials published in Internet resources, scientific articles of foreign authors, statistical data, and author personal experience were used in the work.

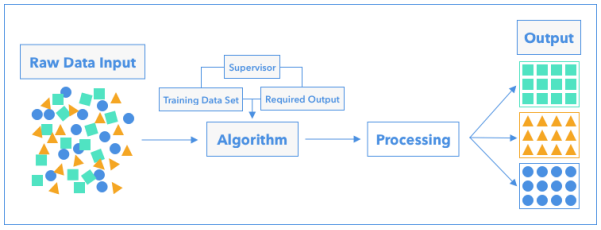
The empirical basis of the research: Worldwide

Study period: June 2019 – November 2020.

Machine learning in economics is still a new subject. Although machine learning is slowly gaining interest among economists, still we see a lack of information. What exactly machine learning entails, what makes it different from classical econometrics and, finally, how economists and businesses along with them can make the best use of it (Athey 2019). There are studies, where is possible to see that data-driven and machine learning prediction in economics is happening. The study was training on 14 years of data, neural networks produce accurate 50-year forecasts. Gaps in these forecasts may reveal macroeconomic regime changes. Failures in otherwise accurate neural network forecasts may thus inform theoretical economic hypotheses through unsupervised machine learning (Chen 2020).

SUPERVISED LEARNING

Supervised learning algorithms include classification and regression. The classification problem is when the output variable is a category, such as “red” or “blue” or “disease” and “no disease”. A classification model tries to draw some conclusions from observed values. Given one or more inputs a classification model will try to predict the value of one or more outcomes. Classification models include logistic regression, decision tree, random forest, gradient-boosted tree, multilayer perceptron. Used in fraud detection, email spam detection, diagnostics, image classification. For example, “Yelp” uses machine learning to organize images in the right categories. “American Express” processes $1 trillion in a transaction and has 110 million cards in operation. They rely heavily on data analytics and machine learning algorithms to help detect fraud in near real-time, therefore saving millions in losses. Additionally, “American Express” is leveraging its data flows to develop apps that can connect a cardholder with products or services and special offers (Marr 2018).



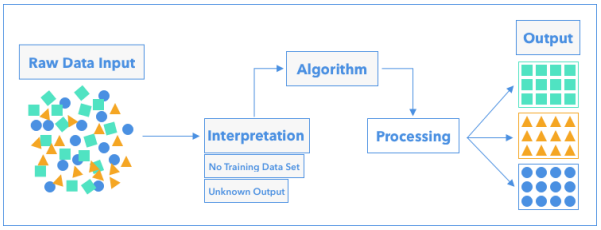
Figures 2: Machine learning clasification: Supervised learning

A regression difficulty is when the output variable is a real or constant value, such as “salary” or “weight”. Many different models can be used, the simplest is the

linear regression. It tries to fit data with the best possible points which go through the points. Used in risk assessment, score prediction. The difference between tasks is the fact that the conditional attribute is numerical for regression and categorical for classification (Weiss and Provost 2001). John Deere is getting data-driven analytical tools and automation into the hands of farmers. Advanced machine learning algorithms allow robots to make decisions based on visual data about whether or not a plan is a pest to treat it with a pesticide. The company already offers automated farm vehicles with pinpoint-accurate GPS systems and its Farmsight system is designed to help agricultural decision-making. Cars are increasingly connected and generate data that can be used in a number of ways. Volvo uses data to help predict when parts would fail or when vehicles need servicing, uphold its impressive safety record by monitoring vehicle performance during hazardous situations and to improve driver and passenger convenience (Marr 2018).

UNSUPERVISED LEARNING

Unsupervised learning algorithms combine clustering the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups. It is the main task of exploratory data mining, and a common technique for statistical data analysis, used in many fields, including pattern recognition, image analysis, information retrieval, bioinformatics, data compression, and targeted marketing (Jain and Dubes 1988). The reduction is mainly used for text mining, face recognition, big data visualizations, image recognition. The North Face uses machine learning to help shoppers find the best outdoor recreation product. Starbucks uses machine learning to recommend drinks thru its app (Bhattacharya 2019).

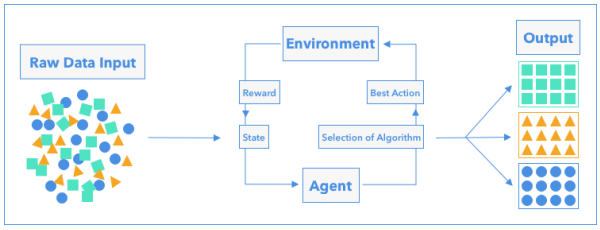


Figures 3: Machine learning clasification: Unsupervised learning

Coca-Cola's global market makes it a large beverage company in the world prospect. The company creates a lot of data, but it has also embraced new technology and puts that data into practice to support new product development, capitalize on artificial intelligence bots and even trialing augmented reality in bottling plants (Marr 2018).

REINFORSMENT LEARNING

Reinforcement learning is an area of machine learning concerned with how program ought to take actions in an environment to maximize some notion of cumulative reward. Due to its generality, the field is studied in many other disciplines, such as game theory, control theory, operations research, information theory, simulation-based optimization, multi-agent systems, swarm intelligence, statistics, and genetic algorithms. Many reinforcement learning algorithms use dynamic programming techniques. Reinforcement learning algorithms are used in autonomous vehicles, gaming, finance sector, manufacturing, inventory management, robot navigation (Qi and Davidson 2009).



Figures 4: Machine learning clasification: Reinforcement learning

Bonsai, recently acquired by Microsoft, offers a reinforcement learning solution to automate and build intelligence into complex and dynamic systems in heating, ventilation, and air conditioning technology of indoor and vehicular environmental comfort, energy manufacturing, automotive and supply chains. The first application in which reinforcement learning gained fame was when AlphaGo, a machine learning algorithm, won against one of the world's best human players in the game Go. Now reinforcement learning is used to compete in all kinds of games (Marr 2018). Reinforcement learning is good for navigating complex environments. It can handle the need to balance certain requirements. Great example is Google's data centers. They used reinforcement learning to balance the need to satisfy our power requirements, but do it as efficiently as possible, cutting major costs (Marr 2018).
Semi-supervised learning falls between unsupervised learning and supervised learning. Many machine-learning researchers have found that unlabeled data, when used in combination with a small amount of labeled data, can produce a considerable improvement in learning accuracy (Kushmerick and Lau 2005).

ADVERTISING WITH MACHINE LEARNING

Self-learning as a machine learning was introduced along with a neural network capable of self-learning. It is learning with no external rewards and no external guides. The self-learning algorithm computes, in a crossbar fashion, both decisions about actions and emotions about consequence situations. It is a system with only one input, situations, and only one output, action. There is neither a separate reinforcement input or an information input from the environment. Exists in two environments, one is a behavioral environment where it behaves, and the other is a genetic

environment, wherefrom it initially and only once receives initial emotions about situations to be encountered in the behavioral environment (Braberman, D'Ippolito, Kramer, Sykes, Uchitel, 2015).

Association rule learning is a rule-based machine learning method for discovering relationships between variables in large databases. It is intended to identify strong rules discovered in databases. Rule-based machine learning is a general term for any machine learning method that identifies, learns, or evolves "rules" to store, manipulate or apply knowledge. The defining characteristic of a rule-based machine learning algorithm is the identification and utilization of a set of relational rules that collectively represent the knowledge captured by the system (Bernadó-Mansilla, Josep, Garrell-Guiu, 2003).

Personalized marketing is a method that utilizes consumer data to modify the user experience to address customers by name, present shoppers with tailored recommendations, and more. Mainly, this is targeted marketing at its most raw. Personalized marketing leverages consumer behavior to present buyers with customized offers (Chittaranjan, Blom, Gatica-Perez, 2013). Spotify has built within their platform daily personalized playlists, daily artist suggestions, and weekly recommendations based on your listening preferences. Customers can even rate the tracks on these playlists based on whether they enjoyed them or not. While this might not directly convert a lead into a sale, it does build brand loyalty and it provides a strong user experience (Perez, 2019). Burberry has been busy reinventing itself and use big data and machine learning to resist counterfeit products and improve sales and customer relationships. The company's strategy for increasing sales is to sustain deep, personal connections with its customers. They have reward and loyalty programs that create data to help them personalize the shopping experience for each customer. They are making the shopping experience at their brick-and-mortar stores just as innovative as an online experience (Marr & Co, 2019).

Marketing automation automatically manages multiple marketing campaigns across several channels. Marketing automation helps with lead generation, segmentation, lead nurturing, lead scoring, customer retention, and more. If it is done correctly, it will increase performance, segment database so users will become clients. For example, chatbots are a unique resource. Through a chatbot, it is quickly discoverable in what stage of the sales tunnel is a user navigating, tracking and analyzing the questions users ask. This allows sales teams to receive the most qualified leads possible. These attributes also can be related to new or returning visitor, temporal variations (time and day of week), channels (mobile and desktop), referral source (social media ad) (Kosinski, Stillwell, Graepel, 2013).

To completely personalize advertisements, it is very important to connect several sales tunnels. The author believes that the unsupervised machine learning algorithm should be used, with relatively large training data. Training data will be collected using statistics, psychological properties, website-specific analytic data, and social media public data. The data amount and specifics will be different depending on websites. The data set will be very large, but it will be helpful, also not only for personalization but also for new pastern appearance. It is unquestionably, that after some amount of time machine learning algorithms will create its patters. The new patterns will be used for a more personalized experience in e-commerce, where consumers will be able to get what they want in relatively fast and entrepreneurs will be able to get most of the profit without any harm to their customers.

Performance optimization is one of the most valuable use cases for machine learning in advertising. Machine learning algorithms are used in commercial solutions to analyze ad performance across specific platforms and recommendations on how to improve performance. In the most exceptional cases, machine learning algorithms can automatically manage ad performance and spend optimization, obtaining decisions entirely on their own regarding whence best to reach advertising KPIs and recommending a fully optimized budget. An example is Google Ads it is possible to use advanced machine learning algorithms in PPC (pay-per-click) campaigns. In bidding, machine learning algorithms train on data at a vast scale to help make more accurate predictions across your account about how different bid amounts might impact conversions or conversion value. These algorithms use a wider range of parameters that impact performance than a single person or team could compute. Bid adjustments allow showing ads more or less frequently based on where, when, and how people search. For example, sometimes a click is worth more than usual if it comes from a smartphone, at a certain time, or from a specific location.

Click fraud is a challenging issue in advertising because it can negatively impact ad budget and harm the integrity of the online advertising market. To detect click fraud, an ensemble learning-based approach is proposed. Click fraud can damage an advertiser's return on investment significantly. It was found that 30% of ad revenue is wasted on click frauds (Choi, Lim, 2020).

Behavioral targeting is used to select the most relevant advertisements for consumers and is based on historical user behavior, such as identifying clicked links, pages visited, searches, earlier purchases from the users browsing history (Choi, Lim, 2020). With the popularity of search engines, such as Google, online searches and web browsing have become two of the most common online behaviours. Web browsing behaviour helps advertisers make assumptions regarding user interests and to define potential audience segments. User online behaviour strengthens the relevance and personalization of advertising messages to wanted consumers. User search queries also help conclude which ads should appear to the user by matching them to the advertiser's keywords.

CONCLUSIONS AND RECOMMENDATIONS

Based on the analysis carried out in the research, the author's personal experience, the analyzes carried out and working on his doctoral thesis, the author came to the following conclusions and recomendations:

1. Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task.
2. Supervised learning algorithms include classification and regression. Classification models include logistic regression, decision tree, random forest, gradient-boosted tree, multilayer perceptron. Used in fraud detection, email spam detection, diagnostics, image classification. Regresion is used in risk assement, score prediction.
3. Unsupervised learning algorithms combine clustering the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.
4. Reinforcement learning is an area of machine learning concerned with how program ought to take actions in an environment to maximize some notion of cumulative reward.
5. Semi-supervised learning falls between unsupervised learning and supervised learning. Many machine-learning researchers have found that unlabeled data, when used in combination with a small amount of labeled data, can produce a considerable improvement in learning accuracy (Kushmerick, Lau, 2005).
6. Self-learning as a machine learning was introduced along with a neural network capable of self-learning. It is learning with no external rewards and no external guides.
7. Many machine-learning researchers have found that unlabeled data, when used in combination with a small amount of labeled data, can produce a considerable improvement in learning accuracy.
8. Association rule learning is a rule-based machine learning method for discovering relationships between variables in large databases.
9. More detailed research should be perpetrated on artificial intelligence tool usage in digital marketing personalization.
10. The author recommends to research data types what can be possibly used in machine learning algorithm training.

Machine learning tasks are classified into several broad categories. In supervised learning, the algorithm builds a mathematical model from a set of data that contains both the inputs and the desired outputs. Unsupervised learning algorithms combine clustering the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups. Reinforcement learning - how to program ought to take actions in an environment to maximize some notion of cumulative reward. Marketing automation automatically manages multiple marketing campaigns across numerous channels. Marketing automation helps with lead generation, segmentation, lead training, lead scoring, customer retention, and more. If done correctly it will increase performance - user converts to clients. Personalization is the most essential thing in e-commerce, as people evaluate their time most. Money has value, but time nowadays is the most valuable thing as people learned to score their time and be productive in all things possible. People were always aimed to get what they want and as fast as possible. Using machine learning algorithms it is possible to create a personal shopping experience. It can be used not only for shopping but also for information. It is important to get relevant news or relevant sponsored adverts. Digitalization is happening now and that means that it is essential to establish new consumer's demands.

REFERENCES

Athey, S. 2019. “The Economics of Artificial Intelligence: An Agenda”. University of Chicago Press. USA. 507 – 547.

Bernadó-Mansilla, E., and Josep, M., Garrell-Guiu. 2003. “Accuracy-based learning classifier systems: models, analysis and applications to classification tasks”. *Evolutionary Computation*. Vol. 11 (No. 3) 209-238.

Bishop, C., M. 2006. “Pattern Recognition and Machine Learning”. *Springer*. *USA*. 10 - 50.

Braberman, V, N, D'Ippolito, Kramer, J., Sykes, D., and Uchitel, S. 2015. “Morph: A reference architecture for configuration and behaviour self-adaptation”. *International Workshop on Control Theory for Software Engineering*. *ACM*. 9-16.

Chen, J., M. 2020. “Economic Forecasting With Autoregressive Methods and Neural Networks”. *SSRN Electronic Journal*. 2 – 40.

Cheeseman, P., and Strutz, J. 1996. “Bayesian Classification: Theory and Results. In Advances in Knowledge Discovery and Data Mining”. *AAI/MIT Press. USA*. 20 - 25.

Choi, J. A., & Lim, K. (2020). Identifying machine learning techniques for classification of target advertising. In ICT Express (Vol. 6, Issue 3, pp. 175–180). Korean Institute of Communications Information Sciences. https://doi.org/10.1016/j.icte.2020.04.012.

Cohn, D., A, Ghahramani, Z., Jordan, M. 1996. “Active learning with statistical models”. *Journal of artificial intelligence research*. Vol. 4 (No. 1). 129-145.

Chittaranjan, G., Blom, J., Gatica-Perez, D. 2013. “Mining large-scale smartphone data for personality studies”. *Personal and Ubiquitous Computing.* Vol. 17 (No. 3). 433-45.

Forbes Insights and Arm Treasure Data. 2020. “Obstacles to Personalization” https://www.forbes.com/sites/insights-treasuredata/2019/05/01/the-path-to-personalization/, viewed 06.04.2020.

Hochreiter, S., and Schmidhuber, J. 1997. “Long short-term memory”. *Neural Computation*. Volume 9 (No. 8). 100 - 288.

Jain, A., K., and Dubes, R., C. 1988. “Algorithms for Clustering Data”. *Prentice Hall, Englewood Cliffs. USA*. 40 -78.

Kushmerick, N., and Lau, T. 2005. “Automated Email Activity Management: An Unsupervised Learning Approach”. *IUI. California.* pp. 67-74.

Kosinski, M., Stillwell, D., and Graepel, T. 2013. “Private traits and attributes are predictable from digital records of human behavior”. *Proceedings of the National Academy of Sciences.* Vol. 110 (No. 15).

Marr, B. 2018. “27 Incredible Examples Of AI And Machine Learning In Practice” https://www.forbes.com/sites/bernardmarr/2018/04/30/27-incredible-examples-of-ai-and-machine-learning-in-practice/#475b95975022, viewed 12.03.2020.

Marr, B. 2018. “Artificial Intelligence: What Is Reinforcement Learning - A Simple Explanation & Practical Examples” https://www.forbes.com/sites/bernardmarr/2018/09/28/artificial-intelligence-what-is-reinforcement-learning-a-simple-explanation-practical-examples/#4bf86d94139c, viewed 12.03.2020.

Marr, B. & Co. 2019. “Burberry: How Big Data and AI is driving success in the fashion world” *https://www.bernardmarr.com/default.asp?contentID=1282,* viewed 12.03.2020.

Perez, S. 2019. “Spotify expands personalization to its programmed playlists” https://techcrunch.com/2019/03/26/spotify-expands-personalization-to-its-programmed-playlists/, viewed 06.04.2020.

Qi, X., Davidson, and B., D. 2009. “Web page classification: Features and algorithms”. *ACM Computing Surveys.* Vol. 41 (No. 2). 1-31.

Rouse, M. 2007. Definition: personalization https://searchcustomerexperience.techtarget.com/definition/personalization, viewed 02.02.2020.

Ryan, M., Talabis, R., Martin, J., L., and Kaye, D. (2015). Information Security Analytics Finding Security Insights, Patterns and Anomalies in Big Data. *Syngress*. *USA*. 1 - 12.

Sugiyama, M. 2016. “Introduction to Statistical Machine Learning”. *Morgan Kaufmann. USA*. 375-390.

Weiss, G., M., and Provost, F. 2001. “The Effect of Class Distribution on Classifier Learning: An Empirical Study”. *Rutgers University. USA. Researchgate Electronic Journal* https://www.researchgate.net/publication/2364670_The_Effect_of_Class_Distribution_on_Classifier_Learning_An_Empirical_Study#fullTextFileContent, viewed 12.01.2020.

Yoganarasimhan, H. 2019. “Search Personalization Using Machine Learning”. *Management Science*. Vol. 66 (No. 3). 1-5.

Yannopoulos, P. 2011. “Impact of the Internet on Marketing Strategy Formulation”. *International Journal of Business and Social Science*. Vol. 2 (No. 18). 1-7.

Zander, S., Nguyen, T., and Armitage, G. 2005. “Automated Traffic Classification and Application Identification using Machine Learning”. *LCN'05. Australia*. 1 - 2.

Zheng, Y. 2019. “Reinforcement Learning and Video Games. University of Sheffield”. 10- 20 https://arxiv.org/pdf/1909.04751.pdf, viewed 06.04.2020.

AUTHOR BIOGRAPHIES

ANNA NIKOLAJEVA was born in Rezekne, Latvia and went to the Rezekne Academy of Technologies, where she studied information technologies and obtained her master degree in 2019. She works in the electronic commerce and digital marketing field last six years and recently started to work as a guest lecturer at Rezekne Academy of Technologies. Her e-mail address is : ann@gmz.lv

ARTIS TEILANS was born in Riga, Latvia and graduated the Riga Technical University, where he studied automatic and remote control and obtained his doctor degree in 1999. In academic field he is a professor and senior researcher at the Rezekne Academy of Technologies. His professional interests include tecniques of system modeling and discrete-event simulation. His e-mail address is : artis.teilans@rta.lv

Finite - Discrete - Element Simulation

INVESTIGATING THE LOAD-BEARING CAPACITY OF ADDITIVELY MANUFACTURED LATTICE STRUCTURES

Dr. János Péter Rádics and Levente Széles
Department of Machine and Product Design
Budapest University of Technology and Economics
Műegyetem rkp. 3., H-1111, Budapest, Hungary
E-mail: szeles.levente@gt3.bme.hu

KEYWORDS

Lattice structure, finite element method, compression test, auxetic, additive manufacturing

ABSTRACT

Additive manufacturing provides unprecedented design freedom from the product's external appearance to the internal structure. Additively manufactured parts, objects can be designed with cellular lattice structures as infills. The application of lattice structures can reduce the required amount of material and desired properties can be assigned to certain objects. There are several different lattice structures each with its own unique, exclusive property or properties. In this study a wide spectrum of so called 'auxetic' and standard lattice structures will be compared using finite element method and compression laboratory tests. The considered auxetic and non-auxetic cellular structures are based on the result of other researches. Along with the aforementioned existing lattices several new structures were proposed. Nine distinct additively manufactured specimens were compared.

INTRODUCTION

Nature inspired cellular structures such as wood and bone are widely used in countless areas of life. Metal foams, carbon fibre reinforced foams and honeycomb based structures are derived from natural cellular materials. (Hang et al. 2019). Owing to their promising mechanical properties, energy absorbing capabilities, impact resistance, high strength and favorable strength to weight ratio lattice structure filled parts and products are used in the automotive industry, aerospace exploration, packaging technology and biomechanics (Hang et al. 2019; Oyindamola and Behrad 2020). Cellular structures are made up from repeating cells, Cellular structures with open cell arrangement are referred to as lattice structures.

Recent advances in additive manufacturing enables the design and production of intricate 3D lattice structures. Knowing the behavior, mechanical, deformational and thermal properties of certain lattice structures we can design parts with prominent desired attributes.

In recent years several studies focused on the outstanding characteristics of lattice structures and their realization with additive manufacturing techniques. In particular, the creation of metamaterials gained outstanding attention. Metamaterial are materials which derives their properties form their structure and not form the actual material they are made from. The Greek word "meta" means "beyond", metamaterials can exhibit properties beyond the product's forming material.

Negative Possion's ratio (NPR) materials, structures and foams are of great interest in recent studies. Foams with negative Poisson's ratio were first created by Lakers in 1987 (Yongguang et al. 2019), Lakers named these materials 'auxetics'. Deformational response of auxetic materials to compression and stretching is ultimately different from traditional materials. Based on the definition of Poisson's ratio: the negative ratio of the transverse strain to the longitudal strain.

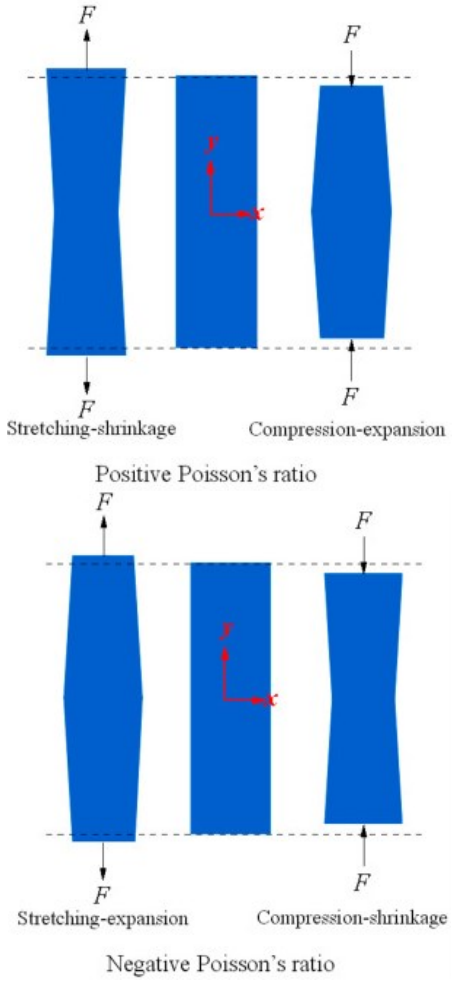


Figure 1: Deformational Response of Auxetic and Non-auxetic material (Chanfang et al 2020)

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

Based on the definition, when the Possion's ration is positive the material's deformation is stretching-shrinkage and compression-expansion. On contrary, when the Poisson's ration is negative the deformation is stretching-expansion and compression-shrinkage (Changfang et al. 2020). Figure 1 illustrates the deformational response of standard and auxetic materials.
Energy absorption, load bearing and deformational capacity can be obtained from compression tests, hence in this study compression tests and simulations were adopted.
The realization of additively manufactured specimens and subsequent laboratory testing is time-consuming and costly. Employing finite element simulations, the effect of certain lattice parameters can be obtained more efficiently. Finite element simulation results must be verified by a series of measurements conducted on real specimens. In this study the result of FEM simulations and laboratory compression tests are compared.

MATERIALS AND METHODS

In this section the reasoning and the process of creating comparable specimens is presented.

Structure of the proposed specimens

In total nine different specimens were created, each built with different elements. Being a comparative study, the specimens were designed from many aspects so that the result of the laboratory tests and FEM result can be compared among specimens.

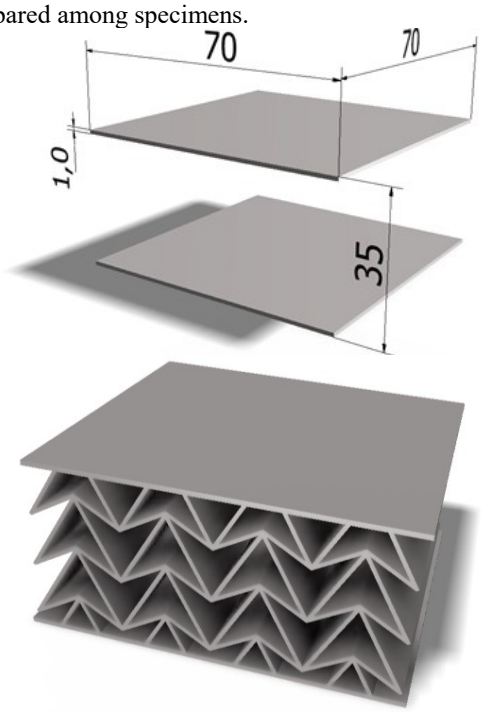
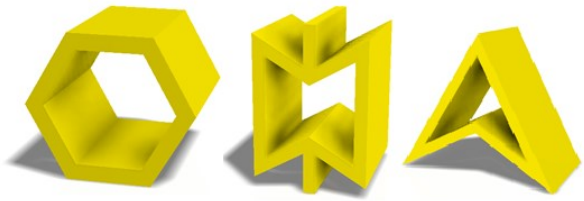


Figure 2: Overall Dimensions of the specimens and one example Specimen with Lattice structure

The overall dimensions are equal among test pieces; the overall height of a specimen is 35 mm, the top and bottom plates are 70 mm by 70 mm and 1 mm in thickness.
As illustrated in figure 2 the space in-between the two planes is filled with different lattice structures. Each specimen is built up using only one structure, there were no combinational experiments considered in this study. Figure 2 illustrates the hollow specimens with general dimensions and an example specimen made up from concave arrow cells.
Specimens were designed to have the same weight thus results are more clearly comparable. More precisely the weight of the test pieces consisting of 2.5 and 3 dimensional elements are the same.

Examined lattice structures

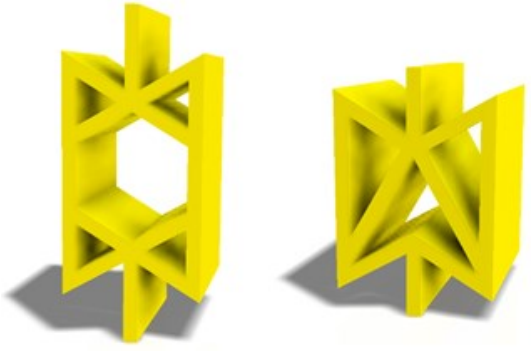
As mentioned in the introduction alongside the existing lattice structures three unique, newly created structures will be compared in this study. Figure 3 represent three existing 2.5 dimensional and broadly examined unit cell geometries. The Regular Honeycomb structure unlike the other two structure on Figure 3 does not demonstrates auxetic behavior.



a) b) c)

Figure 3. a) Regular Honeycomb unit cell; b) Vertical Auxetic honeycomb unit cell; and c) Arrowhead unit cell

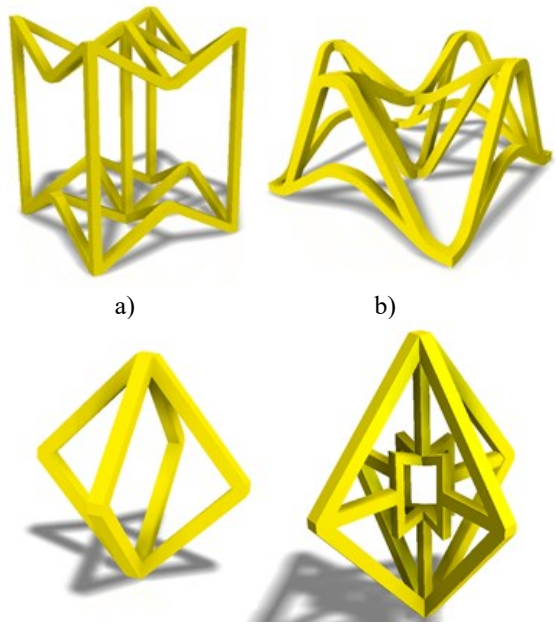
Based on the research result of the unit cell types shown on Figure 3 two new 2.5 dimensional structures were proposed. The created lattice structures are combinations of existing cell structures; a combined honeycomb unit cell and a combined auxetic unit cell was created. Figure 4 illustrates the aforementioned unit cell structures.



a) b)

Figure 4. a) Combined Honeycomb unit cell; b) Combined Auxetic unit cell

Besides 2.5 dimensional unit cell types 3 dimensional ones were considered as well. The so called Octahedron unit cell and the auxetic Double-V and Double-U hierarchical structures (Hang et al.) were studied.
A new unit cell geometry called "Semi Auxetic Octahedron", based on the combination of the Octahedron unit cell and two Vertical Auxetic Honeycomb unit cells is introduced and investigated in this paper.

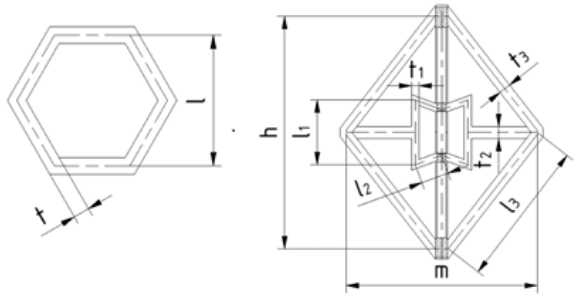


a) b)

c) d)

Figure 5. a) 3D Double-V unit cell; 3D Double-U unit cell; c) 3D Octahedron unit cell; and d) Semi-Auxetic Octahedron unit cell structures

Specimens were created using parametric adaptive computer aided modelling. The previously introduced unit cell geometries can be specified by a series of dimensions, parameters as Figure 6 illustrates. Adaptive modelling enables rapid creation of new specimens for future parameter based investigations.



a) b)

Figure 6. Examples of Parametric Drawings for a) traditional honeycomb unit cell; and b) semi-auxetic octahedron unit cells

The above listed nine structures form the basis of the present study.

Fabrication of specimens

Specimens were realized using selective laser sintering (SLS) technology with an HP Multi Jet Fusion 4200 type printer.
The material used was HP's PA 12 (MJF), material properties are listed in Table 1.

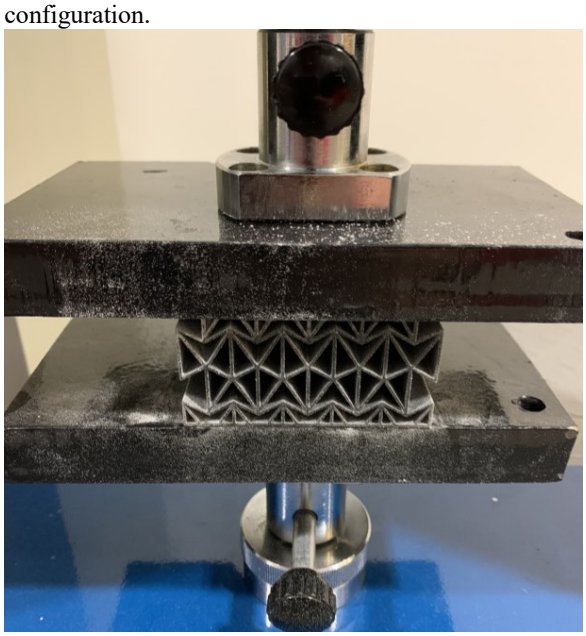
Table 1. Material Properties of PA12

Material constants	**Value**
Density [kg/m^3]	1130
Poisson's ratio [-]	0.35
Tensile modulus [MPa]	1800
Tensile strength [MPa]	49
Elongation at break [%]	20

The printed test pieces were removed from the build unit and also the powder from the lattice structures followed by sandblasting.

Methodology of laboratory testing

Load bearing capacities, deformational response and the absorbed amount of energy can be obtained by subjecting the specimens to compressive load.
Each specimen is compressed by a dual-column (twin-ball screws) tensile testing machine (KINSGEO KJ-1066A type). The conducted measurements were load controlled; force is measured by the S-beam load cell of the testing machine. Displacement is measured via the rotary encoder mounted to the motor.
The measurements are ended when rapid failure begins or when the measurement limit of 5000 N is reached. Force – displacement curves are plotted for each measurement. Figure 7 illustrates the measurement configuration.



Figures 7: Measurement Configuration for Compression tests

The measured values are adjusted by the weight of the black steel plate placed on top of the test samples.

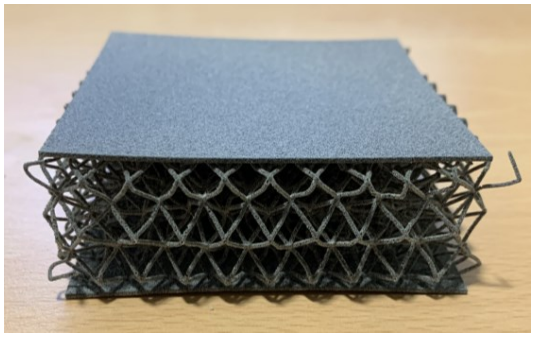


Figure 8: The beginning of Rapid Failure in the octahedron based specimen

Figure 8 illustrates a specimen in which several unit cells broke, resulting in rapid failure.
Figure 9 on the other hand represents a test piece which withstood the measurement limit without any significant failure.

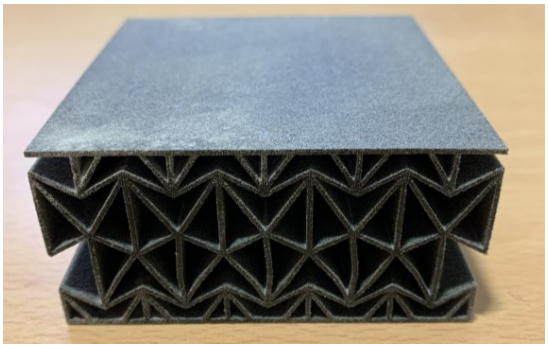


Figure 9. The combined auxetic cell based specimens withstood the measurement limit

Results of the compression test

Measured force-displacement diagrams are shown on figure 10, 11 and 12 for the Regular Honeycomb, the Octahedron and the Vertical Auxetic Honeycomb lattice based specimen respectively.

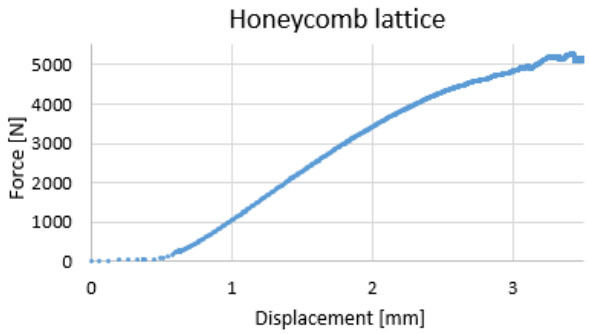


Figure 10. Measured Force – Displacement Curve of the Standard 2.5D Honeycomb lattice.

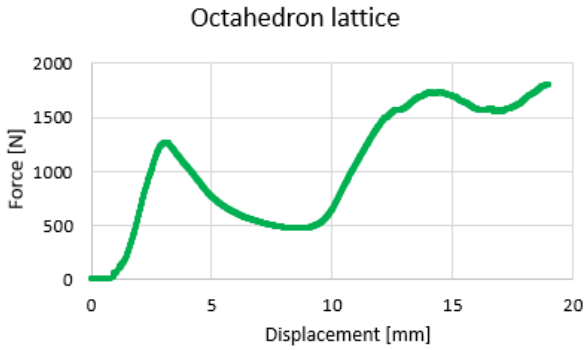


Figure 11. Measured Force – Displacement Curve of the 3D Octahedron lattice

In this study the load-bearing capacity of specimens of the same size and weight were tested. The value of the greatest force endured by the nine examined specimen variations is listed in Table 2. The measurement limit was 5000 N, thus specimens with 5000 N (5300 N) load-bearing capacity have even greater load bearing capabilities.

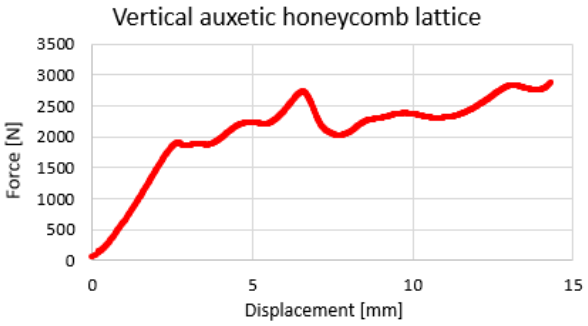


Figure 12. Measured Force – Displacement curve of the Vertical Auxetic Honeycomb lattice

Table 2. The value of the Greatest Force Endured by specific lattice types

Lattice structure type	**Greatest force endured**	**Maximum displacement**
Standard Honeycomb (2.5D)	5300 N	2.5 mm
Vertical Auxetic Honeycomb (2.5D)	4770 N	13.5 mm
Arrowhead (2.5D)	2700 N	12.7 mm
Combined Auxetic (2.5D)	5300 N	12.5 mm
Combined Honeycomb (2.5D)	5300 N	11.5 mm
Octahedron (3D)	2600 N	7.5 mm
Semi-Auxetic Octahedron (3D)	4000 N	6 mm
Double-V (3D)	2400 N	14 mm
Double-U (3D)	750 N	24 mm

The load bearing capacity is just one of the many properties a certain lattice structure can be characterized by. Energy absorption and the shape, characteristics of the force – displacement curve are also important features.
In order to increase the reliability of the laboratory tests three specimens were additively manufactured from each lattice type.

Finite element simulation

Finite element method is used to simulate the response of the specimens under quasi-static compression load using Ansys Workbench. The material properties are listed in Table 1. Isotropic elasticity material model was used for the FEM simulations.
The finite element boundary conditions were set according to the actual compression test; fixed support was used at the bottom plane of the test pieces.
To compare specimens load force of 2550 N was applied on the top plane of the specimens. To achieve uniform results, the force was applied only in the middle 50 mm by 50 mm surface area. Solid element type with 0,3 mm element size was applied for all studies resulting in a fine mesh. Figure 13 and 14 illustrates the deformation and the Von Mises stress distribution in the Vertical Auxetic Honeycomb cell based specimen.

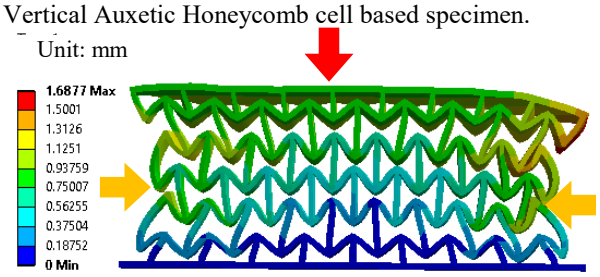


Figure 13. FEM Resulting Displacement

As its name suggests the Vertical Auxetic unit cell shows auxetic properties; the characteristic compression-shrinkage behavior can be obtained on Figure 13. Based on the FEM deformational result, simulations can be deemed acceptable.
Distribution of the Equivalent (Von-Mises) stress was considered in each simulation. Peak mechanical stresses can indicate possible failure segments on the specimens.

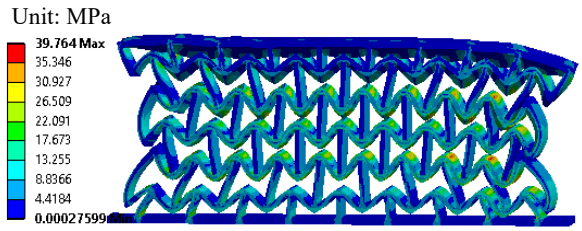


Figure 14. FEM von-Mises Stress Distribution

Results of the FEM simulations are listed in Table 3.

Table 3. Results of the FEM Simulation

Lattice structure type	Maximum displacement	Maximum stress.
Standard Honeycomb (2.5D)	0.96 mm	43.82 MPa
Vertical Auxetic honeycomb (2.5D)	1.69 mm	39.76 MPa
Arrowhead (2.5D)	7.72 mm	123.3 MPa
Combined Auxetic (2.5D)	2.42 mm	50.74 MPa
Combined Honeycomb (2.5D)	1.50 mm	73.98 MPa
Octahedron (3D)	2.59 mm	109.9 MPa
Semi-Auxetic Octahedron (3D)	2.45 mm	160.4 MPa
Double-V (3D)	23.8 mm	490 MPa
Double-U (3D)	33.9 mm	840 MPa

In specimens with maximum Von-Mises stress values greater than the tensile strength of PA12 the high stress areas had a significant extent, thus the specimens are prone to brake during laboratory tests.

COMPARING AND EVALUTAING THE RESULTS

The study focused on determining and comparing the load bearing capacity of different additively manufactured lattice structures. Another aim of this research was to compare the results of the finite element method with the results of the compression laboratory tests. Existing and newly proposed lattices were considered.
The behavior characteristic (compressive shrinkage) of auxetic materials is displayed by the laboratory tests and FEM simulations as well. Among the lattice samples examined the maximum load bearing capacity of non-auxetic lattices is greater. On the other hand, having compared the force-displacement curves of auxetic and non auxetic lattices (Figure 12 and Figure 10) it can be stated that auxetic structures in general have greater energy absorption capabilities.
FEM simulation results and laboratory test results are comparable based on Table 2. and Table 3 the following statements can be made. Those specimens which failed at lower load levels showed greater stress values at FEM simulations next to equal loads (for example 2.5D arrowhead lattice failed at 2700N and the greatest stress value in FEM simulations was 123.3 MPa).
Relative (maximum) displacement shows comparable results on the simulations and measurements as well; auxetic and non-auxetic behaviors can be recognized.
Among the newly proposed lattice structures the "Combined Honeycomb (2.5D)" and the "Semi-Auxetic Octahedron (3D)" geometries exhibited outstanding load-bearing capabilities, next to significant deformation values.

The “Combined Auxetic (2.5D) lattice still presented expressive load-bearing capability; the relatively great enclosed area under the curve characterizes the preeminent total absorb energy (Jianjun et al. 2020).

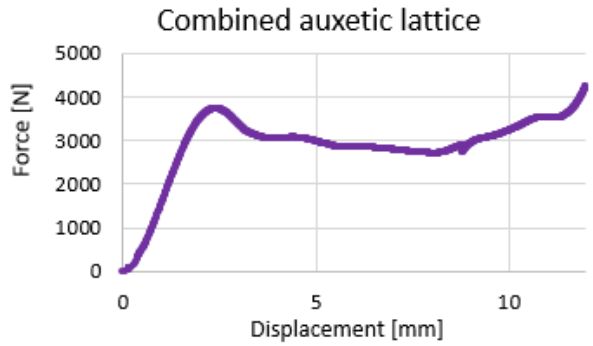


Figure 15 Measured Force-Displacement Curve of the Combined auxetic lattice

Comparing the results of the simulation and the laboratory tests advocates that the behaviors and characteristics of lattice structures can be obtained using finite element computer simulation.

CONCLUSIONS

Based on the Von-Mises stress distribution obtained from FEM simulations we can predict with high confidence the possible failure points. Comparing the stress distribution of FEM simulation especially the high stress areas it can be stated that failure will occur at these regions in real life applications. The finite element simulation method can be applied to study the behavior of existing and newly created lattice structures.

In this study three newly proposed lattice structures were examined. The “Combined Auxetic cell” in our simulations and laboratory tests presented auxetic properties. On the other hand, the two structures combined from auxetic and non-auxetic geometries did not represented auxetic properties. Based solely on our research it can be declared that the combination of auxetic stucutres will result in auxetic behavior, however, further comprehensive research is recommended. In summary the newly created lattice structures proved to be promising, further study and development is recommended.

In further stages of the research it is advised to consider the effect of geometrical parameters on the behavior of specimens. Figure 6 illustrates the geometrical parameters for certain lattice structures, changing geometrical parameters will affect the mechanical properties of a lattice structures. Establishing relationships between geometrical parameters and physical properties can provide a decision preparation basis for choosing the most adequate structure and parameter for a certain application.

ACKNOWLEDGEMENTS

The research reported in this paper and carried out at BME has been supported by the NRDI Fund (TKP2020 NC, Grant No. BME-NCS) based on the charter of bolster issued by the NRDI Office under the auspices of the Ministry for Innovation and Technology.

REFERENCES

Changfang et al. 2020. “The in-plane stretching and compression mechanics of Negative Poisson’s ratio structures: Concave hexagon, star shape, and their combination”

H. M. A. Kolken and A. A. Zadpoor. 2017. “Auxetic mechanical metamaterials”

Hang Yang, Bing Wang and Li Ma. 2019. “Mechanical properties of 3D double-U auxetic structures”

Jianjun Zhang, Guoxing Lu and Zhong You. 2020. “Large deformation and energy absorption of additively manufactured auxetic materials and structures: A review”

Jonathan Simpson and Zafer Kazanci. 2020. “Crushing investigation of crash boxes filled with honeycomb re-entrant (auxetic) lattices”

Oyindamola Rahman and Behrad Koohbor. 2020. “Optimization of energy absorption performance of polymer honeycombs by density gradation”

Yongguang et al. 2019. “Deformation behaviors and energy absorption of auxetic lattice cylindrical structures under axial crushing load”

AUTHOR BIOGRAPHIES

DR. JÁNOS PÉTER RÁDICS is an assistant professor and deputy head at the Department of Machine and Product Design BME. His research area is Design for Additive Manufacturing (DfAM), soil CHG emission, soli-tool wear and DEM simulation of agricultural material behavior. His email address is: radics.janos@gt3.bme.hu.

LEVENTE SZÉLES was born in Balassagyarmat, Hungary. He is PhD student at the Department of Machine and Product Design BME. Levente graduated from BME with BSc and MSc degrees in 2016 and 2020 respectively. His research field is additive manufacturing, specifically creating design guidelines for additive manufacturing. His email address is: szeles.levente@gt3.bme.hu.

FE MODEL OF A CORD-RUBBER RAILWAY BRAKE TUBE SUBJECTED TO EXTREME OPERATIONAL LOADS ON A REVERSE CURVE TEST TRACK

Gyula Szabó, Károly Váradi
Department of Machine and Product Design
Budapest University of Technology and Economics
1-3 Műegyetem rkp., Budapest 1111, Hungary
E-mail:szabo.gyula@gt3.bme.hu

KEYWORDS
filament-wound composite tube, cord-rubber tube, railway brake tube, FE model, sub-zero temperature, reverse curve, draw and buffing gear interaction test

ABSTRACT

In certain cases, rolling stocks and railway vehicle components, i.e. brake tubes need to operate under extreme conditions such as at sub-zero temperature (e.g. -40°C) and on a reverse curve track, when displacements of the suspension points of the tubes cause large deformations in tubes.
In this paper, displacements of the suspension points of the tubes are determined by a kinematic model validated by a draw and buffing gear test [1]. Afterwards, FE simulation has been carried out at minimum and maximum suspension point distance based on these displacements for the investigation of stress, strain states and possible failure considering the case of internal pressure and no internal pressure.
Equivalent strain, stress and Tsai-Hill failure indices are much below the criterial values, so failure is not probable.
The straight section between the curves of opposite curvatures reduces deformation in tubes in the critical positions leading to lower strain, stress and failure index values.

INTRODUCTION

Cord-rubber composite tubes are widely used in applications where relatively high loads need to be withstood with low dead load, and very large deformations occur. These brake tubes are produced mainly by filament-winding due to high fiber precision and good automation capability with low tooling costs [2]. Winding angle of cords is usually ±55° making filament-wound tubes optimal to biaxial tension (when the loads are uniaxial tension and internal pressure) [3]. Filament-wound cord-rubber tubes also find extensive use in railway transportation as railway brake tubes due to their high strength-to-weight ratio, high flexibility and corrosion resistance. Cord-rubber tubes installed on railway carriages can be seen in Figure 1.

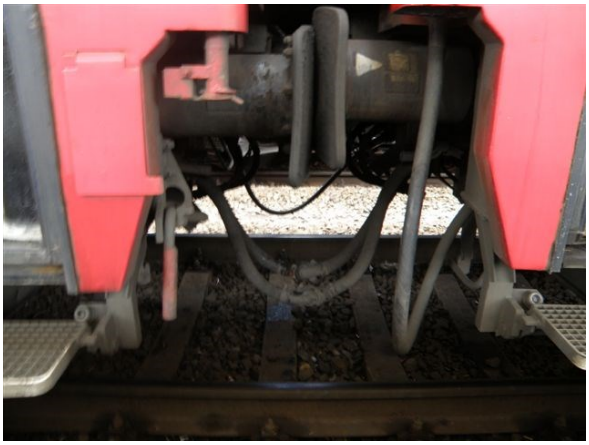


Figure 1. Railway brake tubes installed on railway carriages

Operation on reverse curve tracks (when both rolling stocks are on curves of opposite curvatures) induce large displacement of the suspension points of the brake tubes thus leading to large deformation of the tubes. However, deformation is impeded at sub-zero temperatures (e.g. -40°C) due to the elevated stiffness of the material therefore bringing about higher stresses than at room temperature.

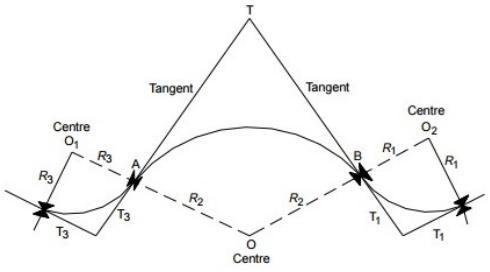


Figure 2. Schematic representation of a reverse curve track [4]

Schematic depiction of a reverse curve can be seen in Figure 2. A reverse curve (also known as S-curve) consists of curves having opposite curvatures.These curves are connected to one another by a transition curve or a straight line. The length of the straight line between the two curves is usually more than 30 m in high-speed applications although the connecting straight lines can be much shorter on maintenance tracks, being

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

a worse case regarding the operational conditions because displacements of the suspension points of the tubes are larger-even though operation on maintenance tracks is not considered in railway rolling stock design practice. [4]
In traditional railway transportation (non-high speed applications), the minium curve radius is 150 m and an intermediate straight section of 6 m must be placed between the curves of opposite curvatures to ensure vehicle stability. In this case, track gauge is 1470 mm. [5].
The aim of this article is to investigate deformation, stress and strain states and possible failure of cord-rubber railway brake tubes subjected to bending and torsion due to operation on an S-curve at -40°C. Materials and methods are similar to those used in [6], however, the railway test track is different which has a significant effect on displacements of the suspension points of the tubes, and thus on the resulting strain and stress states. Furtmore, in this paper, validation of the kinematic simulation has been performed additionally based on [1].

DRAW AND BUFFING GEAR INTERACTION TEST [1]

Railway rolling stock manufactured in the European Union needs to fulfil the requirements of LOC&PAS TSI [7], so draw and buffing gear interaction test [1] of a coach manufactured by MÁV-Start Zrt (Hungarian State Railways) had to be performed. The test was carried out by MÁV Central Rail And Track Inspection Ltd. (MÁV KFV Kft.) on the Szolnok Railway Maintenance Site of MÁV in Szolnok, Hungary between 16 June 2019 and 20 June 2019 on a purpose-built reverse curve test track with a nominal curve radius of 150 m. Length of the vehicle body has been 26100 mm, distance of the bogie pivots has been 19 000 mm, bogie type was Siemens SF 400. Tests were conducted at a speed of 5km/h so that forces exerted by the traction vehicle did not influence the measurements. [1]

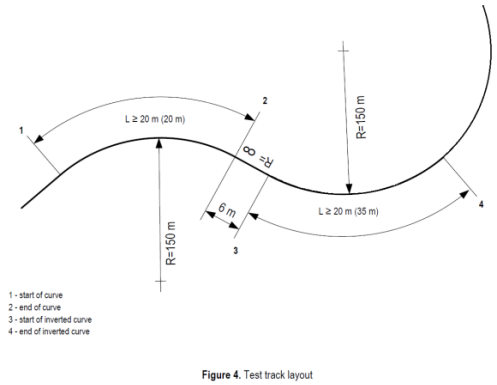


Figure 3. Test track layout [1]

The test track is illustrated in Figure 3. It consists of an initial tangent (straight) section before 1, a 20 m long curve with a radius of 150 m (between 1 and 2), an intermediate straight section of 6 m (between 2 and 3), a 35 m long curve with a radius of 150 m (between 3 and 4) and a curved section after 4 with an approximate radius of 150 m. The test (pulled run is considered hereinafter) commenced at point 1 and lasted until the rear coach has passed point 4.

Figure 4 shows the displacement-type measured quantities. X_1 is the buffer compression on the left side, x_2 is the buffer compression on the right side, whereas y_1 is the lateral deviation at the coupled end (Figure 4).

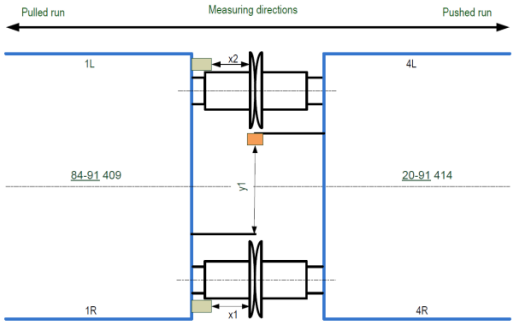


Figure 4. Location of the applied sensors [1]

Lateral displacement y_1 has been used for the validation of the kinematic simulation. It has been measured by a draw-wire displacement sensor as it can be seen in Figure 5. Lateral displacement results can be seen in Figure 6 as a function of measurement time.

Figure 5. Draw-wire displacement sensor utilized for the measurement of lateral displacement y_1.

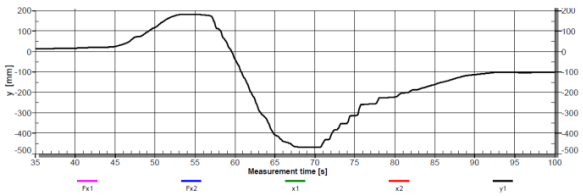


Figure 6. Lateral displacement y_1 as a function of time

KINEMATIC SIMULATION

Kinematic simulation has been performed in PTC Creo 2.0 Mechanism environment based on the dimensions of the draw and buffing gear interaction test. The kinematic model along with the railway track is shown in Figure 7. The purpose of the kinematic model is to obtain the positions of the suspension points of the positioning pins of the tubes at critical cases (minimum and maximum suspension point distance) in order to determine the displacements of these suspension points required for the FE model.

Figure 7. Representation of the railway track in the kinematic model

The assembly can be seen with its constraints in Figure 8. The dimensions (Figure 9) are in accordance with railway standards [8,9] and the test report [1].The railway track is grounded, to which bogies are attached by *slot* constraints at the railway wheels. Bogies and carriages are connected by *pin* contacts at the bogie pivots, which permit only rotational displacement around their axes. Carriages are connected to one another by *the draw gear*, which is able to rotate around fix points on the carriages (represented as *pin* contacts). The assembly is driven by a servo motor attached to one of the slot constraints on the rightmost bogie (Figure 8). [6]
The assembly contains the control points on both carriages required for the validation of the kinematic simulation (Figure 8).

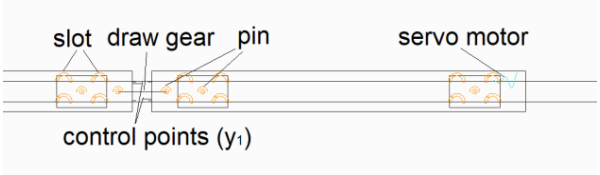


Figure 8. Schematic representation of the kinematic model along with constraints

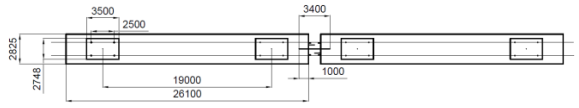


Figure 9. Schematic drawing of the kinematic model with dimensions

Relative displacements of the control points (Figure 8) can be seen as a function of time in Figure 10. Comparing results of the kinematic simulation with results of the draw and buffing gear test (Figure 10 and Figure 6 respectively), a considerably good agreement can be observed. The two curves have the same tendencies and the periodicity is nearly identical. However there is a minor difference in terms of the extremums, (the maximum lateral displacement is 312 mm instead of 200 mm and the minimum lateral displacement is -576 mm instead of approximately -470 mm), this can be attributed to several factors that could not be taken into account in the simulation.
In the test train, there are clearances in the subassemblies, and there is a slight lag in the movement of the bogie pivots.

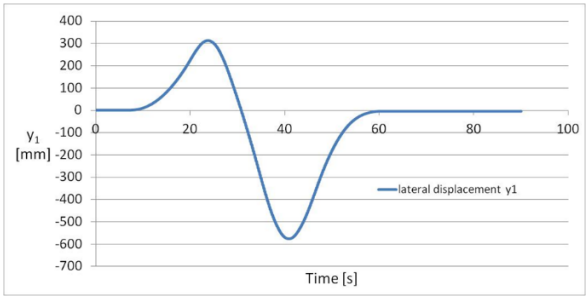


Figure 10. Lateral displacement (y_1)-time

Results of the kinematic model is necessary to estimate the displacements of the suspension points of the positioning pins of the tubes that govern the movement of the tubes [6]. Below, the left pair of tubes is considered.
At the start of the simulation, distance of the suspension points is 840 mm, which firstly diminishes because of being on the inner side of the first curve, until a minimum of 770 mm. Then, the distance increases as the bogies of the first coach begin to negotiate the second curve. After reaching the maximum of 985 mm, the bogies of the second coach get on the second track and the distance between the suspension points of the tubes slightly decreases. The extremums are considered as worst cases regarding the positions of the tubes [6].

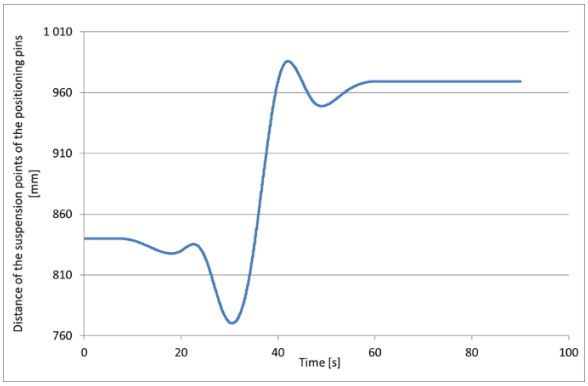


Figure 11. Distance of the suspension points of the positioning pins as a function of simulation time

Minimum and maximum suspension point distances are illustrated in Figure 12 and Figure 13 respectively.

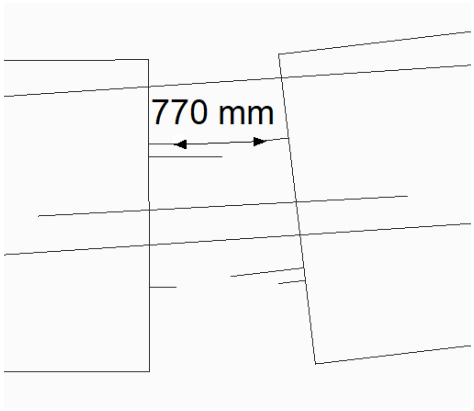


Figure 12. Railway cars at the moment of minimum suspension point distance

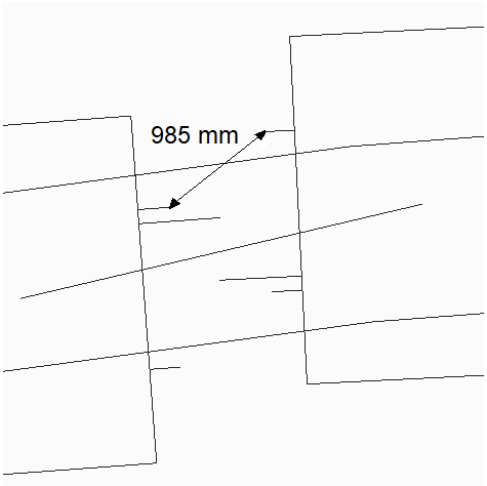


Figure 13. Railway cars at the moment of maximum suspension point distance

The displacements of the suspension points of the tubes in the FE model (see Table 1) have been calculated based on the positions acquired from the kinematic model (Figure 12, Figure 13).

FE MODEL [6]

The objective of the FE simulation is to gain strain, stress and Tsai-Hill failure index distributions for the assessment of failure at -40°C and the prescribed displacements derived from the kinematic simulation on the reverse curve track.

The FE model consists of two tubes and positioning pins on the two sides of the tubes used for actuating the tubes and positioning pins needed to connect the tubes.

Each tube is 620 mm long, consisting of an inner rubber liner, reinforcement layers and an outer rubber liner, shown in Figure 14. The layup is [+55°/-55°/+55°/-55°] and the material coordinate system of the tube is cylindrical.

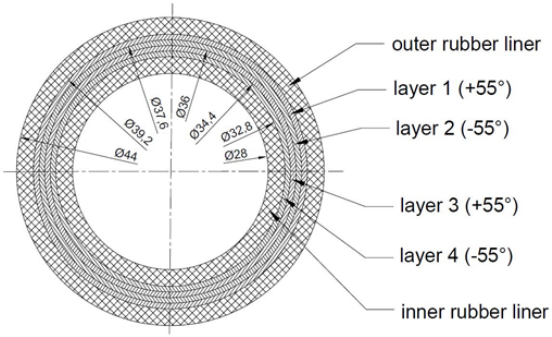


Figure 14. Cross-section of the tube [10]

Material model of the reinforcement layers is transversely isotropic, being a special case of orthotropy. Material properties have been calculated by utilizing the formulae of rules of mixture [11] based on the material properties of the components and fiber volume fraction. Material properties are considered *at -40°C.*

Material properties of the reinforcement layers are the following: modulus of elasticity of fibre is E_f=2961 MPa, Poisson's ratio of fibre is supposed to be υ_f=0.2, modulus of elasticity of rubber matrix is $E_m=E_r$= 19.1 MPa. E_1=1345 MPa, $E_2=E_3$=57 MPa, $\upsilon_{12}=\upsilon_{13}$=0.3637, υ_{23}=0.496, $G_{12}=G_{23}=G_{13}$=19 MPa. Rubber liners, made of EPDM-EVA compound, regarded as incompressible, have been described by a 2 parameter Mooney-Rivlin model with parameters C_{10}=3.34 MPa, C_{01}=1.077 MPa, D=0 1/MPa. These material properties have been validated previously by uniaxial tensile tests performed on test specimens and tube pieces [12] and deflection test carried out at -40°C. [6]

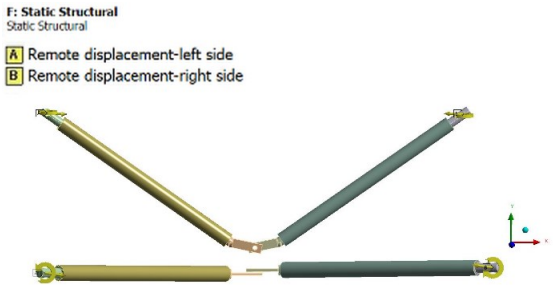


Figure 15. Prescribed displacements in the front view and in top view [6]

Actuation of the tubes is performed by prescribed displacements (Figure 15) of the suspension points of the positioning pins based on the kinematic simulation. The tubes are connected to each other in the middle by a fixed joint of holes of two positioning pins, whose distance is 10 mm in direction Z (see Figure 15).

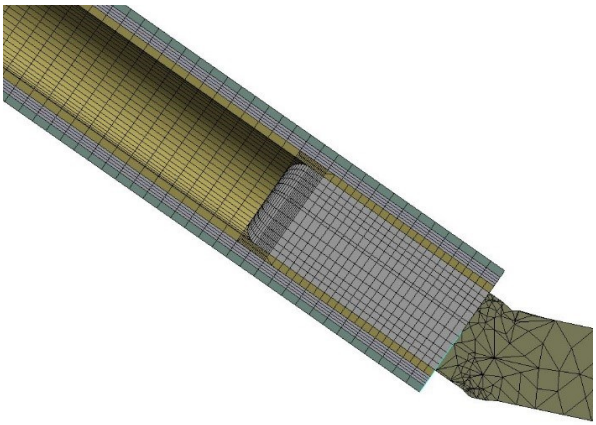


Figure 16. Meshed FE model [6]

The Finite Element model consists of 151548 nodes, 110800 SOLID185 hexahedral elements and 50260 SOLID 187 tetrahedral elements, a detail of the mesh can be seen in Figure 16. (The positioning pin consists of two bodies, bonded to one another, due to meshing considerations.)
At the beginning of the FE simulation (t=0 s), the distance of the suspension points of the positioning pins is 1140 mm. Prescribed displacements for minimum and maximum suspension point distance load cases are listed in Table 1 relative to the initial configuration.

Table 1. Prescribed displacements of FE model without internal pressure

	translation X [mm]	translation Z [mm]	rotation Y
min. s. p. d. left	184.86	0	3.41
min. s. p. d. right	-184.86	0	-3.41
max. s.p. d. left	163.5	278.5	1.55
max. s. p. d. right	-163.5	-278.5	-1.55

where *min. s. p. d. left* is the abbreviation of minimum suspension point distance load case-left remote point *max. s. p. d. right* is the abbreviation of maximum suspension point distance load case-right remote point

The load case of additional internal pressure in case of the minimum and maximum suspension point distances has been further investigated (Figure 17). By utilizing the symmetry of the model, in these examinations, only a half model has been examined (the left tube and its pins). The pressure load is 5 bar, the displacement of the connecting remote point (in the middle) and the displacement of the suspension point of the positioning pin match those of the displacement results of the simulation without internal pressure. This simulation consists of two time steps. In the first one, an internal pressure of 5 bar is applied to the inner lateral surface of the tube, while in the second one, the prescribed displacement results are applied.

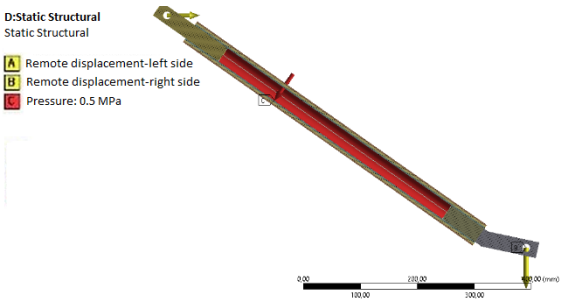


Figure 17. Half FE model with internal pressure [6]

Failure behaviour has been analysed with Tsai-Hill failure criterion. Tsai-Hill criterion is widely utilized for describing failure behaviour of cord-rubber composites [6, 13].
Strength properties in material directions are as follows: X_t=X_c=342.5 MPa , Y_t= Y_c=Z_t= Z_c=34.2 MPa (assuming strength is much lower in transverse directions); Q=R=S=5.5 MPa, derived from the strength of the rubber at -40°C (half of the strength of the rubber) [6]

RESULTS

RESULTS WITHOUT INTERNAL PRESSURE

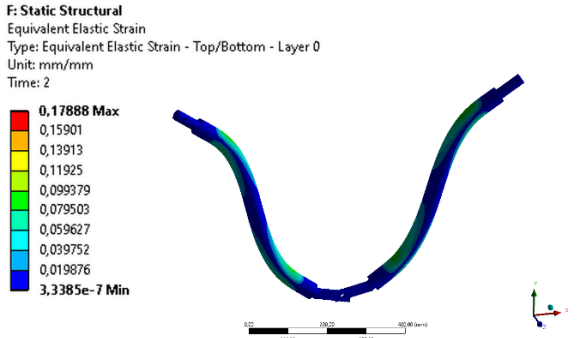


Figure 18. Equivalent strain at minimum suspension point distance (deformation scale 1:1)

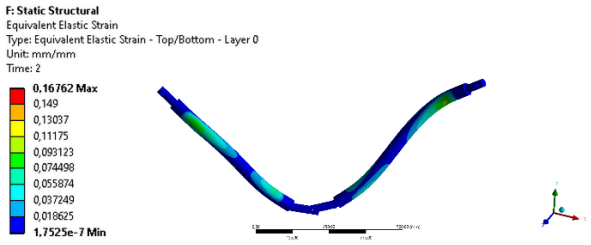


Figure 19. Equivalent strain at maximum suspension point distance (deformation scale 1:1)

Figure 18 shows equivalent strain results at minimum suspension point distance, while Figure 19 shows equivalent strains at maximum suspension point distance. In both cases, maximum equivalent strains are

below 0.2 and are considered as insignificant compared to the elongation at break of the rubber at -40°C, which is 90%.

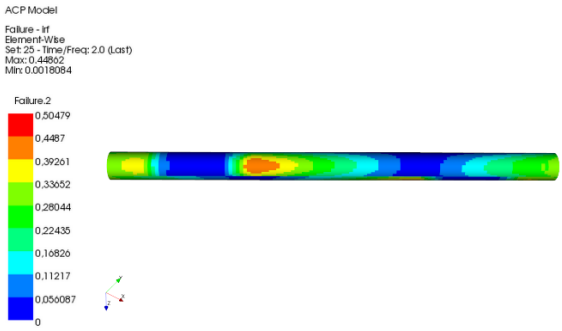


Figure 20. Tsai-Hill failure index distribution at minimum suspension point distance

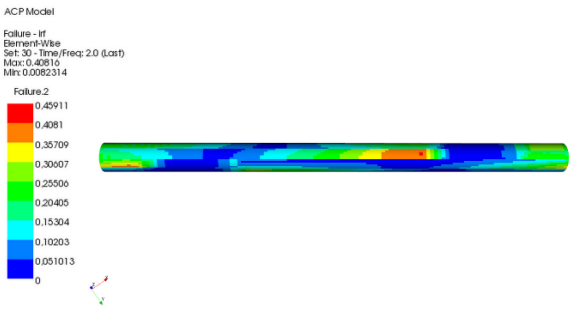


Figure 21. Tsai-Hill failure index distribution at maximum suspension point distance

Maximum failure indices are far below the criteria value of 1 in both minimum (0.45- shown in Figure20) and maximum suspension point distance (0.4- shown in Figure 21), so material failure is not probable in composite layers. These values are considerably lower than maximum values (0.54 and 0.62 respectively) presented in [6].

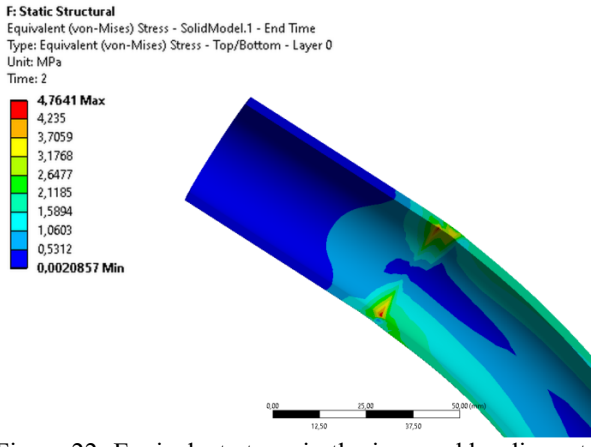


Figure 22. Equivalent stress in the inner rubber liner at minimum suspension point distance

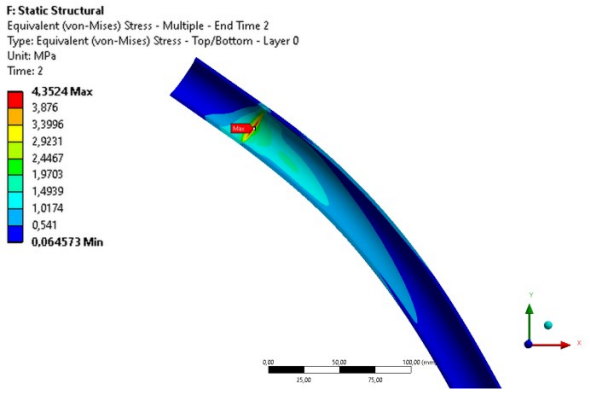


Figure 23. Equivalent stress in the inner rubber liner at maximum suspension point distance

As regards with rubber liners, maximum values arise in the inner rubber liner, at its contact surface with the positioning pin. Maximum values are nearly 4 MPa in both cases being much below the ultimate strength of the rubber at -40 °C (11 MPa).

RESULTS WITH INTERNAL PRESSURE

Tsai-Hill failure indices at minimum suspension point distance can be seen in Figure 24, while failure index values at maximum suspension point distance are shown in Figure 25. Maximum failure index at minimum suspension point distance is not influenced by internal pressure, while at maximum suspension point distance, maxium failure index is higher, although still much lower than the criterial value of 1.

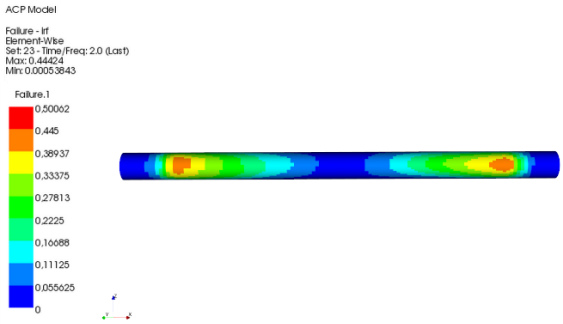


Figure 24. Tsai-Hill failure index distribution at minimum suspension point distance with internal pressure

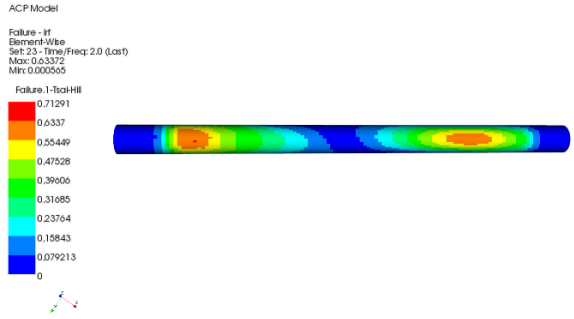


Figure 25. Tsai-Hill failure index distribution at maximum suspension point distance with internal pressure

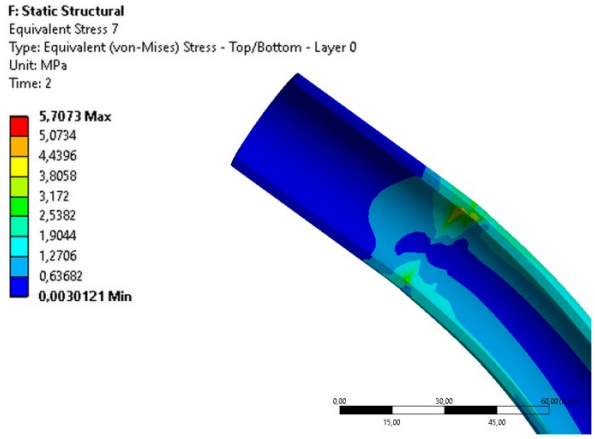


Figure 26. Equivalent stress in the inner rubber liner at minimum suspension point distance

Maximum equivalent stresses at minimum suspension point distance are shown in Figure 26, whereas maximum equivalent stresses at maximum suspension point distance are shown in Figure 27. Maximum stresses are slightly higher with internal pressure although the increase in stresses is not significant.

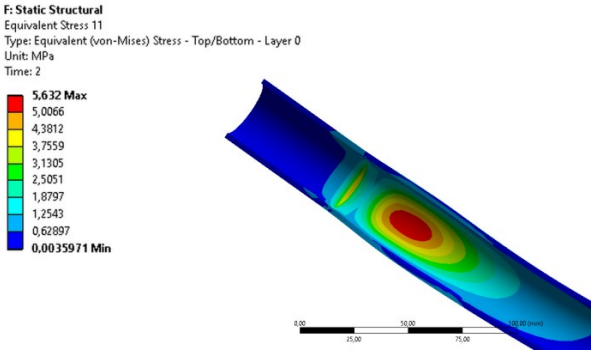


Figure 27. Equivalent stress in the inner rubber liner at maximum suspension point distance

CONCLUSIONS

Extreme operational loads have been considered in the FE model presented in this paper, firstly, because of the large deformations of the tubes as a result of operation on reverse curves and secondly because of stiffer material constants attributed to cold temperatures. Displacments of the suspension points have been determined based on a kinematic simulation validated by a draw and buffing gear interaction test. The utilization of the geometry of the test track led to more realistic carriage positions at minimum and maximum suspension point distance than the positions presented previously in [6]. The existence of a straight section between the curves of opposite curvature makes extrmemums of suspension point distances closer to the initial value (maximum gets lower while minimum gets higher). This also reduces loads arising in the tubes leading to considerably lower maximum stresses and Tsai-Hill failure indices than stresses and Tsai-Hill failure indices reported in [6].

The material properties of the composite layers and the rubber liners have been calculated at -40°C.

In rubber liners, equivalent strains are much below elongation at break (90%), equivalent stress values are also much lower than ultimate tensile strength (11 MPa) and high values are confined to a relatively small zone, so failure is not probable in liners. In composite layers, Tsai-Hill failure indices are also below the criterial value of 1, so failure is not likely in composite layers either.

ACKNOWLEDGEMENT

Authors would like to express their gratitude to Dr. Attila Piros, Department of Machine and Product Design at the Budapest University of Technology and Economics for his valuable help regarding the formulation of the kinematic model.

Authors are also extremely grateful to MÁV Central Rail and Track Inspection Ltd., namely to Csaba Pálfi, for providing them with test report 7420007-19/VÜE-1VJ-1.0-EN.

The research reported in this paper and carried out at BME has been supported by the NRDI Fund (TKP2020 NC, Grant No. BME-NCS) based on the charter of bolster issued by the NRDI Office under the auspices of the Ministry for Innovation and Technology.

REFERENCES

[1] Test report 7420007-19/VÜE-1VJ-1.0-EN, MÁV Central Rail and Track Inspection Ltd.

[2] Mallick, P. K., (1997) Composites Engineering Handbook, CRC Press, , ISBN 9780824793043

[3] Soden PD, Kitching R, Tse PC. (1989) Experimental failure
stresses for ±55° filament wound glass fibre reinforced plastic tubes under biaxial loads Composites; 20 (2): 125–135. DOI: https://doi.org/10.1016/0010-4361(89)90640-X

[4] Railway Engineering: Compound and reverse curve http://www.brainkart.com/article/Railway-Engineering--Compound-and-Reverse-Curve_4228/

[5] UIC 527-1 Coaches vans and wagons-Dimensions of buffer heads- Track layout on S-curves, 2005

[6] Szabó, G and Váradi, K. (2019) FE simulation of a cord-rubber composite tube subjected to bending due to operation on railway track with extremely low curve radius at sub-zero temperature, Proceedings - European Council for Modelling and Simulation, ECMS, 33 (Caserta, Italy,) No. 1, 377-383

[7] Commission Regulation (EU) No. 1302/2014 concerning a technical specification for interopera-bility relating to the 'rolling stock — locomotives and passenger rolling stock' subsystem of the rail system in the European Union (LOC & PAS TSI

[8] UIC 541-1 Regulations concerning the design of break components

[9] UIC 520 -Wagons, coaches and vans-Draw gear-Standardisation

[10] Szabó, G., Váradi, K. and Felhős, D. 2017 Finite Element Model of a Filament-Wound Composite Tube Subjected to

Uniaxial Tension Modern Mechanical Engineering; 7 (4): 91–112. DOI:10.4236/mme.2017.74007
[11] Chawla, K. K. Composite Materials Science and Engineering (3rd Ed.). New York; London: Springer. 2009
[12] Szabó, G. and Váradi, K. 2018 Uniaxial Tension of a Filament-wound Composite Tube at Low Temperature, Acta Technica Jaurinensis, 11 (2), pp. 84-103. doi: 10.14513/actatechjaur.v11.n2.456
[13] Reddy, J.N., Soares C.A.M et al. Mechanics of Composite Materials and Structures, Springer. 1999

GYULA SZABÓ is a PhD Student at the Faculty of Mechanical Engineering, Budapest University of Technology and Economics. His research interests include finite element modelling, composites, particularly cord rubber tubes and diaphragms.
KÁROLY VÁRADI is a professor at the Department of Machine and Product Design, Budapest University of Technology and Economics. His research interests include finite element modelling, structural analyses, composites, fracture mechanics and biomechanics

ANALYSIS OF TIP RELIEF PROFILES FOR INVOLUTE SPUR GEARS

Jakab Molnár*
Attila Csobán
Péter T. Zwierczyk
Department of Machine and Product Design
Faculty of Mechanical Engineering
Budapest University of Technology and Economics
1111, Műegyetem rkp. 3, Budapest, Hungary
E-mail: molnar.jakab@gt3.bme.hu
*Corresponding author

KEYWORDS
involute spur gears, profile modification, tip relief, tooth flank stress, finite element method

ABSTRACT

This research's main goal was to study the influence of involute spur gear tip relief on the contact stress at the engagement meshing point (the beginning point of the line of contact A), as Figure 1. shows.

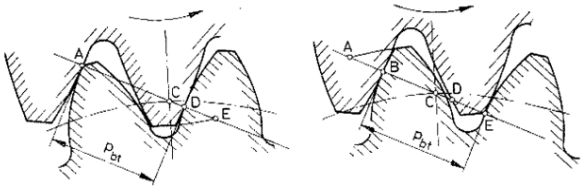


Figure 1. The starting point of single tooth connection (Erney 1983)

Different predefined involute spur gears and modification parameters (amount and length of modification) were already available from previous studies (Schmidt 2019), as Figure 2. shows.

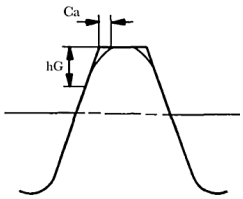


Figure 2. Interpretation of amount of modification and modification length (Schmidt 2019)

In this study, both the drive and the driven gear have tip relief. The modification of the gear profile was achieved through the modification of the gear rack cutter's profile. This way, the gear profiles' profile modification is generated during the gear generation (gear planning) process. The gears have been nitrided, so the heat treatment did not deform the modified gear profile after the gear manufacturing process. The gear modifications were generated in a CAD system, and the calculations were made with FEM. The results show that the tip relief influences the magnitude of the gear contact stress at the first connection point. With the use of tip relief modification, the contact stress of the meshing gears can be reduced at the beginning of the meshing line.

INTRODUCTION

Spur gear drives and transmissions are widely used in mechanical engineering. It is necessary to improve the design life, load capacity of the spur gears, and this way, the transmission's stability. Because of the elastic deformation of gear teeth under heavy load, the base pitch of the drive and driven gear differ from each other. This phenomenon leads to contact shock at the beginning of the meshing, significant fluctuation of the transmission ratio, generation of vibrations and noises, reducing the design life, and the transmission accuracy of the gear drive. With tooth profile modifications, the original true involute profile of the spur gear's teeth was modified by removing material from the potential deformation region of gear teeth.

The tooth profile modification of the involute spur gears can correct the deformation of the gear teeth, thus decreasing the noises and vibrations in the gear drive and the fluctuation of the gear ratio. In this study, we limited our focus just to tip relief modification. Tip relief modification is defined as the material that was removed along the tooth flank with reference to the nominal involute profile at the tip circle. In this study, we used tip relief on both the gear and pinion.
The generating, machining process of the gears was gear planning with MAAG rack type gear-cutter tool (DIN 3972). Profile modification can be achieved through the change of the default gear cutter machine parameters or through the modification of the MAAG gear-cutter tool profile.

METHOD

At the start of our study, both the modification parameters (amount of modification and modification length) and the CAD models of the analysed gear pairs were obtained from previous studies (Schmidt 2019). Only spur gears were analysed with the module of 1 [mm]. The gear pairs have zero backlash and addendum modifications. For the tip relief modifications, Inventor 2018 CAD system, and for the preprocessing and FE studies, ANSYS Workbench 18.2 was used.
The main parameters of the analysed gear pairs can be seen in the following table on the next page (Table 1.).

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

Table 1. Calculated values of
the main geometrical parameters of gear pairs

z_1 [-]	i [-]	z_2 [-]	d_{w1} [mm]	d_{w2} [mm]	a [mm]	b [mm]
17	1	17	17	17	17	10.2
17	4	68	17	68	42.5	10.2
17	6	102	17	102	59.5	10.2
30	1	30	30	30	30	18
30	4	120	30	120	75	18
30	6	180	30	180	105	18
40	1	40	40	40	40	24
40	4	160	40	160	100	24
40	6	240	40	240	140	24

The first goal was to modify the gear profiles with the given modification parameters. As previously mentioned, in this study, we modified the original gear cutter tool profile to achieve tip relief on the gears. The basic idea comes from the paper (Gonzalez et al. 2015) and from the study of the standard DIN ISO 21771. The tool paths of the original rack-cutter tool can be seen in Figure 3.

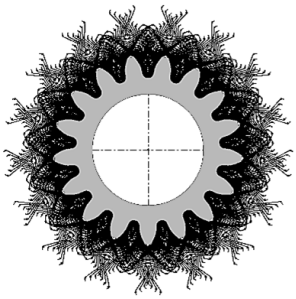


Figure 3. Tool paths of the original gear-cutter tool

The main goal was to make the modification of the rack cutter-tool profile as simple as possible. On the original DIN 3972 rack type gear-cutter profile, the starting point of the gear-cutter tool modification (u) was measured from the tool centerline. From this starting point, the original profile angle (α) was increased on average with 2-3 [°] (α'). This change on the tool profile means that only the required amount of material will be removed at the tip circle. The modification of the tool profile is performed with Electric Discharge Machine (EDM) machine, and because of that, we specified a rounding with the value of 0.3 [mm] at the joining point of the two tool edges with different profile angles. The modified gear cutter tool profile can be seen in Figure 4.

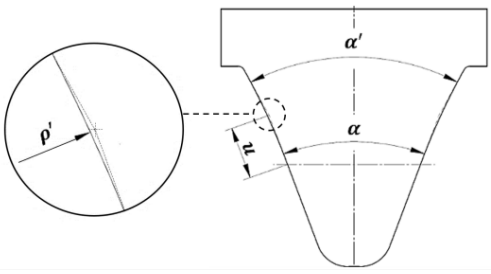


Figure 4. Modified rack-cutter tool profile

Gear profile modification with CAD

Having modified tool profiles, we were able to modify the original gear profiles with CAD. In order to generate a true involute profile in the modification process, we generated tool positions every 0.1 [°]. Where the tool edges crossed, a sketch point was placed in the intersection point.
The intersection points of the tool edges can be seen in Figure 5.

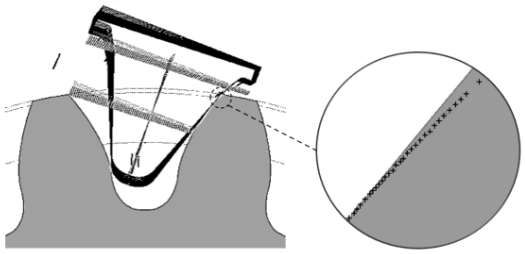


Figure 5. Intersection points of the tool edges

Multiple interpolation splines were placed on the intersection points resulting from the true modified involute profile. The area created by the splines was subtracted from each of the CAD models of the gears. The interpolation splines of the modified involute profile can be seen in Figure 6.

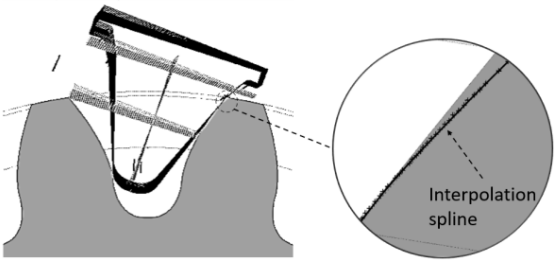


Figure 6. Interpolation splines of the modified profile

As mentioned above, only nitrided gears were analysed, thus the heat treatment process takes place after the machining process of the gears, which already had a modified profile.

Finite Element Model

The goal of the FEM analyses was to determine the changes in the distribution of stress on the contact point of entry's environment on the driving gear while changes were made on driven gears tooth profile, namely the tip relief modification. For the purpose of the analysis, static 2D plane stress was assumed. During the preprocessing part of the analysis, only the midplane surfaces were kept with the original gear thickness. During the studies, the connection of one gear pair was thoroughly analysed, the rest of the gear teeths were kept on each gear to take the stiffening effect of the neighboring teeth into account. In the geometric preparation, each individual tooth was separated from each other with splitting, so multibody parts were created. Because of the need for mesh refinement, the connection region of the analysed gear pairs was separated from the part.

The preprocessing of the CAD models can be seen in figure 7.

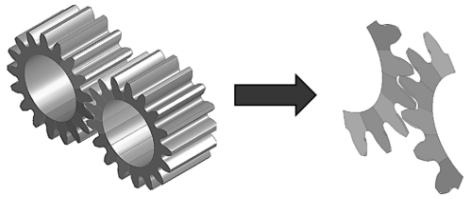


Figure 7. Preprocessing of the CAD models

The applied material for the analysed gears was 16MnCr5 with linear, isotropic, elastic parameters. Thus the linear-elastic study of the stiffness-increasing effect of the heat treatment process was omitted. The material properties can be seen in the following table (Table 2.).

Table 2. Linear, elastic material properties for 16MnCr5

Material property	Value of
Young's Modulus [MPa]	210 000
Poisson's ratio [-]	0.3

Since the friction coefficient is specific to any given gear-pair, a frictionless contact was used between the connecting tooth surfaces in order to generalize the problem definition. The contact between the analysed gear pair was calculated using the Augmented Lagrange method. As previously mentioned, the contact region of the surfaces of the analysed gears was separated, the connection between the teeth body and the meshed region was defined as bonded contact with MPC calculation.
The mesh was constructed using primarily second-order quadrilateral elements, with less than 2 [%] of the mesh consisting of second-order triangular elements. Based on previous studies (Schmidt 2019), the global element size for the gear geometries was selected to 0.1 [mm]. In order to determine the gear root stresses precisely, half of the global element size was set to 0.05 [mm] for the gear roots. In the meshing region of the analysed gear pairs, mesh refinement was used according to the results of evaluated mesh independence studies. The contact region's element size was selected to be 0.006 [mm], two orders of magnitude smaller than the global mesh size to precisely determine the Hertzian contact pressure.
The main parameters of the used FEM mesh are shown in the following table (Table 3.).

Table 3. Main parameters of the FEM mesh

Property	Value
Global element size [mm]	0.1
Element size at the gear root [mm]	0.05
Element size at meshing region [mm]	0.006
Number of nodes [-]	150-300 000
Number of elements[-]	50-100 000
Maximum Aspect ratio [-]	4.3

An example of the structure of the used FE mesh can be seen in Figure 8.

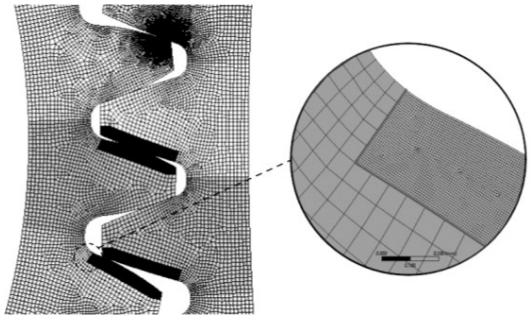


Figure 8. FE mesh

The body of the spur gears was fixed to the z-axis of their coordinate systems as remote points. Only free rotation in Z-axis was allowed. The load of the driving gear was applied to the inner surface of the spur gear. The magnitude of the torque was calculated from the allowable bending stress of the spur gear according to the following equation (Erney 1983):

$$T_{max} = \frac{m \cdot a \cdot b_1 \cdot \sigma_{Flim}}{2 \cdot K_A \cdot 2.7 \cdot 10^6 \cdot (u+1)} \quad (1)$$

, where:
T_{max} [Nm] is the maximum allowable torque
m [mm] is the module
a [mm] is the center distance
b_1 is the width of the driving gear
σ_{Flim} is the allowable bending stress of the spur gear
K_A [-] is the operating factor
u [mm] is the gear ratio

The boundary conditions and the applied load of the gear pairs can be seen in Figure 9.

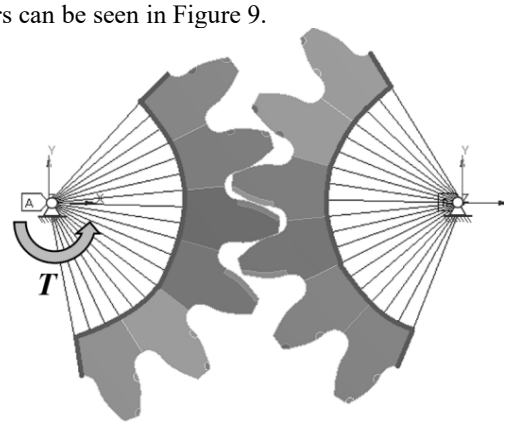


Figure 9. Boundary conditions and applied load

RESULTS

The total displacements showed the expected results, namely, the maximum displacement was located at the tip of the gear tooth where the maximum value was 2 [mm]. The resulted equivalent surface stress field of gear pairs was expected from photoelastic studies. Minimum principal stress was calculated, and the maximum value of the stress was the same as analytically calculated Hertzian pressure.

In Figures 10. and 11., an example for the surface stress field can be seen.

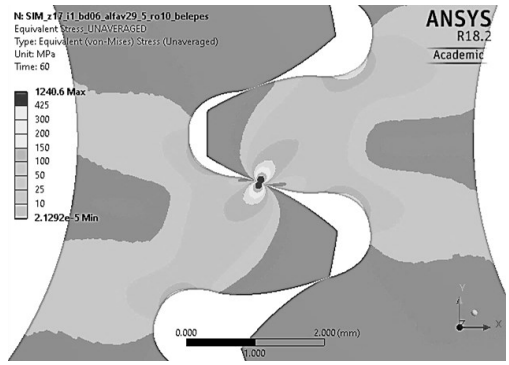


Figure 10. Equivalent stress field of the gear pair

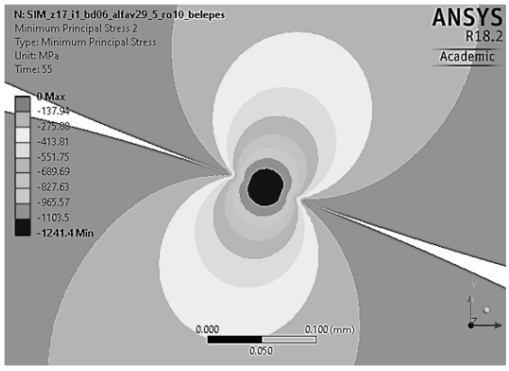


Figure 11. Minimum principal stress field of the gears

Validation

The magnitude of the torque was calculated from the allowable bending stress of the spur gear according to the following equation (Erney 1983):

$$\sigma_{H\_B} = \sqrt{\frac{1}{\pi} \cdot w_n \cdot \left(\frac{1}{\rho_{1,B}} + \frac{1}{\rho_{2,B}}\right) \cdot E_e} \qquad (2)$$

, where:
σ_{H_B} [MPa] is the contact surface stress at the beginning of the single teeth mesh point (point B)
w_n [N/mm] is the line of pressure
E_e [MPa] is the equivalent of Young's modulus
$\rho_{1,B}$ [mm] is the equivalent radius of the driving gear curvature
$\rho_{2,B}$ [mm] is the equivalent radius of the driven gear curvature

In the validation process, the analytically calculated surface contact stress at the beginning of the single teeth meshing (point B, see eq. 2.) were compared to the numerically calculated minimal principal stress (FEM simulations) at the same meshing point. The results show that the difference between the analytically and numerically calculated surface stress at the meshing point B differ by a maximum of 7 [%] from each other. This concludes that the maximum contact stress at the beginning of the gear contact can be validated.

Table 4. Analytically and numerically calculated surface contact stresses

z_1 $[-]$	i $[-]$	z_2 $[-]$	σ_{H-B} $[MPa]$	${\sigma_{H-B}}^{FEM}$ $[MPa]$	${\sigma_{H-A}}^{FEM}$ $[MPa]$
17	1	17	1350	1264	1863
17	4	68	1157	1087	2506
17	6	102	1129	1063	1815
30	1	30	990	933	1892
30	4	120	810	769	1345
30	6	180	786	745	943
40	1	40	855	803	1257
40	4	160	692	654	1366
40	6	240	670	630	1716

Summary

Based on the above-mentioned results, the conclusion was reached that the distribution of load, the noise of the gear system, and the life expectancy of the component are greatly influenced by the modification of the teeth profile. In order to reduce the amount of stress present on the teeth surface at the moment of connection, tip relief is recommended. Of all the different types of tip reliefs that are available, the tip relief modification made by the adjustments of the rack-cutter profile had the most promising results, while the linear profile modification only delayed the emergence of the maximum stress point. The teeth modification made by the linear profile modification does not reduce the stress present on the gear teeth' surface but delays its emergence. With the use of rack-cutter modification, the maximum surface stress (at the gear entry point location *A*) could be reduced up to 50[%], compared to the original shape or tip relief made with chamfer. Increasing the gear ratio reduces the stress at the instantaneous point of contact. By increasing the number of teeth on the driven gear, the teeth profile has a straighter form which causes the growth of the gear root that reduces the amount of stress on the teeth surface.

The comparison of the original gear and chamfer and tool profile modification tip relief can be seen in the following figure (Figure 12.).

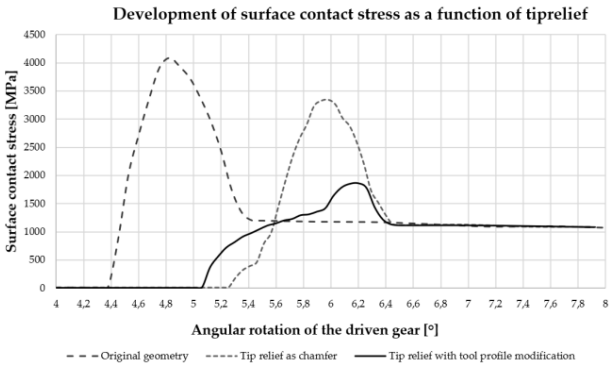


Figure 12. Development of surface contact stress as a function of tip relief

DISCUSSION

The profiles of the tip reliefs examined in our study are partially accurate because of the CAD program. In some cases, a few hundredths of a millimeter difference occurred between the gear profiles' exact positions. The tip relief's spline profile could have been better refined during the analyses, which caused a minimal amount of difference between them. Our examinations would be more precise with further refining of the modification of the tip edge in the simulation program. The amount of time required to run a simulation could be reduced by making submodels that have fewer calculations because of the fewer mesh points. In this study, only connections without backlash were analysed, but it would be recommended to analyze the effect of the gear pair backlash on the stress formation. Thus the stiffness-increasing effect of the heat treatment process and different backlash types should be taken into account.

Due to the short time available for research, only the base problems could be studied. In the future, we would like to continue our research in this field to take a deeper look at the gear optimization possibilities.

ACKNOWLEDGMENT

Hereby, the authors would like to express my thanks to Bence Schmidt, a mechanical engineer and a former student, who contributed hugely to this research by his master's thesis and work.

The research reported in this paper and carried out at BME has been supported by the NRDI Fund (TKP2020 NC, Grant No. BME-NCS) based on the charter of bolster issued by the NRDI Office under the auspices of the Ministry for Innovation and Technology

REFERENCES

György Erney, 1983. , “Fogaskerekek.”, *Műszaki Könyvkiadó*, Budapest, 1983

Li, X., Wang, N., & Lv, Y. 2016. "Tooth Profile Modification and Simulation Analysis of Involute Spur Gear". *International Journal of Simulation Modelling*, vol. 15, pp. 649-662.

Beghini, M., Presicce, F., & Santus, C., 2005. "Proposal for Tip Relief Modification to reduce Noise in Spur Gears and sensitivity to meshing conditions". *Gear Technology,* vol. 23.

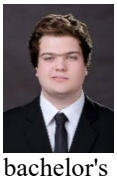
Gonzalez-Perez, I., Roda-Casanova, V. & Fuentes, A., 2015. "Modified geometry of spur gear drives for compensation of shaft deflections". *Meccanica 50*, pp. 1855–1867

DIN ISO 21771:2014-08, Zahnräder - Zylinderräder und Zylinderradpaare mit Evolventenverzahnung - Begriffe und Geometrie (ISO 21771:2007)

DIN 3972:1952-02, Bezugsprofile von Verzahnwerkzeugen für Evolventen-Verzahnungen nach DIN 867

Bence Schmidt, 2019. "Influence of the bearing stiffness on the load distribution of spur gears." *Master's thesis*. (supervisor: Attila, Csobán PhD.)

AUTHOR BIOGRAPHIES



JAKAB MOLNÁR was born in Győr, Hungary, and went to the Budapest University of Technology and Economics, where he studied mechanical engineering and machine design and obtained his bachelor's degree in 2019. He continues his studies at Budapest University of Technology and Economics as a mechanical engineer and machine design master's student. His e-mail address is: molnar.jakab@gt3.bme.hu and his webpage can be found at: http://gt3.bme.hu

ATTILA CSOBÁN Assistant professor at Budapest University of Technology and Economics. Member of the Association of Hungarian Inventors since 2000. Member of the Entrepreneurship Council of the Hungarian Research Student Association since 2006.Member of the public body of the Hungarian Academy of Sciences (MTA) since 2012. Gold level member of the European Who is Who Association since 2013. Research field: gear drives, gearboxes, planetary gear drives, cycloidal drives. His email address is: csoban.attila@gt3.bme.hu, and his webpage can be found at: http://gt3.bme.hu

PÉTER T. ZWIERCZYK is an assistant professor at Budapest University of Technology and Economics Department of Machine and Product Design, where he received his M.Sc. degree and then completed his Ph.D. in mechanical engineering. His main research field is the railway wheel-rail connection. He is a member of the finite element modeling (FEM) research group. His email address is: z.peter@gt3.bme.hu, and his webpage can be found at: http://gt3.bme.hu

IMPLEMENTATION OF BONE GRAFT ADAPTATION'S FE MODEL IN HYPERMESH

Martin O. Dóczi
Péter T. Zwierczyk
Department of Machine and Product Design
Budapest University of Technology and Economics
Műegyetem rkp. 3., Budapest 1111, Hungary
doczi.martin@gt3.bme.hu
z.peter@gt3.bme.hu

Róbert Sződy
Péterfy Hospital
National Institute of Traumatology
Fiumei street 17., Budapest 1081, Hungary
robert.szody@gmail.com

KEYWORDS
Acetabular bone defect, Acetabular cage, Bone graft adaptation, Finite element analysis, HyperMesh

ABSTRACT

Research significance: In the clinical practice, surgeons sometimes must deal with extended bone defects. Among others, bone grafts are used for filling the large absence.
After implantation, the structure of the graft can change, and the graft's load-bearing effect can be significant. This leads to the idea, that during the design of an implant this effect should be taken into account in the finite element simulations.
In this paper, the authors show the implementation of the bone graft adaptation.
Methodology: This programming task was done by using Python, Tcl and the HyperMesh interface. The bone remodeling algorithm and the related parameters were from the literature research. The results are shown with a finite element model prepared for the Optistruct solver, where the geometry models were based on a patient's CT data.
Results: Viewing the bone graft's elemental apparent density, the most loaded areas could be detected.
Conclusion: The model can predict qualitatively the bone graft's change, which can provide additional information for the implant design. Further analyses are required to investigate the sensitivity of the results.

INTRODUCTION

Clinical Overview

Total hip replacement is an effective way in the treating of osteoarthritis. However, after 10-20 years, a revision surgery has to be made, where the damaged prosthesis elements have to be changed.
In some cases the problem is not with the prosthesis, but the patient's bone. Due to some kind of infection or the stress-shielding effect, significant bone degradation can be observed (Figure 1). (Paprosky et al. 1994)

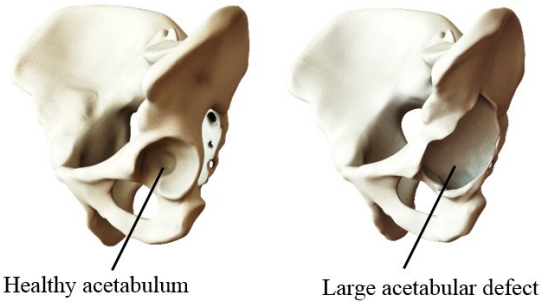


Figure 1. Comparison of healthy acetabulum and large acetabulum defect (Dóczi et al. 2020)

Another example is the tumorous mandible resections, where a large part of the mandible have to be replaced. (Chi Wu et al. 2020)
When the surgeons have to deal with these kind of large bone defects, it is a plausible way using void fillers, which can turn later into living bone tissue. (Sződy et al. 2018), (Ahmad and Schwarzkopf 2015)
It is evident, that in these situations the fixation system can not be so rigid that leads to bone degradation again. However, when the implant is not so stiff, it is usually a weaker construction as well, which can not withstand large forces. This is a trade-off problem.
There is another aspect, which should be considered. If the flexible implant can induce positive bone graft adaptation, it means, that the graft load-bearing effect can be significant later, the overall system can withstand the external forces more easily.
The authors had a different publication (Dóczi et al. 2020) where a simple cantilever beam model and an open-source solver were used to show this algorithm's qualitatively correct behavior. However, in this problem, more complex models required commercial software, which led to further improvements and changes in the process. These will be discussed in this paper.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

Bone Graft Remodeling

Bones and bone grafts can adapt to the loading environment.
There are models, which can describe this phenomenon. One of these claims, that the bone adaptation is related to the strain energy density (SED).
If the given volume part's SED value is divided by the part's density, a so-called stimulus can be obtained. (Chi Wu et al. 2020), (Sue 2016)
The bone's growth response (the density increment for the next step) for this stimulus can be separated into multiple zones, as shown in Figure 2.

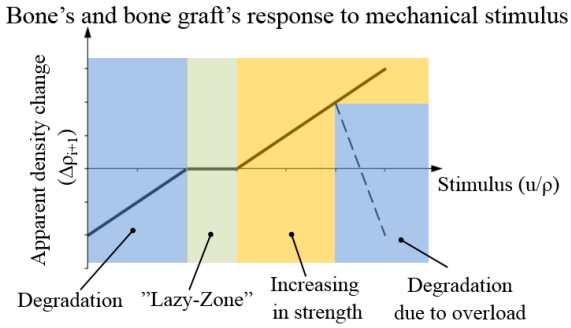


Figure 2. The Bone Graft Remodeling Model

If the mechanical stimulus is low, the bone's apparent density will be decreased, this is the bone degradation. This can happen for large overloads as well. There is a so-called "lazy-zone," a stimulus range, where changes can not be observed in the bone's apparent density. If the stimulus is higher than that, the bone's apparent density is increasing, which means the bone grafts become stiff.
This response is only a model for a pseudo load, which represents a recurring load for a given time period (week, month etc.). The slope of the function has effect of this (pseudo) time. If the slope is too high, this can lead to poor results, and if the slope is too low, the number of the required simulations to show the trends are increasing.
The elastic modulus is the relevant material property for the FE calculations. This can be calculated from an equation by the literature research using the density as shown in Equation (1). E is the elastic modulus, ρ is the density, *b* and *c* are constants from the literature. (Helgason et al. 2008)

$$E = b \cdot \rho^{c} \qquad (1)$$

DATA AND METHOD

Software Environments

Altair HyperMesh is a powerful software for the preprocessing of FE models. Using Tcl, the user can write useful macros for automating tasks. Due to multiple user interfaces, models can be built even for Abaqus, ANSYS etc., solvers.

One of the in-built solvers is the Optistruct, which can be used for FE analysis and optimization as well. The FE input file can be modified as a text file, which can lead to profound customization possibilities. The input files can be separated into different parts, which is helpful during the overwriting because it is not necessary to read and rewrite the entire file.
FE results can be exported as text files as well.
Python is the most popular open-source programing language. Due to the communal improvements, many modules are available. For example, NumPy can be used for manipulating large multidimensional arrays, which is ideal for rewriting FE data.
For the post-processing of the results, pyNastran was utilized.

Implementation of the Graft Remodeling Algorithm

A flowchart can be seen in Figure 3 where the overall implementation is presented.

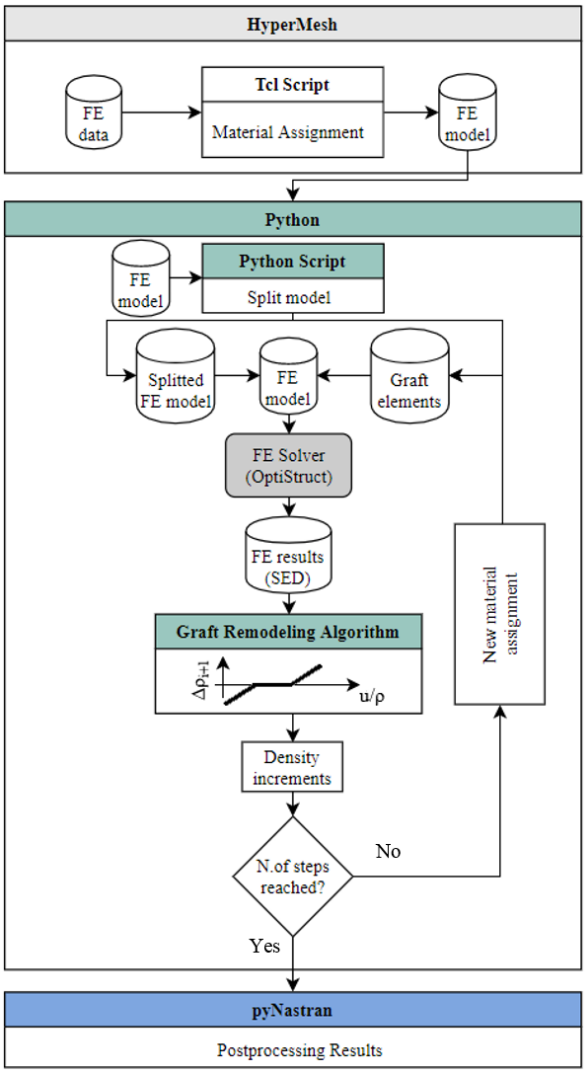


Figure 3. The Flowchart of the Implementation

The workflow is the following. The FE model should be prepared in HyperMesh as usual. A set for the elements of the graft with ID 1 must be created.
For the material assignment, the graft's potential elastic modulus range should be discretized to a large number of ranges; in other words a lot of material data and properties have to enter. Obviously, this should be done by a Tcl script.
After the material assignment, the input file must be saved.
The structure of the saved OptiStruct input file is the following. First, the coordinates of the nodes can be seen. After that, the sets and the elements are presented.
The property ID of an element, which consist the element's material data as well is written after the ID of the element. This number have to be changed during the remodeling process.
In order to make fast changes in the input file, it should be split into parts. The graft data can be separated in another text file with a python script due to the aforementioned set definition.
The required FE results are the strain energy densities of the graft elements. These can be exported to a Nastran ".pch" file, which is easily readable.
The graft remodeling script can calculate the stimulus array from the FE result file and the growth increments of the density for every graft element.
In the next step, the separated input file of the graft elements is rewritten so the elements have new density and elastic modulus value. The input file is solvable again to the pre-defined calculation steps.
The end-results are the graft elements and their density-elastic modulus distribution. For the effective visualization of these plots, a freely available program, pyNastran is used.

Finite Element Model

In this paper, the implementation of the algorithm is the main focus, and the authors investigate the benefits and the disadvantages of the discussed modeling process.
In order to get detailed observations, a pelvis model with a large acetabular defect was used. The geometry model was from a patient's CT data. After the segmentation and the CAD work, a hemipelvis model was generated.
The surgeon prescribed the center of the acetabular cage. (Sződy et al. 2018) The graft's geometry model was mainly in the direction of the maximum amplitude force vector from the gait cycle. This is the most common loading of the implant and the graft. (Bergmann et al. 2001)
The geometry models can be seen in Figure 4.
The FE preprocessing was done in the HyperMesh preprocessing software. HyperMesh. The hemipelvis and the graft were meshed with 10 node tetrahedral elements. Near the acetabular defect, a homogenous bone model was used. In the healthy areas of the pelvis, the material model was separated into a spongiosus and cortical parts. The cortical parts were represented as 6-node triangular shell elements with a 1 mm thickness. (Plessers and Mau 2016)

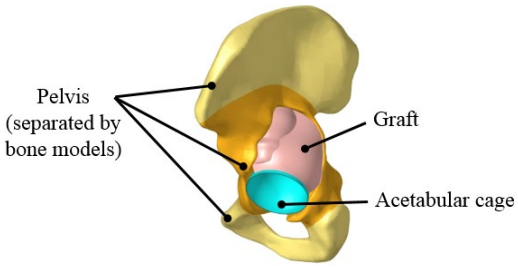


Figure 4. The Parts of the Geometry Model

The acetabular cage has steel properties and it is modeled using 6-node triangular shell elements as well, but with 1.5 mm thickness.
Information about the FE mesh can be seen in Table 1.

Table 1: The Data of the FE Mesh

Number of nodes	210 296
Number of elements	151 139
Number of 10 node tetra elements	140 470
Number of 6 node trias	10669

All of the materials had homogenous, linear elastic, and isotropic properties.
The material properties can be seen in Table 2. The bone's material properties are from the literature research. (Anderson et al. 2005), (Ravera et al. 2016)

Table 2: Elastic Material Properties

	Young's modulus [MPa]	Poisson's ratio [-]
Steel (AISI 316L)	192000	0.3
Cortical bone	17000	0.3
Trabecular bone	100	0.3
Homogenous bone	7000	0.3

In order to investigate the trends, a simplified model was used with bonded connections everywhere.
The modeled acetabular cage has no flange, so it can not connect to the pelvis. After the initial graft density definition, the load was a prescribed displacement by the authors' choice (-0.1 mm; 0.1mm; 0.5mm in the X, Y, Z directions, respectively), at the center of the acetabular cage, transferred with rigid bars. In this simulation, the reaction forces were calculated, the graft's initial density was the minimum density, 382 kg/m^3. The resultant reaction force will be used as a pseudo load, for the graft remodeling calculations. The authors think this approach can be used for eliminating the effect of other irrelevant contacts and the graft's changes can be separately viewed, because from the perspective of the bone graft, just the connecting parts are important.

Fix boundary conditions were used at the sacroiliac joint and the pubic symphysis, according to the literature research. (Plessers and Mau 2016)
The resultant force vector's components and the overall FE model with the boundary conditions can be seen in Figure 5.

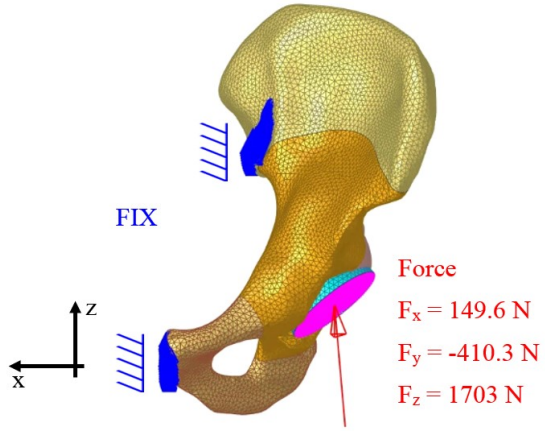


Figure 5. The Finite Element Model

Parameters of the Graft Remodeling Algorithm

The parameters used for the graft remodeling algorithm can be seen in Table 3. The presented graft remodeling parameters only have demonstration goals. Further investigations required to define these numbers.

Table 3: Parameters for the graft remodeling algorithm

	Value	Dimension
Density coefficient (*b*)	1,8	m^2/s^2
Density exponent (*c*)	3	-
"Lazy zone" lower	0,05	m^2/s^2
"Lazy zone" upper	0,1	m^2/s^2
Min density	382	kg/m^3
Max density	2322	kg/m^3
Slope	10	$kg \cdot s^2/m^5$

The bone graft resorption due to possible overload was not examined.
Another value was defined, which name was apparent mass. It is the summarized value of graft element's volumes (V_i) multiplied by their densities (ρ_i). It represents the evolving new bone structure quantitatively, and further comparisons can be made with it. The equation can be seen in Equation (2), where '*i*' is the index of a graft element.

$$m_{app} = \Sigma \rho_i \cdot V_i \qquad (2)$$

Different simulations were made to investigate the effect of the graft's initial density, which means different initial elastic modulus as well. The graft's initial densities were 382 kg/m^3; 500 kg/m^3, and 618 kg/m^3, respectively.
The number of calculation steps was set to 20.

RESULTS

The new density distribution of the graft can be seen in Figure 6. In this model, the graft's initial density was 618 kg/m^3.

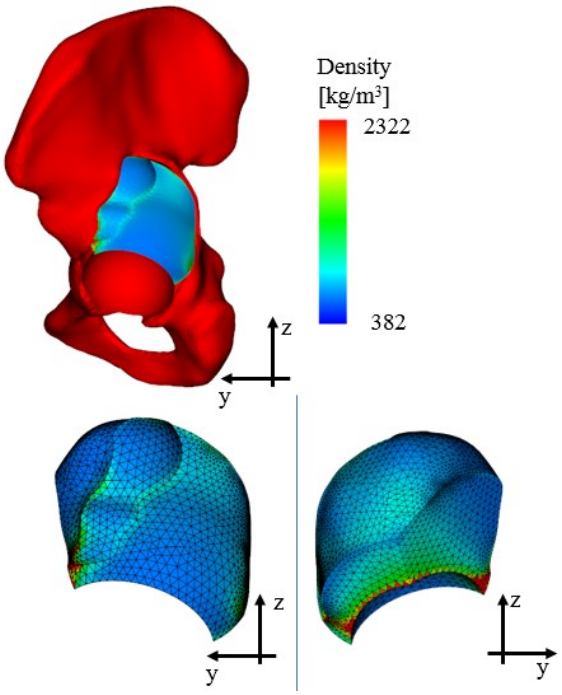


Figure 6. The Density Distribution of the Graft

Using Equation (2), and the apparent mass approach, the different models can be compared. In Figure 7, it can be seen the changing of the apparent mass during calculation steps, with different initial graft's densities.

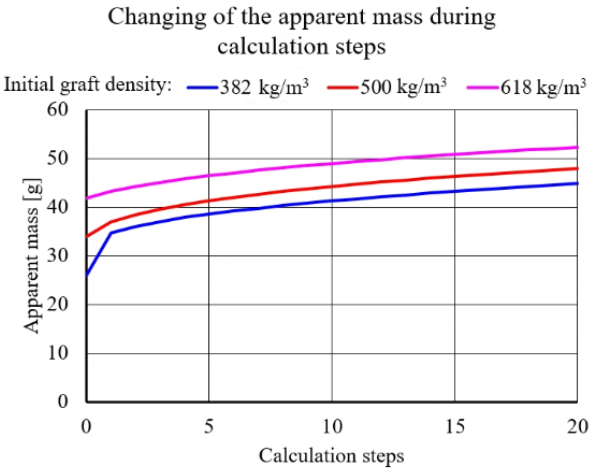


Figure 7. Changing of the Apparent Mass During Calculation Steps

DISCUSSION

The implementation of the graft remodeling algorithm into the HyperMesh-Optistruct interface was successful.
The results in Figure 6 shows that the most loaded areas of the graft, where the elements have the largest density value. These are the areas where shear strain occurs. Further investigations are required to analyze this phenomenon; the authors want to implement the overload-degradation effect into the model, which possible eliminates these results. The other important aspect is the other loaded area on the back of the graft, in the force vector's direction, which is a qualitatively correct result. With the use of second-order elements, the checkerboard problem is not revealed. (Bendsoe 2003), (Rahman et al. 2013)
The results in Figure 7 shows that bigger initial density means a bigger apparent mass. At first glance, it is evident, but it can be seen that the difference between the apparent masses becomes a constant value as the number of the calculation steps is increasing. This information can be used for further implant development because it suggests that the same implant design can produce the same apparent mass growth after a given time, regardless of the graft's initial density and elastic modulus in the investigated range.
Further sensitivity analyses have to be done. Including the effect of the slope of the algorithm, the boundary values of the "lazy-zone", and the effect of the overload and the X,Y,Z values of the prescribed displacement.
The authors want a validation for the model, so actual X-ray images or CT data will be investigated.
It is obvious, that there are many parameters that have effects on the results. However, implant development's main task is to find the trends, which have a significant impact on the design.

ACKNOWLEDGMENT

The research reported in this paper and carried out at BME has been supported by the NRDI Fund (TKP2020 NC, Grant No. BME-NCS) based on the charter of bolster issued by the NRDI Office under the auspices of the Ministry for Innovation and Technology.

REFERENCES

Ahmad, A. and Schwarzkopf, R. 2015. "Clinical evaluation and surgical options in acetabular reconstruction: A literature review." *Journal of Orthopaedics* 12 (2): S238-S243

Anderson, A. et al. 2005. "Subject-Specific Finite Element Model of the Pelvis: Development, Validation and Sensitivity Studies." *Journal of Biomechanical Engineering* 27 (3): 364-373

Bendsoe, M. 2003. "Aspects of topology optimization and bone remodeling schemes." http://biopt.ippt.gov.pl/Minipapers/Bendsoe.pdf 2020.10.15. 15:56

Bergmann, G. et al. 2001. "Hip contact forces and gait patterns from routine activities." *Journal of Biomechanics* 34 (7): 859-891Chi Wu et al. 2020. "Time-dependent topology optimization of bone plates considering bone remodeling."

Chi Wu et al. 2020. "Time-dependent topology optimization of bone plates considering bone remodeling." *Computer Methods in Applied Mechanics and Engineering*. 359: 112702

Dóczi, M., Sződy, R., Zwierczyk, P. 2020. "Finite element modeling of the changing of bone grafts using HyperMesh-Calculix interface." *GÉP* LXXIV: 15-18

Helgason, B. et al. (2008): "Mathematical relationships between bone density and mechanical properties: A literature review". *Clinical Biomechanics* 23 (2): 135-146

Paprosky, W., Perona, P. and Lawrence, J.1994. "Acetabular defect classification and surgical reconstruction in revision arthroplasty: A 6-year follow-up evaluation." *The Journal of Arthroplasty* 9 (1): 33-44

Plessers, K. and Mau, H. 2016. "Stress Analysis of a Burch-Schneider Cage in an Acetabular Bone Defect: A Case Study." Reconstructive review. 6 (1): 37-42

Rahman, K. et al. 2013. "Structural topology optimization method based on bone remodeling." *Applied Mechanics and Materials* 432-426: 1813-1818

Ravera, E. et al. 2015. "Combined finite element and musculoskeletal models for analysis of pelvis throughout the gait cycle." *Conference: 1st Pan-American Congress on Computational Mechanics and XI Argentine Congress on Computational Mechanics*

Sue, A. 2016. Bone remodeling. http://web.aeromech.usyd.edu.au/AMME5981/Course_Documents/files/Lecture%208%20-%20Bone%20Remodelling.pdf 2021.03.17. 12:10

Sződy, R. et al. 2017. (in hungarian) "Csípőprotézis revízióikor alkalmazott „custom made" vápakosár tervezése és készítése, három esetben alkalmazott eljárás." In *7. Hungarian Conference of Biomechanics* (Szeged, HU, okt 6-7) *Biomechanica Hungarica* 10(2): 20

AUTHOR BIOGRAPHIES

MARTIN O. DÓCZI is a Ph.D. student at the Budapest University of Technology and Economics Department of Machine and Product Design, where he studied mechanical engineering and obtained his degree in 2019. His research area is numerical biomechanics and implant development. His e-mail address is: doczi.martin@gt3.bme.hu and his web-page can be found at http://www.gt3.bme.hu.

RÓBERT SZŐDY is a orthopeadic and traumatology physician. He got his degree at the Semmelweis University in 1995. He made a traumatology professional examination in 2000 and an orthopeadics professional examination in 2005. He works as a surgeon at Péterfy Hospital and Manninger Jenő National Institute of Traumatology. His e-mail address is: robert.szody@gmail.com.

PÉTER T. ZWIERCZYK is an assistant professor at Budapest University of Technology and Economics Department of Machine and Product Design where he received his M.Sc. degree and then completed his Ph.D. in mechanical engineering. His main research field is the railway wheel-rail connection. He is a member of the finite element modelling (FEM) research group. His e-mail address is: z.peter@gt3.bme.hu and his web-page can be found at: http://gt3.bme.hu

Simulation and Optimization

Real-time digital twin of research vessel for remote monitoring

Pierre Major, Guoyuan Li, Houxiang Zhang, Hans Petter Hildre
NTNU Ålesund
Largårdsvegen 2,
6025 Ålesund, Norway
{pierre.major, guoyuan.li, hozh, hans.p.hildre}@ntnu.no

KEYWORDS

Virtual Prototyping; Digital Twin; Remote Monitoring

ABSTRACT

Real-time digital twins of ships in operation find many applications such as predictive maintenance, climbing the ladders of ship autonomy, and offshore operational excellence. The literature describes a focus on digital twinning of individual equipment such as navigation, propulsion, engine and power system, or crane. Yet, digital twinning and virtual prototyping for offshore operations are in their infancy and the onboard digitisation hardware and the telecommunication infrastructure are becoming accessible and affordable. Previous work has failed to address the need for building a holistic model and thus contextualising the equipment with the state of the whole vessel. A prototype of an online digital twin of a research vessel is proposed, its architecture described and its suitability for virtual prototyping demonstrated in a remote control centre. The study shows a viable proof of concept for remote monitoring and crew assistance in nominal and contingency response for offshore crane operations.

INTRODUCTION

Offshore operations in wind blown areas such as wind mill parks often involve a lot of downtime for offshore service companies, which have to wait up to 8 weeks at quay to have a proper weather window for installation. The saying is "99 % boredom, 1 % action". To increase the asset utilisation, offshore crews have to optimize installation, maintenance, and decommissioning procedures, test the limits of the system, and design contingency plans. As it is too expensive to be performed with the real assets, the state-of-art is to create digital twins of the system: {ship + equipment + machinery + payload} and use them to simulate the operations in their socio-technical context with hardware-in-the-loop (HIL) and humans-in-the-Loop (HITL), Major et al. [2020]. Digital twins of offshore systems integrate thus physical models of various domains such as the ship's hydrostatics and hydrodynamics, power management systems (PMS), propulsion, ballasting system, dynamic-positioning (DP) system, and machinery such as offshore cranes and winches. Furthermore, the operational procedures to be designed often involve chains, wires, cables, risers and umbilicals. This increases the complexity of the simulation. There is thus a need for integration of multi-domain physics with interaction between rigid bodies and wire-like entities on one side and hydro- and aerodynamics on the other side. Finally, to be useful for hardware integration and human training and design, the performance of the simulation should be real-time or faster, without impairing its fidelity. To respond to these stringent requirements, a modular approach is needed.

As autonomy is gradually becoming a reality for cargo, ferries, and passenger ships, a system of remote monitoring centers will be necessary to watch the remote systems' trajectory, health, and overall functioning. Such an infrastructure is already common in the aerospace industry, with earth crew monitoring the health and activities of space-borne systems 24/7 from launch to decommissioning. Much like air traffic control, vessel traffic service (VTS) centres are a network of onshore based centres monitoring the traffic near the coasts and in vicinity of offshore platforms. The service relies on voice communication and mainly on automatic identification service (IAS) to transmit information mainly limited to navigation and draught and excluding the health of the waterborne systems and their sub-systems. Many research projects are thus tackling the task of building monitoring systems of the remote systems: power, propulsion, ballast, etc. Such an approach allows for predictive maintenance, incident and fault prevention, better fuel consumption through better route planning and less port congestion, and safer offshore operations in an industry where between 75% and 96% of maritime incidents are related to human error All [2019].

This study goes a step further by creating the digital twin of a research vessel, integrating its crane system and transmitting the whole state of the ship via a 4G communication line to an onshore simulator and remote control centre (SRCC). The whole scene is then reconstructed, visualised with a truthful digital twin of the {ship+crane} system, together with a simulation of the system for navigational purposes and a simulation of the crane system.

This paper is organized as follows. Related works are first presented, after which the framework and infrastructure is introduced. The results of a live demonstration are then presented. Finally concluding remarks

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

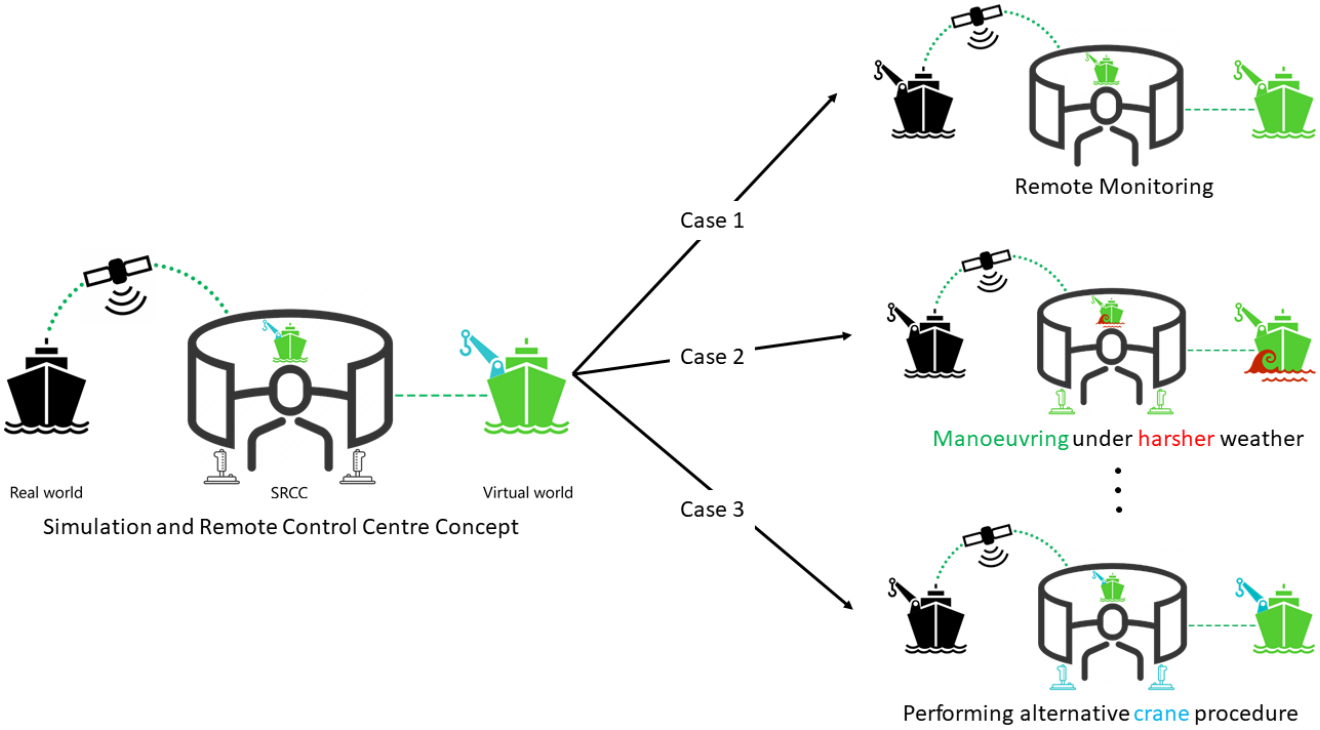


Fig. 1: Digital twin and remote control centre concept, with illustration of cases

and future work are presented.

RELATED WORK

Digital twins are mostly used during at design-time for virutal prototyping. To mention just a few recent publications: Nikolopoulos and Boulougouris [2020] present ship design using an holistic digital twin, Perabo et al. [2020] take profit of the functional mockup interface (FMI) for co-simulation to design and build a testable virtual prototype of a ship with its propulsion system. Likewise Chu et al. [2015] introduces a design system for cranes using FMI. Digital twins find also applications during the operative phase for repeatable operations: Listou Ellefsen et al. [2020] presents an online onboard and onshore fault-prediction and remaining useful life estimation system, Green [2016] showcases an onboard fault prediction maintenance system, finally Coraddu et al. [2019] illustrate the use of data-driven methods for bio-fouling detection and fuel efficiency. Furthermore, Li et al. [2016] present an Agx-based virtual prototyping framework for offshore operations. But in this study, we address unique and non-repeatable operations based on the digital twin of an offshore system. A first of its kind offshore operation was monitored from an onshore remote control center in real-time operation-time via a satellite link, as reported by Time and Torpe [2016]. Underwater Remotely Operated Vehicles (ROVs) operations can not only be performed from the offshore system but also from onshore remote operation centers for ROVs [Oceaneering]. This is case, only the ROV systems are monitored and remotely controlled and not the entire {ship+crane+ROV launcher+ROV} system. Finally, to measure the surrounding state of the ship, Halstensen et al. [2020] illustrates the use of radar-based short term wave prediction for an onboard decision support system using a digital twin of a crane and ship, but without onshore control centre and analysis of scenarios. In this paper we propose a remotely monitored digital twin of the ship and crane systems and illustrate its benefits for advanced offshore operations.

CONCEPT AND ARCHITECTURE

Fig. 2: Crane and Ship Control SRCC

The stretched dome depicted in Figure 2 is one of the SRCCs of NTNU Ålesund research laboratory. Equipped with one crane control chair for commanding a crane with crane joysticks (right on the picture) and one control chair (left on the picture) for controlling the propellers of the ship with maritime lever, it can perform virtual prototyping and remote monitoring of offshore operations, as depicted in Figure 3, where the experimental setup is composed of a sailing ship (left) and

the SRCC (right). The ship's systems are monitored by two onboard management systems, one for navigation information (OLEX server) and for the crane system (MQTT broker). The navigation server gathers data from sensors via signals following the MMEA protocol, which is a text-based low rate protocol, at the rate of 1Hz. The sensed data include global positioning system (GPS), wind speed and direction, and motion reference unit (MRU). The state of the OLEX Server is cloned to an onshore mirror (OLEX Mirror), via a 4G connection and the NMEA signals are interpreted by the OSC Simulator. The simulator can thus reconstruct the current state of the ship's position, orientation and their first and second derivatives (speed and acceleration). A textured and detailed digital elevation model (DEM) of the environment with bathymetry, topography, and built infrastructure is used to contextualise the operation near the shores. The digital twin can thus be placed in the virtual world with the correction position (latitude and longitude) and orientation (roll, pitch, and yaw). Furthermore more, the Navigational Screen (Nav Screen) displays contextualized information such as sea-bottom depth and AIS-based surrounding ship traffic information, and provides even more contextualized remote monitoring information.

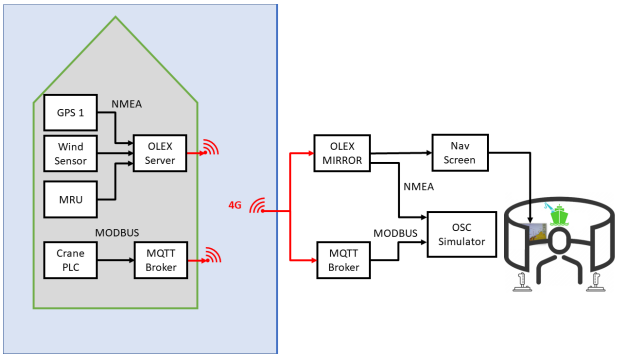


Fig. 3: System Infrastructure

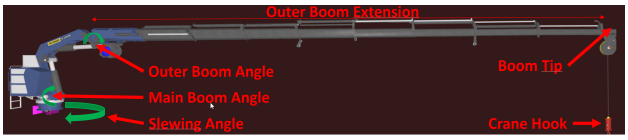


Fig. 4: Palfinger Crane In Simulator

The state of the offshore crane is replicated onshore in a similar fashion via another system and following the Modbus protocol through an MQTT infrastructure mirrored over 4G. The OSC simulator polls the state of the crane at regular intervals (1Hz) and reconstructs the crane in the virtual world based on the slew angle of the crane relative to the ship, the angles of the booms and the extension of the boom tip (in meters), as shown in Figure 4. Figure 1 schematizes the concept: the virtual environment, ship, and crane mirror the real world systems and allows different scenarios. To fully take profit of the simulator centre and simulation engine, it is possible to decouple the visualized models from their real data streams and simulated their behaviours based on physics engines and user control command. Case 2 of Figure 1 illustrates such a case where the position of the virtual crane relative to the ship mirrors the real crane, but the ship responds to harsher environmental conditions (waves, wind, and current), as waves are depicted in red and the ship thrusters are controlled by joysticks (in green). Another possibility is to mirror environment and ship, but control the crane via joystick, as shown in case 3 of Figure 1, with the virtual crane pedestal following the ship movement via mathematical constraints.

The software architecture of the simulation engine (OSC Simulator) is schematized in Figure 5, the data from the real sources or from the mathematical models are fed into an abstraction layer which allows various feeds, with various frequencies and spacial resolutions to be combined into one coherent simulation. Table I summarizes the data source for each case. In case 1 of Figure 1, the visualised data mirrors the offshore ship and crane. In case 2 of Figure 1, the onshore personnel controls the wave height and direction, and the virtual ship behaviour is controlled by a ship engine called FhSim and the handles control the ship's propellers. Finally, in case 3 of Figure 1 the virtual crane is commanded by onshore personnel via crane chair joystick, with the behaviour computed in the physics simulator AgX and the virtual ship truthfully follows the offshore ship.

TABLE I: Case data or physical model source

	Crane	Ship	Environment
Case 1	Real Data	Real Data	Real Data
Case 2	Real Data	FhSim	Instructor
Case 3	AgX Model	Real Data	Real Data

RESULTS

The experiment was performed November 24th 2020, when the RV Gunnerus was stationed in Trondheim Norway and chartered by the Ocean Space Department of NTNU. Figure 7 shows images from case 1: 7 A, is a snapshot of the simulation, 7 B is a live-feed from phone camera, and 7 C is a picture taken in the dome during the experiment, with one of the developers inspecting the crane behaviour and the viewpoint of the simulation taken from a "free-flight" view. If the live-feed was sometimes faster, it experiences more jitter than the digital twin. This seems paradoxical since, as described in the previous section, the data stream for the digital twin goes through more nodes than the video stream (phone to phone) incurring inevitable latency, but the bandwidth usage on the 4G system of the digitized state has a much lower footprint than the video stream. As a matter of fact, parallel channels of a few kbit/s (NMEA and Modbus messages) are used for the digital twin, while the video feed require 100kbit/s to a few Mbit/s on a single channel. Furthermore, once they have reached the onshore simulator centre, the states

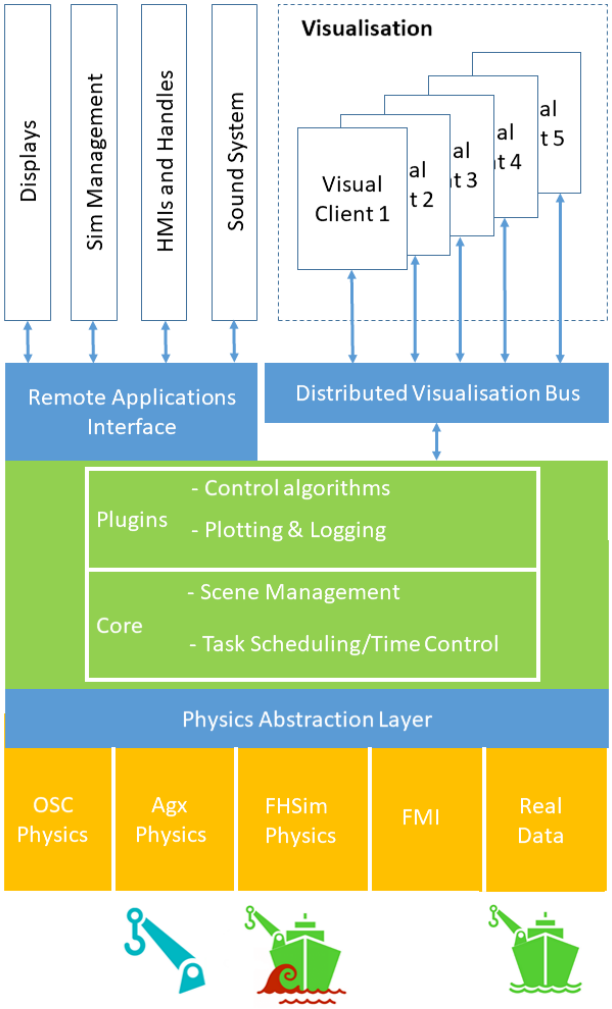


Fig. 5: Architecture and Models for the cases

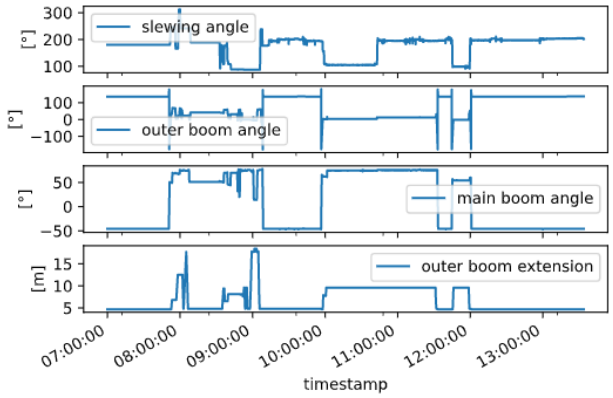


Fig. 6: Experimental data received in real-time

Fig. 7: A) Digital twin viewed from instructor panel, B) Visualisation the a stretched dome of the NTNU Ocean Space Lab, C) Live-feed from the ship during operation

of the ship and crane are filtered in time and space (via physical constraints in Agx) to smooth the visuals. If the few seconds latency are inevitable, the quality of the digital twin visualisation is comparable to the quality of the video feed: it is hard to distinguish the real from the virtual in Figure 7. Furthermore, bandwidth efficiency is an advantage when using satellite links.

Figure 8 shows a map with the scatter plot of the position of the Gunnerus vessel during operations, the color levels correspond to different outer boom extension ranges. The green color denotes the crane in standby, the blue color indicates that outer boom is extended until 10m (mid range) and the red dot corresponds to the peak when crane boom reached its maximum extension 14.8m as show in Figure 6 at 8:00 and 9:00. The ship and crane were both in activity between 10 and 12 (blue line).

The system presented finds many applications. Figure 9 illustrates difference between case 1 and case 3. In case 3, it is possible to run the crane independently and add overlays marking the safe weigh limits. One can see the boom crane of the green ship is higher than the mirror ship. Virtual prototyping applications such as just-in-time operation preparation, tool-box-talk, alternative operational path, and contingency procedures can thus be tested by senior onshore personnel and communicated to the offshore crew. One senior officer could thus stay onshore and be in charge of multiple ships in service. This is both a productivity boost for the service company and an improvement of work life balance of the officer, since she does not have to work many weeks offshore.

As depicted in Figure 10, for case 2 the sea is rougher with higher waves than in the real and mirror case. This allows onshore personnel to test the limits of the equipment and operation and determine the remaining safety margins if the weather was getting worse. This also allows to visualise the effects of performing the operation outside the safety zone such as reaching the safe working load on the crane due to splash zone effect where the immersed crane load in the wave zone is experienced to be much heavier than it own weight due to unfavourable hydrodynamic pressure and rolling of the ship.

CONCLUSIONS

A concept of simulation and remote control centre (SRCC) of {ship + crane } system was demonstrated in three different cases, the experimental setup and ar-

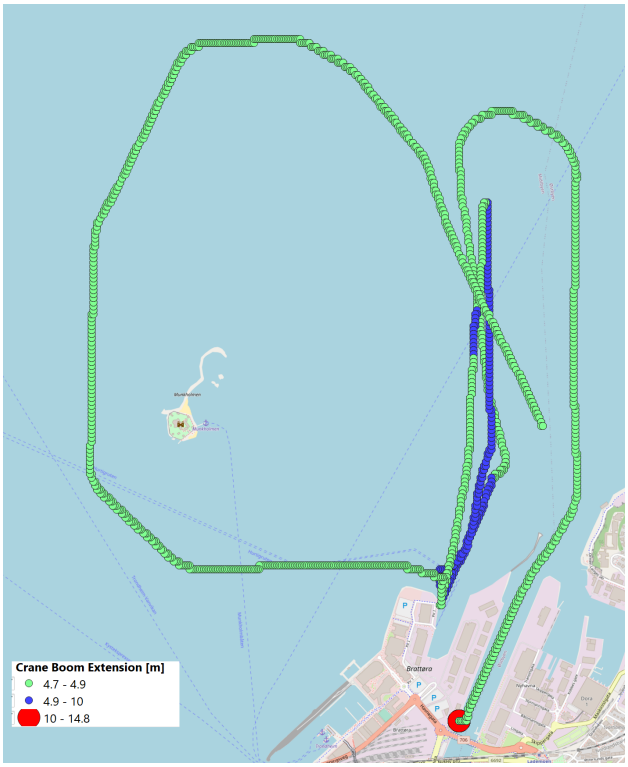


Fig. 8: Geolocalised scatter plot showing crane boom extension during operation. Map credit: Open Street Map

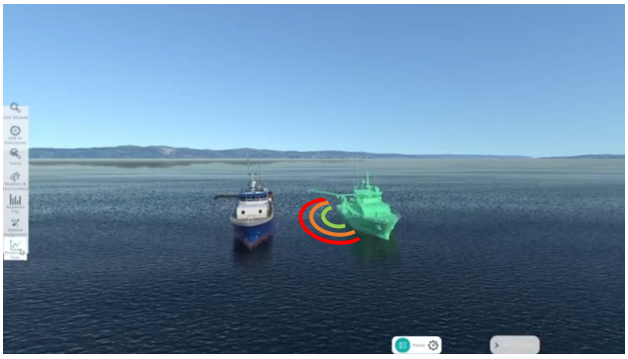


Fig. 9: Mirror digital twin (left), Case 3 (right) with overlayed SWL

chitecture were presented and the results illustrated in form of various visualisations. The main potential applications of such a system are remote monitoring and virtual prototyping aided by augmented reality. The potential can also be further developed by integrating more onboard systems such as propulsion, PMS, and alarm systems.

ACKNOWLEDGMENT

The research presented in this paper is supported by the Norwegian Research Council, industrial PhD. under Grant Number 285949, and Offshore Simulator Centre (OSC). The project is highly related and supported in part by the Project "Remote Control Centre for Autonomous Ship Support", under Grant 309323 from Research Council of Norway. We are also thankful for the logistic help of André Listou Ellefsen from NTNU/DIPAI AS and Finn Tore Holmeset from NTNU. Palfinger kindly contributed to the research by sharing the model of the crane.

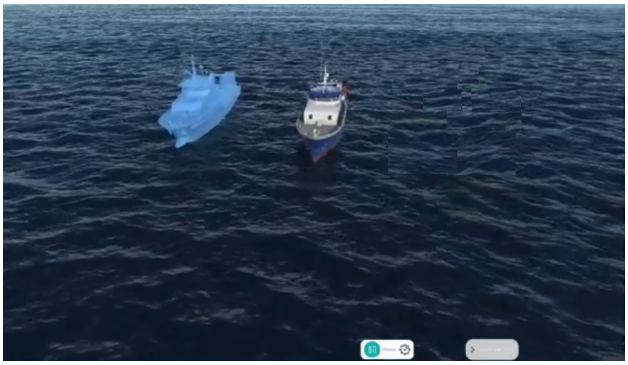


Fig. 10: Case 2 (left) experiencing harsher weather, Case 1 (right) mirroring real weather

References

Onshore Remote Operations Center. URL `https://www.oceaneering.com/rov-services/rov-technology/`.

An annual review of trends and developments in shipping losses and safety. Technical report, Allianz Global Corporate & Specialty 8,862, 2019. URL `https://www.agcs.allianz.com/content/dam/onemarketing/agcs/agcs/reports/AGCS-Safety-Shipping-Review-2019.pdf`.

Y. Chu, L. I. Hatledal, F. Sanfilippo, V. Asoy, H. Zhang, and H. G. Schaathun. Virtual prototyping system for maritime crane design and operation based on functional mock-up interface. In *MTS/IEEE OCEANS 2015 - Genova: Discovering Sustainable Ocean Energy for a New World*. Institute of Electrical and Electronics Engineers Inc., 9 2015. ISBN 9781479987368. doi: 10.1109/OCEANS-Genova.2015.7271342.

A. Coraddu, L. Oneto, F. Baldi, F. Cipollini, M. Atlar, and S. Savio. Data-driven ship digital twin for estimating the speed loss caused by the marine fouling. *Ocean Engineering*, 186, 8 2019. ISSN 00298018. doi: 10.1016/j.oceaneng.2019.05.045.

J. Green. Machina Research Strategy Report The Smart City Playbook: smart, safe, sustainable. Technical report, 2016.

S. O. Halstensen, L. Vasilyev, V. Zinchenko, and Y. Liu. 'Next minutes' ocean waves and vessel motion predictions for more efficient offshore lifting operations. In *SNAME Maritime Convention 2020, SMC 2020*. Society of Naval Architects and Marine Engineers, 2020.

G. Li, P. B. Skogeng, Y. Deng, L. I. Hatledal, and H. Zhang. Towards a virtual prototyping framework for ship maneuvering in offshore operations. In *OCEANS 2016 - Shanghai*. Institute of Electrical and Electronics Engineers Inc., 6 2016. ISBN 9781467397247. doi: 10.1109/OCEANSAP.2016.7485650.

A. Listou Ellefsen, P. Han, X. Cheng, S. Member, F. Tore Holmeset, V. AEsøy, H. Zhang, and S. Member. Online Fault Detection in Autonomous Ferries: Using Fault-type Independent Spectral Anomaly Detection. *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, 2020. doi: 10.1109/TIM.2020.2994012.

P. Major, H. Zhang, H. Petter Hildre, and M. Edet. Virtual prototyping of offshore operations: a review. *Ship Technology Research*, pages 1–18, 10 2020. ISSN 0937-7255. doi: 10.1080/09377255.2020.1831840. URL `https://www.tandfonline.com/doi/full/10.1080/09377255.2020.1831840`.

L. Nikolopoulos and E. Boulougouris. A novel method for the holistic, simulation driven ship design optimization under uncertainty in the big data era. *Ocean Engineering*, 218:107634, 12 2020. ISSN 00298018. doi: 10.1016/j.oceaneng.2020.107634.

F. Perabo, D. Park, M. K. Zadeh, O. Smogeli, and L. Jamt. Digital Twin Modelling of Ship Power and Propulsion Systems: Application of the Open Simulation Platform (OSP). In *IEEE International Symposium on Industrial Electronics*, volume 2020-June, pages 1265–1270. Institute of Electrical and Electronics Engineers Inc., 6 2020. ISBN 9781728156354. doi: 10.1109/ISIE45063.2020.9152218.

N. P. Time and H. Torpe. Subsea compression - Åsgard subsea commissioning, start-up and operational experiences. In *Proceedings of the Annual Offshore Technology Conference*, volume 4, pages 3212–3231, Houston, 2016. Offshore Technology Conference. ISBN 9781510824294. doi: 10.4043/27163-ms.

AUTHOR BIOGRAPHIES

PIERRE MAJOR received his M.Sc. degree in Electrotechnique and Information Technology from the Swiss Federal Institute of Technology of Zürich (ETHZ) in 2005. Industrial PhD. on "data-driven models for fast virtual prototyping" at the Department of Ocean Operations and Civil Engineering, Norwegian University of Science and Technology (NTNU), Ålesund Norway. His domains of interest are virtual prototyping of demanding offshore operations and graphical digital twins of systems such as cities or ships.
Email: pierre.major@ntnu.no

ASSOC. PROF. GUOYUAN LI received the Ph.D. degree in computer science from the Institute of Technical Aspects of Multimodal Systems, Department of Informatics, University of Hamburg, Hamburg, Germany, in 2013. From 2014, he joined the Intelligent Systems Laboratory, Department of Ocean Operations and Civil Engineering, Norwegian University of Science and Technology, Norway. In 2018, he became an Associate Professor of Ship Intelligence. He has published more than 60 papers in the areas of his research interests which include modelling and simulation of ship motion, autonomous navigation, intelligent control, optimization algorithms, and locomotion control of bioinspired robots.

PROF. HOUXIANG ZHANG received the Ph.D. degree in mechanical and electronic engineering, in 2003, and the Habilitation degree in informatics from the University of Hamburg, in February 2011. Since 2004, he has been with the Department of Informatics, Faculty of Mathematics, Informatics and Natural Sciences, Institute of Technical Aspects of Multimodal Systems (TAMS), University of Hamburg, Germany. He joined the NTNU, Norway, in April 2011, where he is currently a Professor of robotics and cybernetics. His research interests lies on two areas: one is on biological robots and modular robotics and the other is on virtual prototyping and maritime mechatronics.

PROF. HANS PETTER HILDRE is professor and head of the Department of Ocean Operations and Civil Engineering at the Norwegian University of Science and Technology (NTNU). His area of interest is product design and system architecture design. Hans-Petter is Centre Director for Centre for Research Driven Innovation (SFI-MOVE) within marine operations. This is cooperation between NTNU, SINTEF, University Sao Paulo and 15 companies at the west coast of Norway. Professor Hildre is head of research in national program Global Centre of Expertise Blue Maritime, project leader in several research projects, member of the board in 5 companies, and has several patents.

COMPARATIVE EVALUATION OF *LACTOBACILLUS PLANTARUM* STRAINS THROUGH MICROBIAL GROWTH KINETICS

Georgi Kostov*, Rositsa Denkova-Kostova**, Vesela Shopska*, Bogdan Goranov***, Zapryana Denkova***
*Department of Wine and Beer ** Department of Biochemistry and Molecular Biology, *** Department of Microbiology
University of Food Technologies, 4002, 26 Maritza Blvd., Plovdiv, Bulgaria
E-mail: george_kostov2@abv.bg; rositsa_denkova@mail.bg; vesi_nevelinova@abv.bg; goranov_chemistry@abv.bg; zdenkova@abv.bg

KEYWORDS

Probiotics, growth kinetics, modeling, optimization

ABSTRACT

The study of the growth kinetics of lactobacilli with pronounced probiotic properties in their batch cultivation is essential. Various models based on the logistic curve model, containing parameters showing the influence of the accumulating lactic acid on the biosynthesis of the product, as well as parameters showing the sensitivity of the cells to lactic acid were used to model the growth kinetics in the present work. The rate constant of adaptation of the studied strains to the used nutrient medium and the induction period were also determined. The kinetics of lactic acid synthesis was determined according to the Weibull model.

INTRODUCTION

The bacteria most commonly included as components of probiotic preparations are lactic acid bacteria (*Lactobacillus* sp., *Enterococcus* sp., *Pediococcus* sp., *Streptococcus* sp., *Lactococcus* sp., *Leuconostoc* sp.) and bifidobacteria. They are also used in the production of probiotic foods (Gibson, 2004), with the largest share being that of the lactobacilli.

Not all species of lactobacilli, as well as not all strains of the same species can be included in the composition of probiotics, but only those that have certain properties (Saarela et al., 2002): to be part of the natural microflora in humans and animals; to be able to adhere and colonize the intestinal mucosa to compete with enteropathogenic bacteria for adhesion sites and nutrients; to survive and maintain their activity in the conditions of the gastrointestinal tract; to be able to reproduce in the gastrointestinal tract; to have high antimicrobial activity in order to suppress and expel pathogenic and toxigenic microorganisms from the biological niche; to allow industrial cultivation - to maintain their activity during production and storage; to modulate the immune response and to be safe for clinical and nutritional use. *Lactobacillus plantarum* is a flexible and versatile species of lactic acid bacteria, which is often found in many probiotic, functional and fermented foods and beverages (cheeses, fermented milk, pasta, sausages and various vegetable juices) or is used as a probiotic (Gobbetti et . al., 1994a; Gobbetti et al., 1994b; Corsetti and Gobbetti, 2002; Guidone et al., 2014). This is due to its flexible metabolism, its ability to adapt to different environmental conditions and the wide range of antimicrobial activity it possesses (Di Cagno et al., 2009).

Along with its antimicrobial activity, the active cells of *L. plantarum* 13M5 have the ability to destroy the mycotoxin patulin at a concentration of 5 mg/dm^3 as a result of the synthesis of a bacteriocin called plantaricin (Todorov et al., 1999; Wei et al., 2020). *Lactobacillus plantarum* YJ7 shows antihyperglycemic potential and reduces insulin resistance, so it can be used in the composition of drugs targeted at people suffering from diabetes (Zhong et al., 2020). In experimental animals, *Lactobacillus plantarum* strains, and in particular *Lactobacillus plantarum* LP33, have been shown to reduce liver damage due to lead intoxication (Hu et al., 2020).

The main metabolite of lactic acid fermentation is lactic acid. It is known that its increasing concentration during fermentation has an inhibitory effect on the growth of the microbial population. The sensitivity to the accumulating lactic acid is strain-specific (Bouguettoucha et al., 2011; Gordeev et al., 2017). The selected mathematical models contain parameters characterizing the influence of lactic acid on lactobacilli. It is also important to determine both the induction period - the time from the lag phase, during which the cells begin to synthesize the necessary cellular structures and enzymes and to move from unadapted to adapted state to the composition of the medium and culture conditions, and the rate constant of adaptation (Warfolomeev and Gurevich, 1999). One of the important conditions for comparing the kinetic characteristics of the models is the initial conditions - inoculum and acidity of the medium – to be the same. Since this is difficult to achieve, it is necessary to measure the data of the biomass and the titratable acidity in the models (Tishin and Fedorov, 2016; Tishin and Golovinskaya, 2015). As a result of the above-mentioned features, the following mathematical models were chosen to model the kinetics of growth and acid formation:

$$\frac{dX_b}{d\tau} = \mu_{max}\left(1 - \frac{P_b}{P_{bm}}\right)^c X_b \quad (1)$$

$$\frac{dX_b}{d\tau} = \mu_{max}\left(1 - \frac{X_b}{X_{bm}}\right)^n X_b \quad (2)$$

$$\frac{dX_b}{d\tau} = \mu_{max}\left(1 - \frac{X_b}{X_{bm}}\right)^{n_1}\left[\left(1 - \frac{P_b}{P_{bm}}\right)^q\right] X_b \quad (3)$$

$$\ln\frac{M}{N_0} = \mu\tau + \ln\left\{\frac{k_0}{k_0+\mu}\left[1 + \frac{\mu}{k_0}e^{[-(k_0+\mu)\tau}\right.\right. \quad (4)$$

$$K_T = a - be^{-(q_p\tau)^\delta} \quad (5)$$

where: μ_{max} - maximum specific growth rate, h^{-1}; X_b, P_b, X_{bm} and P_{bm} are the biomass, the lactic acid amount, the final concentration of the biomass and the lactic acid, respectively, in dimensionless form; M - current

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

biomass concentration, cfu/cm^3; N_0 - initial biomass concentration, cfu/cm^3; τ_a - induction period, h; k_0 - rate constant of cell adaptation to the medium and culturing conditions, h^{-1}; c - a parameter taking into account the inhibitory effect of the accumulating product (lactic acid) on the cell growth; n and n_1 - coefficients taking into account the influence of lactic acid on the cells, respectively showing the resistance (sensitivity) of the cells to the increasing concentration of the product; q - coefficient showing the inhibitory effect of the product, lactic acid, on its own synthesis; K_T - titratable acidity in dimensionless form; *a* - maximum value of the titratable acidity in dimensionless form; *b* - coefficient equal to the difference between the maximum and initial titratable acidity in dimensionless form; q_P - specific rate of acid formation, h^{-1}; δ - an indicator determining the change in the shape of the curve or the change in the rate of accumulation of lactic acid over time; τ - cultivation time, h.

The presented models make it possible to determine the parameters of the fermentation process analytically. Moreover, they allow the assession of the influence of cultivation conditions and the accumulation of lactic acid on the microbial population.

The aim of the present work was to study the lactic acid fermentation process with selected probiotic *Lactobacillus plantarum* strains by applying modified dependences of the logistic curve type and assessing the influence of acid formation on the lactic acid fermentation process.

MATERIALS AND METHODS

Microorganisms and cultivation conditions

The study was conducted with four different strains of *Lactobacillus plantarum*: *Lactobacillus plantarum* 4/17, *Lactobacillus plantarum* 3, *Lactobacillus plantarum* 10 and *Lactobacillus plantarum* 1/18, isolated from spontaneously fermented vegetables. The 4 strains were identified by molecular-genetic identification method – 16S rDNA sequencing – as representatives of the *Lactobacillus plantarum* species (unpublished data).

Cell cultivation was performed under static conditions in flasks using LAPTg10-broth medium. Samples were periodically taken to determine the titratable acidity of the medium and the number of viable lactobacilli cells.

Nutrient media

- LAPTg10-broth;
- MRS-agar;
- Saline solution.

Methods of analysis

- Determination of titratable acidity (ISO/TS 11869:2012);
- Number of viable lactobacilli cells (ISO 7889:2005).

Identification of the model parameters

The logistic curve models from 1 to 3 are solved numerically using the Runge-Kuta method of the 4th row, and the parametric identification in them is performed by minimizing the sum of the squares of the difference between the experimental data and the data obtained from the corresponding model in Microsoft Excel (Choi et al., 2014). The parametric identification of model 4 and the Weibull model was performed using the software Curve Expert Professional by nonlinear regression.

RESULTS AND DISCUSSION

Table 1 presents the data from the determination of the induction period and the rate constant of adaptation, determined according to equation 4.

Table 1: Induction period and rate constant of adaptation

Strain	τ_a, h	k_0, h^{-1}
L. plantarum 4/17	0.36	0.256
L. plantarum 3	0.73	0.253
L. plantarum 10	0.88	0.227
L. plantarum 1/18	1.43	0.103

The strains *L. plantarum* 4/17 and *L. plantarum* 3 have a significantly shorter induction period (0.36 h and 0.73 h, respectively) and higher values of the rate constant of adaptation (0.256 h^{-1} and 0.253 h^{-1}, respectively), compared to the other two strains studied (Table 1). The longest induction period of 1.43 h and the lowest rate constant of adaptation (0.103 h^{-1}) was observed for *L. plantarum* 1/18, while *L. plantarum* 10 occupied an intermediate place with an induction period of 0.88 h and a rate constant of adaptation of 0.227 h^{-1}.

From the studies conducted it can be concluded that *L. plantarum* 4/17 would most quickly adapt to the fermentation medium and cultivation conditions, followed by *L. plantarum* 3 and *L. plantarum* 1/18. This shows that the fermentation medium used for these strains has an optimal composition and is suitable for their growth. The slower adaptation of *L. plantarum* 10 compared to other strains may be due to the lack of some substrates in the medium, values of the redox potential and temperature regime different from the optimal ones for the specific strain studied. Another reason for the lower values of the rate constant of adaptation and the longer induction period is probably the static nature of the medium, which is characterized by a lack of surface aeration, which for some species of *L. plantarum* has certain stimulating effect, as many strains of this species are microaerophiles.

Table 2A presents the kinetic parameters of the used logistic curve models (equation 1 to equation 3). Table 2B presents the correlation coefficients and errors of the models. The comparison of the models with the experimental data is presented in Fig. 1 to Fig. 4. As a general conclusion, the three models used describe the experimental data from the cultivation of the four *L. plantarum* strains with very high accuracy. Similarly, the model used to describe the kinetics of lactic acid accumulation describes very well the experimental data (equation 5). This general conclusion can be explained by the fact that, unlike numerical methods for determining the kinetic parameters in the analytical solution of differential equations, the number of parameters in the model is minimized, and the obtained parameters have a

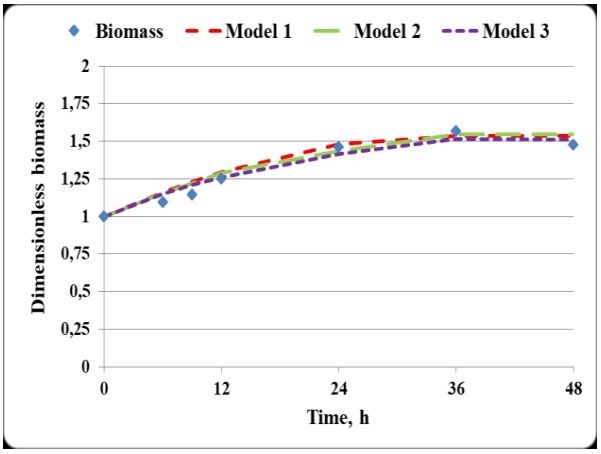
clear biological meaning. It is this biological meaning that we will demonstrate by interpreting the results for the four strains studied. Another reason for increasing the accuracy of the models is the dimensionless form of the biomass, which reduces the identification error. In this way, the influence of dimensionality and the influence of random errors in the enumeration of microorganisms according to the methodology for determining the concentration of viable cells is avoided.

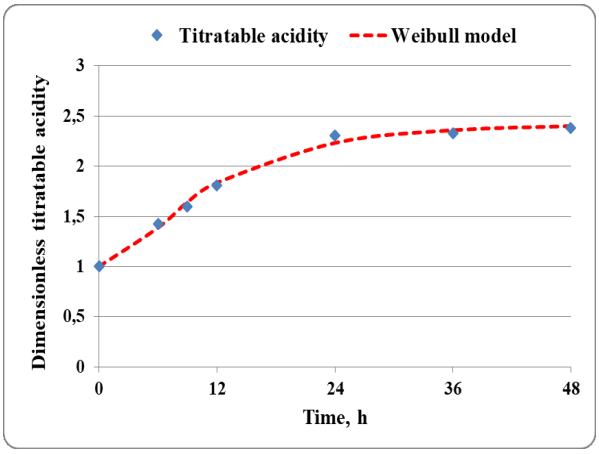
Table 2A: Kinetic parameters in the different logistic curve models in the cultivation of the *Lactobacillus plantarum* strains

Strain	Mathematical model										
	Model 1 (eq.1)			Model 2 (eq.2)			Model 3 (eq.3)				
	μ_m, h^{-1}	c	P_m	μ_m, h^{-1}	n	X_m	μ_m, h^{-1}	n_1	q	P_m	X_m
4/17	0.0198	0.346	1.98	0.0210	3.661	2.41	0.0181	3.208	0.117	2.81	2.30
3	0.0190	0.329	2.48	0.0476	1.703	2.83	0.0383	1.257	0.116	2.78	2.54
10	0.0220	0.368	2.38	0.0740	1.135	2.16	0.0400	0.980	0.130	2.85	2.45
1/18	0.0190	0.448	2.83	0.0156	0.857	1.87	0.0470	0.478	0.621	1.91	2.64

Table 2B: Correlation coefficients and errors

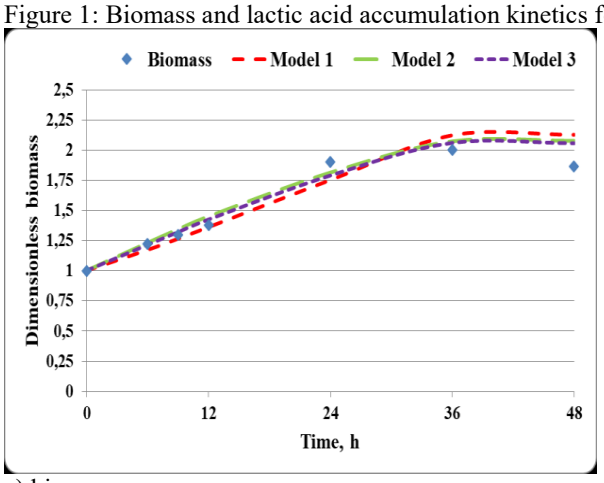
Strain	Mathematical model					
	Model 1 (eq.1)		Model 2 (eq.2)		Model 3 (eq.3)	
	R^2	Error	R^2	Error	R^2	Error
4/17	0.9364	0.078	0.9405	0,078	0.9474	0.074
3	0.9115	0.174	0.9378	0.160	0.9461	0.159
10	0.9022	0.135	0.8200	0.155	0.9406	0.159
1/18	0.8853	0.153	0.9497	0.122	0.9495	0.123

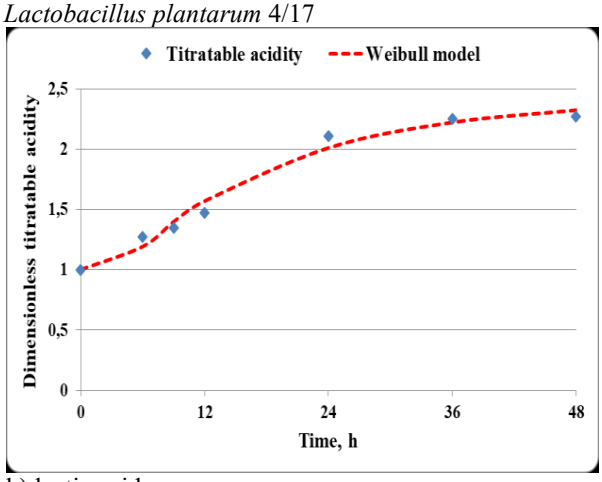




a) biomass b) lactic acid

Figure 1: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 4/17



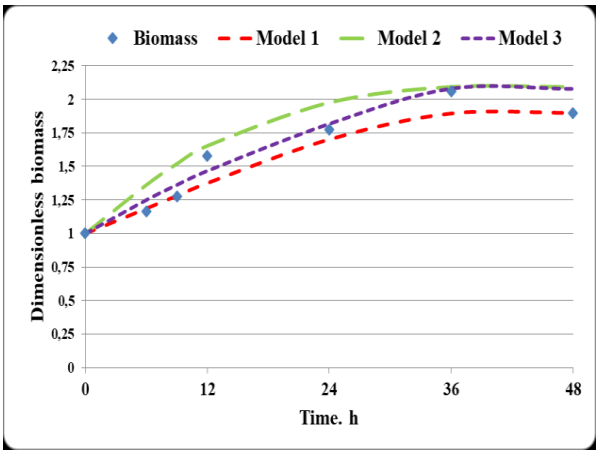


a) biomass b) lactic acid

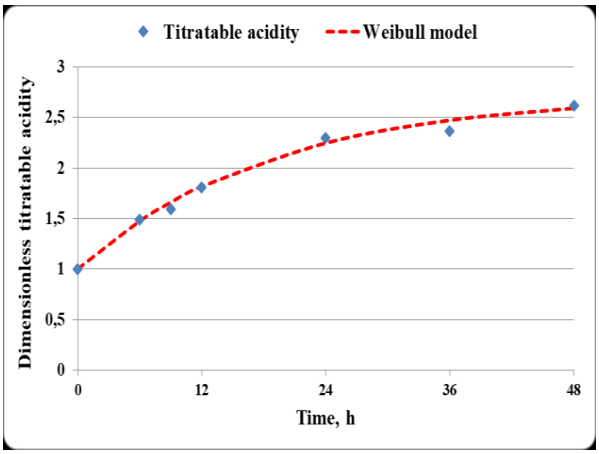
Figure 2: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 3

In all four studied strains the increasing concentration of lactic acid would have a less pronounced inhibitory effect on the maximum specific growth rate (Table 2). This is underlined by the relatively low values of the parameter ***c*** in model 1. In *L. plantarum* 4/17, *L. plantarum* 3 and *L. plantarum* 10 this parameter has comparable values - 0.346, 0.329 and 0.368, respectively. A higher value of the parameter ***c*** is observed in *L. plantarum* 1/18 - 0.448. The observed higher value of the parameter can be explained by the higher concentration of lactic acid accumulated by the strain. This is evidenced by the value of the parameter **Pm**, which is 2.83 and its value is the highest one among the **Pm** values in all the four *L. plantarum* strains studied. The lowest final concentration of lactic acid in dimensionless form is observed for *L. plantarum* 4/17 - 1.98, while for *L. plantarum* 3 and *L. plantarum* 10 it has comparable values - 2.48 and 2.38, respectively. According to the data from model 1, *L. plantarum* 10 has the highest maximum specific growth rate of 0.0220 h^{-1}. The remaining three strains are characterized by lower and commensurable maximum specific growth rates, varying in the range from 0.0190 h^{-1} to 0.0198 h^{-1}.

It is interesting to determine the effect of lactic acid on the cultivation process. Model 2 in which the parameter **n** is subjected to identification is used to achieve this goal. This parameter shows the effect of lactic acid on the biomass, or, more precisely, the resistance of the cells to the accumulating lactic acid. The data are summarized in Table 2.

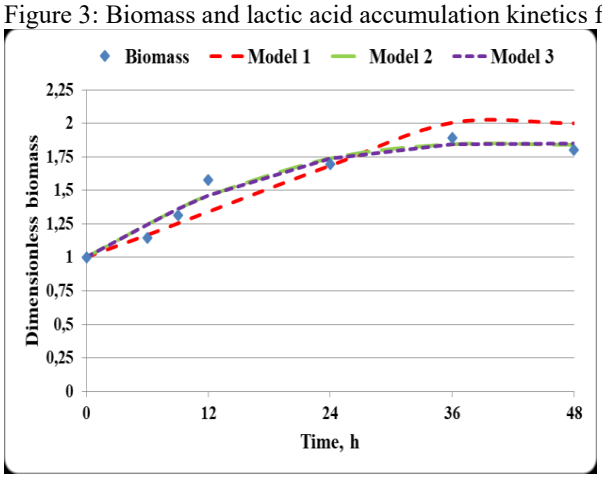


a) biomass

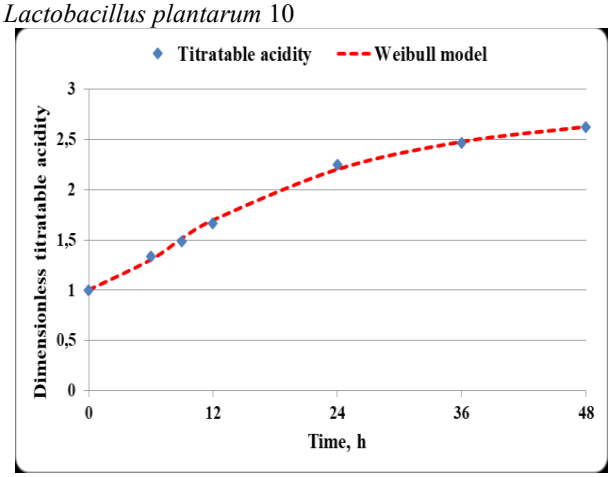


b) lactic acid

Figure 3: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 10



a) biomass



b) lactic acid

Figure 4: Biomass and lactic acid accumulation kinetics for *Lactobacillus plantarum* 1-18

The cells of *L. plantarum* 4/17 (Table 2) show the highest sensitivity, respectively the lowest resistance to the increasing concentration of lactic acid. For this strain the parameter **n** is 3.661. *L. plantarum* 1/18 shows the lowest sensitivity to the increasing concentration of lactic acid (n = 0.857). *L. plantarum* 3 and *L. plantarum* 10 have intermediate resistance to the metabolic product, with the values of **n** being 1,703 and 1,135, respectively. According to model 2, the highest maximum specific growth rate was observed for *L. plantarum* 10 (μ_m = ***0.0740 h^{-1}***), followed by *L. plantarum* 3 (μ_m = ***0.0470 h^{-1}***). *L. plantarum* 4/17 and *L. plantarum* 1/18 have lower values of the maximum specific growth rate - μ_m = ***0.0210 h^{-1}*** and μ_m = ***0.0156 h^{-1}***, respectively.

According to the data from model 2, the highest final concentration of biomass in dimensionless form is achieved by *L. plantarum* 3 - 2.83. In *L. plantarum* 4/17 and *L. plantarum* 10 the maximum concentration of biomass in dimensionless form is 2.41 and 2.16, respectively, and in *L. plantarum* 1/18 it is 1.87.

The meaning of parameter $\mathbf{n_1}$ in Equation 3 is similar. According to this model, *L. plantarum* 4/17 is again characterized by the highest sensitivity to the accumulating lactic acid. In this strain $\mathbf{n_1}$ is 3.208. The next strain in terms of cell sensitivity to the increasing concentration of lactic acid is *L. plantarum* 3, which is also characterized by a high value of the parameter $\mathbf{n_1}$ (1.257). However, in *L. plantarum* 10 model 3 predicted a value of the parameter $\mathbf{n_1}$ below 1, namely 0.980 (the value of the analogous parameter **n** in model 2 is 1.135). This shows that according to model 3, the strain is characterized by increased resistance (reduced sensitivity) to the increasing concentration of lactic acid. The highest resistance to lactic acid is demonstrated by *L. plantarum* 1/18, characterized by the lowest value of the parameter $\mathbf{n_1}$ (0.478) (Table 2A).

The comparative assessment of the resistance of different *Lactobacillus plantarum* strains to lactic acid is important for the characterization of the strains. In general, it can be assumed that strains that have higher resistance to lactic acid would also have a higher survival rate at low pH in the human gastrointestinal tract. This is especially important for the selection of strains, to be included in the composition of probiotic preparations, because resistance to low pH is one of the most important requirements to potentially probiotic strains.

An important parameter in model 3 is the parameter **q**. It reflects the inhibitory effect of lactic acid on its synthesis rate. In *L. plantarum* 4/17, *L. plantarum* 3 and *L. plantarum* 10 this parameter has low values - 0.117, 0.116 and 0.130, respectively (Table 2). This in turn shows high intensity of acid formation and a relatively weak inhibitory effect of the acid on its synthesis rate.

In contrast, in *L. plantarum* 1/18 the value of q is 0.621. This means that the increasing acid concentration would have stronger inhibitory effect on the lactic acid synthesis rate. The high value of **q** characterizes this strain with lower energy of acid formation and therefore the process of acid formation would be more moderate, and the least amount of lactic acid in dimensionless form is accumulated in the medium - 1.91.

The other three strains are characterized by high values of the final concentration of lactic acid in dimensionless form – from 2.78 to 2.85. Unlike model 2, here the values of the biomass in dimensionless form at the end of the fermentation process vary in a relatively small range - from 2.30 to 2.64 (Table 2A).

In order to confirm the assumptions about the acid-forming ability of the studied strains, the specific rate of acid formation q_P and the degree of change in the intensity of lactic acid accumulation over time in general (δ) were calculated. The results of the conducted modeling are presented in Table 3 and Table 4 and the models for the studied strains are shown in real form.

The results in Table 3 once again confirm the conclusions made about the acid-forming ability of the studied strains. According to the Weibull model, *L. plantarum* 4/17, *L. plantarum* 3 and *L. plantarum* 10 have high values of the parameter δ - 2.55, 1.55 and 1.88, respectively. This indicates that in these strains the acid-formation process would proceed with a higher intensity over time in general, compared to *L. plantarum* 1/18. In *L. plantarum* 1/18 δ has a value less than 1, namely 0.99, which again confirms that in this strain the acid formation would be more moderate in time as a whole, although in this strain the Weibull model predicts a slightly higher rate of lactic acid synthesis (0.054 h^{-1}) compared to the other strains. In the other strains, the specific rate of acid formation occupies close values and varies in the range from 0.027 h^{-1} to 0.054 h^{-1} for the different strains.

Table 3: Kinetic parameters in the Weibull model in the cultivation of the *Lactobacillus plantarum* strains

Strain	a	b	q_p, h^{-1}	Δ	R^2	Error
4/17	2.42	1.82	0.037	2.55	0.9933	0.20
3	2.42	1.77	0.041	1.51	0.9803	0.20
10	2.81	2.04	0.027	1.88	0.9981	0.23
1/18	2.72	1.72	0.054	0.99	0.9900	0.22

Table 4: Weibull's mathematical models in real form

Strain	Models in real form
L. plantarum 4/17	$K_T = 2{,}42 - 1{,}82e^{(-0{,}032\tau)^{2{,}55}}$
L. plantarum 3	$K_T = 2{,}42 - 1{,}77e^{(-0{,}047\tau)^{1{,}51}}$
L. plantarum 10	$K_T = 2{,}81 - 2{,}04e^{(-0{,}027\tau)^{1{,}88}}$
L. plantarum 1/18	$K_T = 2{,}72 - 1{,}72e^{(-0{,}054\tau)^{0{,}99}}$

CONCLUSION

Some important conclusions can be drawn for the modeling of the fermentation processes and in particular lactic acid fermentation, from the obtained results. The data show that the use of only one kinetic model does not show all aspects of the lactic acid fermentation process. Combining several mathematical dependencies makes it possible to consider different aspects of the process. For example, equation 4 allows the estimation of the time for adaptation of the culture and the possibility for the real process to start faster. Equations 1 to 3 make it possible to assess the various aspects of the fermentation process - the accumulation of biomass, the influence of lactic

acid, both on the biomass growth and on the acid-formation rate.

The possibility of the models used to assess the sensitivity of the strains to their own metabolic product allows the selection of high-resistant strains to be used in the composition of probiotic preparations, but also the selection of strains that produce less lactic acid and can be used in food development and production.

Therefore, the modeling of the fermentation process must be done with at least two dependencies that reflect the different aspects of the modeled process. Thus, is it possible to achieve a complete interpretation of the various aspects of fermentation. In addition, the dependencies proposed in the present paper allow the estimation of kinetic parameters to be done through simple analytical dependencies. This allows for faster process management decisions.

The main purpose of the present study was to allow the evaluation of different strains of lactic acid bacteria with a view to their use in the production of different types of functional foods. Knowledge of the fermentation kinetics and the behavior of the strains under different cultivation conditions makes it possible to model the fermentation process, and hence the composition of the obtained functional foods.

In this regard, the results allow the strains to be divided into two groups - strains with high growth rate (strain 10), strains with moderate growth rate and high rate of acid formation (strain 1/18) and strains with moderate growth rate and moderate acid formation (strains 4/17 and 3). Depending on the specific food production, the choice may fall on different groups of strains. In some cases, the functional characteristics of a specific food product are determined by the high concentration of viable cells, while in other cases - by the lactic acid produced by the lactic acid bacteria strains and, hence, accumulated in the food product. In this sense, the combination of different models to describe the kinetics of microbial growth allows for improved options for selection and management of the process of functional food production.

REFERENCES

Bouguettoucha, A., B. Balannec and A. Amrane. 2011. "Unstructured Models for Lactic Acid Fermentation: A Review." *Food Technol. Biotechnol.*, 49 (1), 3–12.

Choi, M., M. Saeed Al-Zahrani, and S. Y. Lee. 2014. "Kinetic model-based feed-forward controlled fed-batch fermentation of *Lactobacillus rhamnosus* for the production of lactic acid from Arabic date juice." *Bioprocess Biosyst Eng*., 37, 1007–1015. https://doi.org/10.1007/s00449-013-1071-7

Corsetti, A. and M. Gobbetti. 2002. "*Lactobacillus plantarum*". In *Encyclopedia of dairy sciences* (H. Proginsli, J.W. Fuquay, and P.F. Fox Eds.), New York: Academic Press Ltd., 1501-1507.

Di Cagno, R., M. De Angelis, R. Coda, F. Minervini, and M. Gobbetti. 2009. "Molecular adaptation of sourdough *Lactobacillus plantarum* DC400 under co-cultivation with other lactobacilli." *Research in Microbiology,* 160, 358-366. https://doi.org/10.1016/j.resmic.2009.04.006

Gibson, G. R. 2004. "From probiotics to prebiotics and a healthy digestive system". *J. Food Science*, 69 (5), M141- M143. https://doi.org/10.1111/j.1365-2621.2004.tb10724.x

Gobbetti, M., A. Corsetti, and J. Rossi. 1994b. "The sourdough microflora, evolution of soluble carbohydrates during the sourdough fermentation." *Microbiologie Aliments Nutrition*, 12, 9–15.

Gobbetti, M., A. Corsetti, J. Rossi, F. La Rosa, and M. De Vincenzi. 1994a. "Identification and clustering of lactic acid bacteria and yeasts from wheat sourdoughs of central Italy." *Ital J Food Sci*, 1, 85–93.

Gordeev, L., A. Koznov, A. Skichko, and Y. Gordeeva. 2017. "Unstructured mathematical models of the lactic acid biosynthesis kinetics: A Review." *Theoretical Foundations of Chemical Engineering*, 51 (2), 175-190.

Guidone, A., T. Zotta, R. P. Ross, C. Stanton, M. Rea, E. Parente, and A. Ricciardi. 2014. "Functional properties of *Lactobacillus plantarum* strains: A multivariate screening study." *LWT - Food Science and Technology*, 56, 69-76. https://doi.org/10.1016/j.lwt.2013.10.036

Hu, T., J. Song, W. Zeng, J. Li, H. Wang, Y. Zhang, and H. Suo. 2020. "*Lactobacillus plantarum* LP33 attenuates Pb-induced hepatic injury in rats by reducing oxidative stress and inflammation and promoting Pb excretion." *Food and Chemical Toxicology*, 143. Paper ID: 111533. https://doi.org/10.1016/j.fct.2020.111533

ISO/TS 11869:2012. Fermented milks — Determination of titratable acidity — Potentiometric method

ISO 7889:2005. Yogurt — Enumeration of characteristic microorganisms — Colony-count technique at 37 degrees C

Saarela, M., L. Zahteenmaki, R. Crittenden, S. Salminen, and T. Mattila-Sandholm. 2002. "Gut bacteria and health foods – the European perspective." *Int. J. Food Microbiol*. 78, 99-117.

Tishin, V. B., and A. V. Fedorov. 2016. The peculiarities of mathematical modelling for the kinetics of microorganisms' cultivation." *Processes and Food Production Equipment*, 9 (4), 65-74. (in Russian).

Tishin, V. B., and O. V. Golovinskaia. 2015. *Experiment search and mathematical models of the kinetics of biological processes. Textbook*. St. Petersburg, University ITMO Publ., p. 111. (in Russian)

Todorov, S., B. Onno, O. Sorokine, J. M. Chobert, I. Ivanova, X. Dousset. 1999. "Detection and characterization of a novel antibacterial substance produced by *Lactobacillus plantarum* ST31 isolated from sourdough." *Int. J. Food Microbiol*., 48, 167–177. https://doi.org/10.1016/S0168-1605(99)00048-3

Warpholomeew, S. D and K. G. Gurevich. 1999. *Biokinetic –practical course*", Fair-Press, Moscow, ISBN: 5-8183-0050-1, p.720. (in Russian).

Wei, C., L. Yu, N. Qiao, S. Wang, F. Tian, J. Zhao, H. Zhang, Q. Zhai, and W. Chen. 2020. "The characteristics of patulin detoxification by *Lactobacillus plantarum* 13M5." *Food and Chemical Toxicology*, 146, Paper ID 111787. https://doi.org/10.1016/j.fct.2020.111787

Yoha, K.S., J. A. Moses, and C. Anandharamakrishnan. 2020. **"**Effect of encapsulation methods on the physicochemical properties and the stability of *Lactobacillus plantarum* (NCIM 2083) in synbiotic powders and *in-vitro* digestion conditions." *J. Food Eng.*, 283, Paper ID: 110033. https://doi.org/10.1016/j.jfoodeng.2020.110033

Zhong, H. Abdullah, Y. Zhang, M. Zhao, J. Zhang, H. Zhang, Y. Xi, H. Cai, and F. Feng. 2020. **"**Screening of novel potential antidiabetic *Lactobacillus plantarum* strains based on *in vitro* and *in vivo* investigations." *LWT – Food science and technology*, 139, Paper ID: 110526. https://doi.org/10.1016/j.lwt.2020.110526

ACKNOWLEDGEMENTS

This work were supported by the Bulgarian Ministry of Education and Science under the National Research Programme "Healthy Foods for a Strong Bio-Economy and Quality of Life" approved by DCM № 577/17.08.2018 and by the project "*Strengthening the research excellence and innovation capacity of University of Food Technologies - Plovdiv, through the sustainable development of tailor-made food systems with programmable properties*", part of the European Scientific Networks National Programme funded by the Ministry of Education and Science of the Republic of Bulgaria (agreement № Д01-288/07.10.2020).

AUTHOR BIOGRAPHIES

GEORGI KOSTOV is Professor at the Department of Wine and Beer Technology at the University of Food Technologies, Plovdiv. He received his MSc degree in Mechanical Engineering in 2007, a PhD degreein Mechanical Engineering in the Food and Flavor Industry (Technological Equipment in the Biotechnology Industry) in 2007 at the University of Food Technologies, Plovdiv, and holds a DSc degree in Intensification of Fermentation Processes with Immobilized Biocatalysts. His research interests are in the area of bioreactor construction, biotechnology, microbial population investigation and modeling, hydrodynamics and mass transfer problems, fermentation kinetics, and beer production.

VESELA SHOPSKA is Head Assistant Professor at the Department of Wine and Beer Technology at the University of Food Technologies, Plovdiv. She received her MSc degree in Wine-making and Brewing Technology in 2006 at the 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-KOSTOVA is Head Assistant Professor at the Department of Biochemistry and Molecular Biology at the University of Food Technologies, Plovdiv. She received her MSc degree in Industrial Biotechnologies in 2011 and a PhD degree in Biotechnology (Technology of Biologically Active Substances) in 2014. Her research interests are in the area of isolation, biochemical and molecular-genetic identification and selection of probiotic strains and development of starters for functional foods.

BOGDAN GORANOV is an assistant at the department of Microbiology at the University of Food Technologies, Plovdiv. He received his PhD in 2015 from the 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 the Food Industry". His research interests are in the area of bioreactor construction, biotechnology, microbial population investigation and modeling, hydrodynamics and mass transfer problems, and fermentation kinetics.

ZAPRYANA DENKOVA is a professor at the department of Microbiology at the 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.

Using Semantic Technology to Model Persona for Adaptable Agents

Johannes Nguyen, Thomas Farrenkopf,
Michael Guckert
Kompetenzzentrum für Informationstechnologie
Technische Hochschule Mittelhessen
61169, Friedberg, Germany
{johannes.nguyen,thomas.farrenkopf,
michael.guckert}@mnd.thm.de

Simon T. Powers, Neil Urquhart
School of Computing
Edinburgh Napier University
EH10 5DT, Edinburgh, United Kingdom
{s.powers,n.urquhart}@napier.ac.uk

KEYWORDS

Adaptable Agents, Persona, Agent Modelling, Semantic Technologies

ABSTRACT

In state of the art research a growing interest in the application of agent models for the simulation of road traffic can be observed. Software agents are particularly suitable for the representation of travellers and their goal-oriented behaviour. Although numerous applications based on these types of models are already available, the options for modelling and calibration of the agents as goal-oriented individuals are either simplified to aggregated parameters or associated with overly complex and opaque implementation details. This makes it difficult to reuse available simulation models. In this paper, we demonstrate how the combination of persona models together with semantic methods can be applied to achieve a well-structured agent model that allows for improved reusability.

INTRODUCTION

Computer-based simulation is an accepted means for researching transportation questions, which has been used as early as the 1970s [1], [2]. The number of existing simulators is significant, with each of the tools focusing on different aspects of the transport system and differing in the underlying methods. There is a variety of simulators that range from more general purpose applications (e.g. [3], [4], [5], [6], [7]) to systems designed for specific research questions (e.g. [8], [9], [10]). In practical research on transportation, researchers are faced with the issue of finding appropriate simulators. [7] have described that even though general purpose applications such as MATSim [5] and SUMO [4] offer a lot of potential for reusability and sharing of common traffic concepts (e.g. modelling of road network, vehicles, traffic flow), in many cases researchers have instead implemented their own simulation models from scratch. A reason for this may be that customisation options in available simulators are either too limited or too complex to be implemented. This is the case when customisation requires advanced programming or a deep understanding of the underlying system. A structured design with a clear separation of concerns (see [11]) for modelling software agents using persona models and semantic methods can help to improve reusability of simulation models and reduce complexity for customisation.

This paper is organised as follows: The following sections provides a short introduction into the theoretical background of persona models, which are usually applied to areas in which focus lies on user-centricity such as *Human–Computer interaction (HCI)* or *marketing*. Furthermore, an overview of the semantic instruments used in this work is given, namely ontologies implemented in OWL (Web Ontology Language) [12] and SWRL (Semantic Web Rule Language)[13]. Following this, we discuss related work. We then present a modelling method that allows for less complex customisation using the concepts of persona models and ontologies. As proof of concept, we perform simulation of two example scenarios using the AGADE Traffic simulator [14]. The scenarios fundamentally differ in types of mobility, which is often the case when specific research questions at hand deviate from the main focus of available simulators. Thus, we demonstrate how customisation or extensions to the model can be implemented with the proposed method. Finally, summary and conclusions are given as well as indications for future work.

PRELIMINARIES

The following section briefly introduces background knowledge on the concept of persona models and semantic methods.

Persona Models

Persona models are an instrument for analysing and modelling groups of individuals sharing similar behaviour. They are often applied in the field of Human–Computer interaction (HCI) and for marketing purposes. In practical applications persona are usually created with segmentation or clustering methods based on collected customer or user data. [15] has discussed the origins of persona models as an approach to goal-oriented software design. Reference is given to Cooper's definition of persona models as *"a precise description*

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

of [a] user and what he wishes to accomplish" (see [16], p.123). A more detailed description is given by [17] who describe persona as *"fictional, detailed archetypical characters that represent distinct groupings of behaviours, goals and motivations observed and identified during the research phase"*. It can be summarised that persona are fictional characters representing groups of individuals. They are identified by a unique name and carry additional descriptive information of relevance for the perspective that is to be modelled, e.g. appearance, private background, preferences, habits and goals in order to make a group of individuals more comprehensible and manageable and to convey their personality and motivations.

Ontologies and Rules

Ontologies are an expressive tool to model a domain in machine readable form and provide an explicit, shared specification of a conceptualisation [18]. Ontologies typically consist of a taxonomy of concepts each with properties and relations. *OWL (Web Ontology Language)* is a standardised implementation of a description logic based ontology language [12]. As description logic is object centered, formulation of simple if-then rules is limited. These rules can be expressed using Semantic Web Rule Language (SWRL). SWRL is also standardised by W3C. Inference engines derive knowledge by evaluating OWL and SWRL expressions.

RELATED WORK

We have reviewed a wide range of available traffic simulators in detail (*inter alia* [3], [5], [19], [8], [20], [6], [7], [21], [22], [23]). In particular, the AgentPolis approach stands out as it also reflects on the shortcomings of reusability in available simulation models. AgentPolis is a fully agent based traffic simulator that focuses on the simulation of interaction-rich transport scenarios [24], [7]. For example, simulation of on-demand mobility services (e.g ridesharing) requires interaction between service providers and customers but numerous other forms of interaction between travellers are possible. Despite the fact that general purpose traffic simulators such as MATSim and SUMO provide a variety of modelling concepts (e.g. road network, vehicles, traffic flow), the authors of AgentPolis identified the gap of transport scenarios with significant interaction between travellers and their immediate surrounding. The authors of AgentPolis concluded that similarities between simulation models have not been exploited sufficiently due to existing tools not taking into account the multi-agent nature of interaction-rich transport systems. Reference is given to work in which model-specific simulation tools have been developed from scratch (see [10], [25], [9]). AgentPolis addresses these deficiencies and provides a set of abstractions, code libraries and software tools for building simulation models [7]. While focus of the project was on the modelling of interaction-rich transport systems, a technical solution has also been implemented to facilitate the reuse of common transportation concepts. For this purpose, AgentPolis integrates a *modelling abstraction ontology*. The theoretical concept of this component is to separate defined modelling abstractions from implementations of specific modelling elements. It uses an ontology in order to define more general concepts of multi-agent systems. This approach results in a tailored structure for object-oriented programming that simplifies extending the simulation models for research-specific scenarios.
In this paper, we revisit this idea of reusable modelling concepts using ontologies in traffic simulations and further expand on the modelling capabilities of semantic methods. Furthermore, we will go one step further and place the individual and its decision-making at the center of attention in our modelling rather than solely defining general modelling abstractions in the ontology for common transportation concepts such as traffic lights, etc.

MODELLING

The application of agent-based models for simulating road traffic is an established method. Traffic is an emergent phenomenon in which global system behaviour is determined by a large set of individuals, each with their own goals and preferences. As [26] describe, software agents are *closed computer systems that are situated in some environment, and that are capable of autonomous action in this environment in order to meet their designed objectives*. This autonomous and goal-oriented behaviour also applies to travellers in the real world which is why software agents are particularly suitable for representing travellers in computer-based simulation models. The modelling of these individuals and their decision-making behaviour is often complex, and closely depends to the research question at hand. As a result, agent behaviour needs to be adjusted. For example, choice of transport mode in sightseeing scenarios differs from the choice in everyday commuting to work as travellers value time differently. It is precisely these adjustments in agent modelling that transportation researchers have to implement in order to be able to properly simulate their research scenarios in the first place. In this context, various researchers are repeatedly confronted with difficulties, as options for modelling and calibration of the agents are either simplified to aggregated parameters, or are associated with complex programming that often requires a deep understanding of the underlying software architecture. The problem does not only relate to researchers with a background in computing science, but also to those who would rather deal with traffic engineering issues exclusively. Consequently, it can be anticipated that these researchers will be overburdened when customising existing models, which is why new ontological concepts are needed to simplify this process. Otherwise, these researchers will start to develop their own simulation models from scratch as illustrated by [7].
As agent modelling essentially depends on the scenario being investigated, agents are usually modelled specifically for one particular scenario. For flexible reuse of agents in different scenarios, we need methods that en-

able generalisation of agent behaviour. A similar problem can be observed in general problem solving which is a subcategory of artificial intelligence [27]. [28] analysed implementations of domain-specific problem solving, in order to identify abstraction patterns that can define different methods of general problem solving. These patterns have served as the basis for numerous subsequent research. Particularly, the *CommonKADS* project is one of the outcomes [29]. The project created its own abstraction patterns for general problem solving and also expanded on concepts of knowledge engineering. Based on this, [27] describe the *expertise* of a system as the combination of knowledge about the contexts of the observation subject at hand and the ability to draw conclusions. An example is given of the knowledge acquisition process for building an domain-specific problem solver that performs fault detection on bicycles: (1) First, a mechanic that specialises in bicycles is interviewed about his working methods. (2) In addition, the same mechanic is observed while at work in order to also capture implicit knowledge that cannot be expressed and described with words and that is needed for such a diagnosis. (3) Furthermore, documents such as repair manuals or measurement tables can be included. Collected knowledge can be merged into a unitary model of expertise. [27] point out that different types of knowledge are involved. More particularly, knowledge on the assembly of bicycles, about the mechanics, as well as knowledge about possible faults and their causes, and knowledge about the procedure for recognising and repairing faults. The CommonKADS project has defined a layered model for distinguishing the different types of knowledge (see figure 1). The lowest layer describes *Domain Knowledge*. In this layer, domain-specific concepts and simple relations are defined. Considering the example of fault detection for bicycles, information on this layer may include what a bicycle is, which parts it consists of, which possible faults may occur, as well as possible causes of faults and corrective measures. *Inference Knowledge* is located in the layer above. This layer contains information about the logical contexts of the concepts defined in the domain knowledge. Based on this, conclusions can be drawn using various methods and algorithms. Finally, at the top layer there is *Task Knowledge* in which information from the lower levels is brought together in order to perform decision-making and determine actions.

We now propose an architecture analogous to the CommonKADS knowledge structure as knowledge base for adaptable agents in traffic simulations (see Figure 2). This knowledge base is implemented by means of OWL ontologies extended with SWRL rules. We distinguish between two types of domain knowledge: *travel* and *activity* information. Concepts of the first type of domain knowledge are relevant for traffic related aspects such as mode options. To facilitate reusability, they are encoded in a separate ontology which we call the *travel ontology*. The *travel ontology* exclusively contains knowledge on common traffic concepts for example *transportation modes, road signs, etc.* The second type of domain knowledge can be referred to as *activity* information and extends the knowledge of the agents by concepts that are relevant to model research-question-specific activity. For example, when simulating a sightseeing scenario agents require completely different activity information compared to a grocery shopping scenario. This method allows for a flexible extension of agent knowledge. Agents are not bound to one type of activity information, but may also integrate several activity ontologies for broader decision-making and simulation of more complex scenarios. Regarding the layer of inference knowledge, all ontologies containing travel or activity related information are consolidated (imported) into a central *person ontology*. This ontology contains information about person-specific concepts such as census properties. This enables the implementation of decision-making in various domains using only one software agent. The idea matches the individual in the real world, that is constantly required to make multi-criteria decisions based on preferences from various aspects in life. The defined concepts in the domain knowledge can be used to formulate a set of logic based inference rules that enables the application of computer based reasoners. By employing these established reasoning mechanisms, we use census information as input data to infer domain-specific preferences that can be used as criteria for agent decision making. For example, in a grocery shopping scenario travellers have to make a decision as to which supermarket they want to approach. This decision not only depends on traffic-related preferences (transport mode, shortest distance, etc.) but also on personal food preferences. This reflects different domains of knowledge. Travellers who particularly value organic and sustainable products would possibly be willing to travel to a specialist store for organic food even if the distance is a bit longer. For determining these preferences, rules can be defined according to the following scheme:

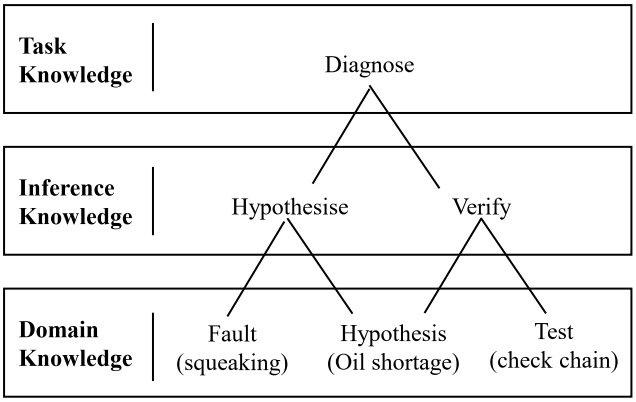


Fig. 1. CommonKADS: Types of Knowledge (see [20],[22]).

$$Person(?p) \wedge hasCensusProperty(?p, ?cprop) \wedge$$
$$swrlb : equal(?cprop, specificProperty) \wedge$$
$$Preference(?pr) \wedge hasPreference(?p, ?pr)$$
$$\Rightarrow Person(?p) \wedge hasValue(?pr, assigned\_value)$$

The rule states that if a person p has a specific census property *cprop*, then it can be inferred that this person holds the value *assigned_value* for a preference pr. An example may look as follows: If a person p has an age of 18-25 years (*cprop*), then it can be concluded that the person p has a preference for organic food pr of 5. Assuming that pr is for example measured on a *Likert* scale from 1 to 5 [30]. For reasons of comprehensibility, a simple example rule has been formulated. In practical modelling, preferences should be inferred using probability distributions as even within the age group of 18-25, there are various types of travellers with varying preferences. Moreover, the same preference pr can be inferred from different census properties. The multiple inference of values for the same preference pr results in probabilities for all attributes of the Likert scale that can be considered in final decision making. With our approach, researchers that are looking to customise the simulation model for research question specific purposes no longer have to deal with complex programming, but instead can make use of the benefits of semantic modelling. Using tangible persona models, settings for different agent types can be captured in a comprehensible form. Agents are assigned to persona types and are mirrored as individuals into the ontology. This means that individual conclusions can be drawn for the particular agent, and the inferred preferences can be incorporated into the decision-making behaviour of the agent.

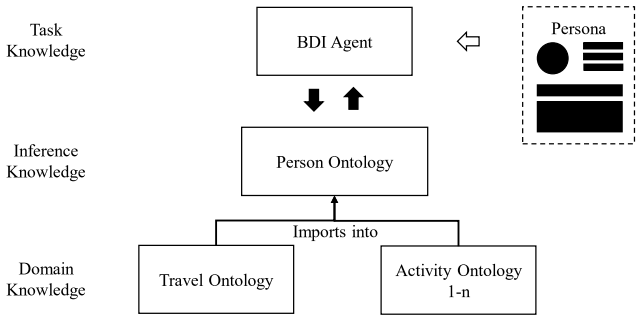


Fig. 2. Modelling Concept.

Finally for implementing task knowledge, the BDI model as a well established paradigm for implementing intelligent agents is particularly suitable. It enables software agents to perform action decisions (intentions) on the basis of defined goals (desires) and their modelled knowledge of their external world (beliefs) [31]. The BDI model is well suited to model actors in traffic scenarios: not only destinations of journeys but also optimisation goals e.g. minimal travel time, minimal emissions can be formulated as desires. Travel preferences and other parameters are potential beliefs that can be used to determine e.g. the selection of means of transportation. In our own previous work, we have given proof that a separation of general agent activity logic from aspects of modelling agent knowledge is an efficient and effective approach [32].

PROOF OF CONCEPT

To demonstrate the benefits of our proposed modelling method we have selected two example scenarios that fundamentally differ in types of mobility. In practical application this will be the case when specific research questions at hand deviate from the main focus of available simulators. We use the proposed modelling method and perform simulation as proof of concept. For both simulations we have exported geographical map data from *OpenStreetMap* [33] for the area around the city of Wetzlar which is located in Hesse, Germany, in addition we use data provided by the German census of 2011 [34]. Regarding the different types of travellers, for both scenarios, we created 12 persona based on a classification provided in [35] (see figure 3). The classification is based on various stages in life (*age/occupation status*) as well as *family status* and *social strata/income* (as illustrated in [36]) in order to represent the most significant groups of people in the German demographic. AGADE Traffic provides an option to create this type of persona using the web frontend.

For the first simulation, a commuter scenario has been modelled in which individuals start from various residential areas with all having the same target location. In real world scenarios, this is the case for example, when large gatherings take place or a large number of persons is commuting to the same workplace. We have marked the event arena in Wetzlar as the venue and thus, the desired target location for all agents. Furthermore, markers for each of the residential areas in the surrounding area of Wetzlar have been defined. The distribution of traveller agents starting from each residential area is based on data provided by the German census of 2011. A commuter scenario of this type primarily deals with knowledge about the traffic domain. In this context, route choice problems are commonly studied to determine current effects on the infrastructure or immediate surroundings. For example, research on transportation usually attempts to relieve particularly crowded road sections by improving traffic management, which is supposed to evenly distribute travel volume across alternative routes. For simulating this type of route choice problems, AGADE Traffic provides a default simulation model. The default simulation model generally assumes that all agents are travelling by car and performs routing based on the A* algorithm (see [37]) that uses a cost functions based on shortest distance and additional geographical information. However in this context, the question of mode choice, e.g. travelling by car, bicycle, or walking, is just as relevant. Therefore, we perform customisation to the supplied default simulation model, just as researchers would like to do with research specific problems. The authors created an example ontology using OWL for modelling domain knowledge on *traffic* concepts. Using the ontology, agents obtain knowledge about different transportation modes available to them. For ease of exposition, in this paper we limit this ontol-

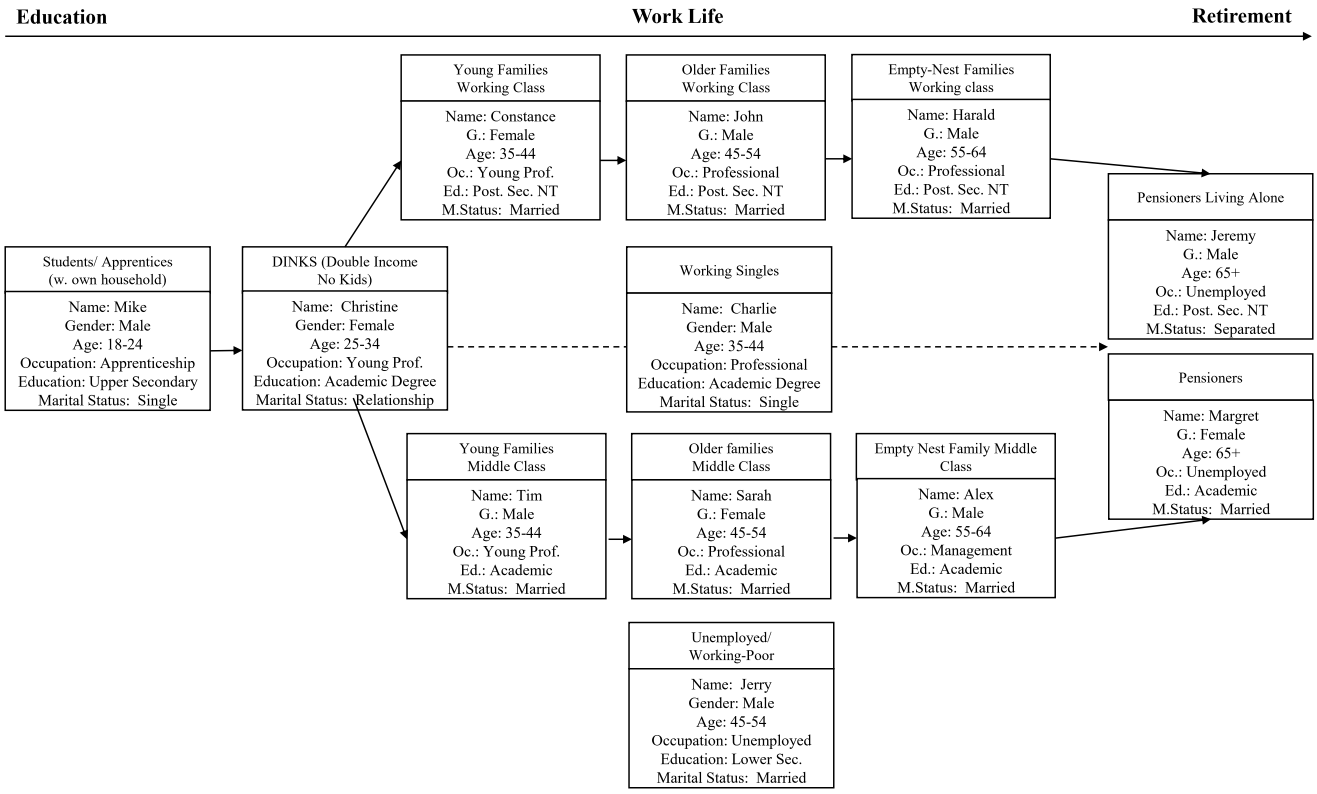


Fig. 3. Persona Models.

ogy to the concepts of various travel mode options. In particular, information on *cars, bicycles, walking* and *public transport* has been modelled. For more complex scenarios that require an expanded knowledge of the traffic domain, such as the simulation of *Intelligent Transportation Systems (ITS)* or testing of traffic light algorithms, this ontology can be extended. According to the modelling structure illustrated in figure 2, this ontology is equivalent to the *travel ontology*. Within this simulation, no activity ontology has been implemented as domain knowledge about traffic concepts is sufficient. In addition, we have extended the *person ontology* to define rules that reflect the decision-making behaviour regarding travel mode choice. The rules are created using semantic methods and do not require complex programming capabilities. The *person ontology* describes the traveller agent as a person concept which is itself described by various census properties. Furthermore, preferences are modelled in the person ontology which can be included as criteria for decision making. Using survey data, rules can be formulated that infer real values for the preferences of the agents based on the census properties defined in the respective personas. For this simulation, we have used data from [38]. For the integration of the inferred knowledge into the layer of Task Knowledge, it is not possible to completely avoid programming. Using our proposed modelling structure we have reduced the amount of programming required to the minimum. AGADE Traffic is written in Java and implements BDI agents using the JADEX framework [39]. For customisation purposes, AGADE Traffic makes use of the advantages of object oriented programming, and provides a central interface within the agent to implement decision-making algorithms or cost functions based on the inferred decision criteria from the ontology. For selection of travel mode, we have implemented a simple utility function that determines a personal utility score for each agent and mode based on utility values of the mode for various dimensions (monetary costs, eco-friendliness, etc.) and the inferred personal preference: Assuming I being the set of modelled preferences in the ontology with $i \in I$, n being the number of preferences in I, $U_i(m)$ being the utility value of a transportation mode m for preference dimension i, and p_i the inferred value of the personal preference of dimension i for an agent. Based on this, we define $UtilityScore(m) = \sum_{i=0}^{n} U_i(m) * p_i$. Furthermore, we implemented mode selection based on $Max(UtilityScore)$. This concludes the customisation performed for the first simulation scenario.

We have created a second scenario in which we simulate mobility related to grocery shopping. The characteristics of this scenario differ significantly from the first simulation. While all agents in the first simulation had a common target location, the grocery shopping scenario features different shopping locations that agents can travel to. Agents are assigned a generated list of food items to purchase and are then required to make decisions about the selection of supermarkets as well as mode of travel. It should be noted that supermarkets not only differ in product supply, but also

available stock may vary in product quality and sustainability. Consequently, in some cases agents will not be able to purchase all items on the assigned grocery list at a selected grocery store, which requires them to visit subsequent target locations. In comparison to the first simulation, the decision-making process and the number of decision criteria involved are much more diverse. Using our proposed modelling structure, we demonstrate necessary customisation.

With regard to the difference in agent decision-making, it can be noted that agents have to decide on two major aspects; firstly, the selection of target locations (supermarkets) and secondly the selection of the travel mode. Decision criteria includes preferences not only regarding travel related aspects but also food related properties. Therefore, domain knowledge has to be extended by a separate ontology that provides information on various types of food and grocery stores, as well as information on available product inventory and further product related properties such as quality, sustainability, price tendency, etc. Considering the proposed modelling structure illustrated in Figure 2, this *food ontology* matches an *activity ontology* that researchers have to append when customising the provided default model for research specific scenarios. For this simulation, we thus make use of the same *travel ontology* from the first scenario, but append a new *food ontology* to the domain knowledge. We then extended the *person ontology* by rules that conclude information on food preferences. For this, we make use of polling data provided by [40]. With this, it is possible to infer all necessary preference information regarding both travel and food related aspects. Finally, we can use the provided programming interface within the agent to implement algorithms regarding decision-making of agents. The selection of supermarkets can for example be implemented in a similar manner using utility functions as demonstrated for travel mode selection. Given that the focus of this example is the description of the customisation process, at this point we will not further elaborate on the precise algorithm that we have implemented for this scenario. However, we will make source code and simulation data available.[1] The algorithms for the implementation of the decision behaviour can be kept arbitrarily complex or simple depending on the research question at hand. With our proposed modelling structure, we create the basis for capturing all necessary decision preferences without complex programming and at the same time allow for flexible and adaptive scaling of the domain knowledge.

CONCLUSION AND FUTURE WORK

As customisation options in available traffic simulators are either simplified to aggregated parameters or associated with complex programming, existing simulation models have not been reused to their full potential. As a result, researchers dealing with specific research questions have rarely made used of available simulators, but instead created their own simulation environment from scratch. Based on the ideas of the CommonKADS project and application of persona models and semantic methods, we have created a modelling structure that facilitates easy reuse by reducing required programming to the necessary minimum. Moreover, our modelling structure allows for adaptable modelling of agent knowledge as well as decision behaviour. For future work, modelling of both travel and activity related knowledge can be expanded. The creation and combination of further activity models for various domains may result in an open source library of activity knowledge that can be flexibly integrated, reused and customised for modelling complex research specific simulations.

[1] see https://github.com/kite-cloud/agade-traffic

ACKNOWLEDGEMENT

This research has been supported by a grant from the Karl-Vossloh-Stiftung (Project Number S0047/10053/2019).

REFERENCES

[1] M. Poeck and D. Zumkeller, "Die anwendung einer massnahmenempfindlichen prognosemethode am beispiel des grossraums nürnberg," in *DVWG-Workshop Policy Sensitive Models, Giessen*, 1976.

[2] K. Axhausen and R. Herz, "Simulating activity chains: German approach," *Journal of Transportation Engineering*, vol. 115, pp. 316–325, may 1989.

[3] Texas Transportation Institute, "Early deployment of transims: Issue paper," 1999.

[4] D. Krajzewicz, G. Hertkorn, C. Rössel, and P. Wagner, "Sumo (simulation of urban mobility)-an open-source traffic simulation," in *Proceedings of the 4th middle East Symposium on Simulation and Modelling (MESM20002)*, pp. 183–187, 2002.

[5] A. Horni, K. Nagel, and K. Axhausen, *The Multi-Agent Transport Simulation MATSim*. Ubiquity Press, aug 2016.

[6] J. Auld, M. Hope, H. Ley, V. Sokolov, B. Xu, and K. Zhang, "POLARIS: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations," *Transportation Research Part C: Emerging Technologies*, vol. 64, pp. 101–116, mar 2016.

[7] M. Jakob and Z. Moler, "Modular framework for simulation modelling of interaction-rich transport systems," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, pp. 2152–2159, IEEE, IEEE, oct 2013.

[8] M. Treiber and A. Kesting, "An open-source microscopic traffic simulator," *IEEE Intelligent Transportation Systems Magazine*, vol. 2, no. 3, pp. 6–13, 2010.

[9] S. Cheng and T. Nguyen, "TaxiSim: A multiagent simulation platform for evaluating taxi fleet operations," in *Advanced Agent Technology*, vol. 2, pp. 359–360, Springer Berlin Heidelberg, 2012.

[10] M. Horn, "Multi-modal and demand-responsive passenger transport systems: a modelling framework with embedded control systems," *Transportation Research Part A: Policy and Practice*, vol. 36, pp. 167–188, feb 2002.

[11] B. De Win, F. Piessens, W. Joosen, and T. Verhanneman, "On the importance of the separation-of-concerns principle in secure software engineering," in *Workshop on the Application of Engineering Principles to System Security Design*, pp. 1–10, 2002.

[12] D. McGuinness, F. Van Harmelen, *et al.*, "Owl web ontology language overview," *W3C recommendation*, vol. 10, no. 10, p. 2004, 2004.

[13] I. Horrocks, P. Patel-Schneider, H. Boley, S. Tabet, B. Grosof, M. Dean, *et al.*, "Swrl: A semantic web rule language combining owl and ruleml," *W3C Member submission*, vol. 21, no. 79, pp. 1–31, 2004.

[14] J. Geyer, J. Nguyen, T. Farrenkopf, and M. Guckert, "AGADE traffic 2.0 - a knowledge-based approach for multi-agent traffic simulations," in *Advances in Practical Appli-*

cations of Agents, Multi-Agent Systems, and Trustworthiness. The PAAMS Collection, pp. 417–420, Springer International Publishing, 2020.
[15] S. Blomkvist, "The user as a personality: A reflection on the theoritical and practical use of personas in hci design," *Proceedings of the Technical report*, pp. 1–13, 2006.
[16] A. Cooper, "The inmates are running the asylum. indianapolis, ia: Sams," *Macmillan*, 1999.
[17] S. Calde, K. Goodwin, and R. Reimann, "SHS orcas," in *Case studies of the CHI2002|AIGA Experience Design FORUM on - CHI '02*, pp. 2–16, ACM Press, 2002.
[18] N. Guarino, D. Oberle, and S. Staab, "What is an ontology?," in *Handbook on Ontologies*, pp. 1–17, Springer Berlin Heidelberg, 2009.
[19] A. Bazzan, M. do Amarante, T. Sommer, and A. Benavides, "Itsumo: an agent-based simulator for its applications," in *Proc. of the 4th Workshop on Artificial Transportation Systems and Simulation. IEEE*, p. 8, 2010.
[20] B. Torabi, M. Al-Zinati, and R. Wenkstern, "MATISSE 3.0: A large-scale multi-agent simulation system for intelligent transportation systems," in *Advances in Practical Applications of Agents, Multi-Agent Systems, and Complexity: The PAAMS Collection*, pp. 357–360, Springer International Publishing, 2018.
[21] M. Adnan, F. Pereira, C. Azevedo, K. Basak, M. Lovric, S. Raveau, Y. Zhu, J. Ferreira, C. Zegras, and M. Ben-Akiva, "Simmobility: A multi-scale integrated agent-based simulation platform," in *95th Annual Meeting of the Transportation Research Board Forthcoming in Transportation Research Record*, 2016.
[22] V. Chu, J. Görmer, and J. Müller, "Atsim: Combining aimsum and jade for agent-based traffic simulation," in *Proceedings of the 14th Conference of the Spanish Association for Artificial Intelligence (CAEPIA)*, vol. 1, 2011.
[23] S. Thulasidasan, S. Kasiviswanathan, S. Eidenbenz, E. Galli, S. Mniszewski, and P. Romero, "Designing systems for large-scale, discrete-event simulations: Experiences with the fasttrans parallel microsimulator," in *2009 International Conference on High Performance Computing (HiPC)*, pp. 428–437, IEEE, IEEE, dec 2009.
[24] M. Jakob, Z. Moler, A. Komenda, Z. Yin, A. Jiang, M. Johnson, M. Pěchouček, and M. Tambe, "Agentpolis: towards a platform for fully agent-based modeling of multimodal transportation," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3*, pp. 1501–1502, International Foundation for Autonomous Agents and Multiagent Systems, 2012.
[25] L. Quadrifoglio, M. Dessouky, and F. Ordóñez, "A simulation study of demand responsive transit system design," *Transportation Research Part A: Policy and Practice*, vol. 42, pp. 718–737, may 2008.
[26] M. Wooldridge, "Agent-based software engineering," *IEE Proceedings - Software Engineering*, vol. 144, no. 1, p. 26, 1997.
[27] M. Guckert and T. Péus, "Problemlösungsmethoden reloaded: Integration von domänenwissen zur anwendung allgemeiner lösungsstrategien," *Integration und Konnexion*, p. 70, 2013.
[28] W. Clancey, *Classification problem solving*. Stanford University Stanford, CA, 1984.
[29] G. Schreiber, H. Akkermans, A. Anjewierden, N. Shadbolt, R. de Hoog, W. Van de Velde, R. Nigel, B. Wielinga, *et al.*, *Knowledge engineering and management: the CommonKADS methodology*. MIT press, 2000.
[30] R. Likert, "A technique for the measurement of attitudes.," *Archives of psychology*, 1932.
[31] M. Bratman, D. Israel, and M. Pollack, "Plans and resource-bounded practical reasoning," *Computational Intelligence*, vol. 4, pp. 349–355, sep 1988.
[32] T. Farrenkopf, M. Guckert, N. Urquhart, and S. Wells, "Ontology based business simulations," *Journal of Artificial Societies and Social Simulation*, vol. 19, no. 4, 2016.
[33] M. Haklay and P. Weber, "OpenStreetMap: User-generated street maps," *IEEE Pervasive Computing*, vol. 7, pp. 12–18, oct 2008.
[34] Statistische Ämter des Bundes und der Länder, *Zensus 2011: Methoden und Verfahren*. Wiesbaden, Hesse, Germany: Statistisches Bundesamt, 2015.
[35] GfK Consumer Panels, *Consumers' choice '17 - neue Muster in der Ernährung: die Verbindung von Genuss, Gesundheit und Gemeinschaft in einer beschleunigten Welt : eine Publikation anlässlich der Anuga 2017*. GfK Consumer Panels and Bundesvereinigung der Deutschen Ernährungsindustrie e.V., 2017.
[36] N. Pestel and E. Sommer, "Analyse der verteilung von einkommen und vermögen in deutschland," tech. rep., Institute of Labor Economics (IZA), 2016.
[37] P. Hart, N. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.
[38] U. Engel and M. Pötschke, "Mobilität und verkehrsmittelwahl 1999/2000." GESIS Datenarchiv, Köln. ZA4203 Datenfile Version 1.0.0, https://doi.org/10.4232/1.11591, 2013.
[39] L. Braubach, W. Lamersdorf, and A. Pokahr, "Jadex: Implementing a bdi-infrastructure for jade agents," 2003.
[40] Statista, "Lebensmittelkauf in deutschland," 2020.

AUTHOR BIOGRAPHIES

JOHANNES NGUYEN is a research assistant at Technische Hochschule Mittelhessen in Friedberg from which he received his master's degree. He is also a PhD student at Edinburgh Napier University. His research interests include multi-agent systems, mechanism design and semantic modelling particularly applied to the field of smart cities and future mobility.

THOMAS FARRENKOPF is a lecturer at the at Technische Hochschule Mittelhessen. He completed a PhD degree in the School Of Computing at Edinburgh Napier University examining the use of software agents and ontologies for business simulation, applied to business games. His research areas include multi-agent systems, semantic modelling and computer science applied to industry projects.

MICHAEL GUCKERT is a Professor of Applied Informatics at Technische Hochschule Mittelhessen and a head of department at Kompetenzzentrum für Informationstechnologie (KITE). He received a PhD in Computer Science from Philipps University Marburg. His research areas are multi-agent systems, model driven software development and applications of AI.

Simon T. Powers is a lecturer at Edinburgh Napier University. He gained his PhD from the University of Southampton, examining evolutionary explanations for cooperative social group formation. His research interests include computational social, political, and economic science in which he investigates the links between institutions, computer science, and multi-agent systems.

NEIL URQUHART is a lecturer at Edinburgh Napier University where he is Programme Leader for Computing Science. He gained his PhD from Edinburgh Napier University in 2002, examining the use of Software Agents and Evolutionary Algorithms to solve a real-world routing optimisation problem. His research interests include Evolutionary Computation and Agent-based Systems and their application to real-world problems.

DIFFERENTIAL EVOLUTION ALGORITHM IN MODELS OF TECHNICAL OPTIMIZATION

Roman Knobloch and Jaroslav Mlýnek
Department of Mathematics
Technical University of Liberec
Studentská 2, 46117 Liberec
The Czech Republic
E-mail: roman.knobloch@tul.cz, jaroslav.mlynek@tul.cz

KEYWORDS
Optimization, search space, cost function, differential evolution algorithm, global convergence, asymptotic convergence, mathematical model.

ABSTRACT

At present, evolutionary optimization algorithms are increasingly used in the development of new technological processes. Evolutionary algorithms often allow the optimization procedure to be performed even in cases where classical optimization algorithms fail (e.g. gradient methods) and where an acceptable solution is sufficient to solve the optimization task. The article focuses on possibilities of using a differential evolution algorithm in the optimization process. This algorithm is often referred to in the literature as a global optimization procedure. However, we show by means of a practical example that the convergence of the classic differential algorithm to the global extreme is not generally assured and is largely dependent on the specific cost function. To remove this weakness, we designed a modified version of the differential evolution algorithm. The improved version, named the modified differential evolution algorithm, is described in the article. It is possible to prove asymptotic convergence to the global minimum of the cost function for the modified version of the algorithm.

INTRODUCTION

New technological procedures are often developed using mathematical models describing the essential features of the solved problem. The model is then used to transform the real world problem into an optimization task. Strong assumptions are often required when using classic optimization methods (e.g. convexity of the searched space, convexity of the evaluation function, knowledge of the appropriate position of the initial solution). Otherwise, these methods do not often lead to the required solution.
Evolutionary optimization algorithms are primarily utilized in situations when other usual methods fail to converge to the optimized state. Recently, use of the evolutionary optimization algorithms has been considerably expanding, see e.g. (Simon 2013), (Affenzeller et al. 2009)). The evolutionary algorithms are in particular appropriate for problems with a complicated structure of the search space and in case of intricate cost functions. Evolutionary algorithms are in general more computationally demanding and they are therefore suitable for calculations that are not time limited (e.g. off-line calculations of trajectories of an industrial robot, see (Mlýnek et al. 2020)). Their use in time critical calculations is rather limited. For example, their utilization for online decision making processes (e.g. online calculations of trajectories of industrial robot depending on the evaluation of current conditions) is not so frequent. Nevertheless, parallel programming tools are often used to speed up calculations with good results. Nowadays, parallel programming tools form a part of most used programming languages.
The differential evolution algorithm is one of the frequently used algorithms for solving practical optimization tasks. This algorithm was first introduced by Storn and Price in (Storn and Price 1997) and (Price et al. 2005). This algorithm is often referred to as a global optimization method (see (Storn and Price 1997), (Price 1996)). However, such statements are always justified. We demonstrate by an example of a specific cost function that this algorithm is prone to premature local convergence and its convergence to the minimum of the cost function is not assured. The issue of suitable choice of optional algorithm parameters is solved, for example, in (Červenka and Boudná 2018). In this article we propose a suitable modification of the differential evolution algorithm that eliminates the premature convergence to a local minimum. Additionally, it is possible to prove asymptotic convergence to the global minimum of the cost function.
The differential evolution algorithms now constitute a larger group of similar algorithms that differ in implementation details. We concentrate on the standard *DE/rand/1/bin* algorithm which is best known and mostly used. That is why it is termed as the classic differential evolution algorithm in (Price at al. 2005). Hereafter it is referenced to as CDEA. The new proposed modification of CDEA is termed to as the modified differential algorithm denoted by abbreviation MDEA.

CLASSIC DIFFERENTIAL EVOLUTION ALGORITHM AND GLOBAL CONVERGENCE

In this part we briefly describe the operation of CDEA. Generally, CDEA seeks for the minimum of the cost function by constructing whole generations of

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

individuals. Each individual is an ordered set of specific values corresponding to one point in the cost function domain. In this way each individual represents a potential solution to the optimization task. The quality of this individual is determined by the evaluation of the cost function corresponding to this individual. The next generation is formed from the existing generation by means of mutation and crossover operators. Specifically, we go successively through all individuals in the generation *G*. To each individual y_m^G (termed as the *target individual*) we select randomly three other (different) individuals y_{r1}^G, y_{r2}^G, y_{r3}^G from the current generation. We form in a specific way (including randomness) a combination of these three random individuals and the target individual. This combination is termed as the *trial individual* and denoted y_m^{trial}. Then we evaluate the cost function for the target y_m^G and trial individual y_m^{trial} and compare the results. The individual with lower value of the cost function advances to the position of the target individual of the next generation y_m^{G+1}. When this procedure is completed for all target individuals in generation *G*, we have constructed the new generation of individuals numbered $G + 1$.
The next part illustrates CDEA operation in a definite way in the form of pseudo code.

Input:
Optimization task parameters:
f denotes the cost function, *D* is the dimension of the cost function domain, $< x_{j\min}, x_{j\max} >$ is a domain of each cost function variable x_j.

CDEA parameters:
NP denotes the generation size (the number of individuals in each generation), *NG* is the number of calculated generations, *F* stands for mutation factor ($F \in < 0, 2 >$), and CR denotes the crossover probability ($CR \in < 0, 1 >$). The symbol *G* stands for the generation number, index *m* is the number of the individual in the generation, index *j* describes the *j*-th component of a specific individual y_m.

Computation:
1. Create an initial generation ($G = 0$) of *NP* individuals y_m^G, $1 \le m \le NP$, (e.g. by use of relation (1)).
2. a) Evaluate all individuals y_m^G of the *G*-th generation (calculate $f(y_m^G)$ for each individual y_m^G). b) Store the individuals y_m^G and their evaluations $F(y_m^G)$ into matrix **B** (each matrix row contains parameters of individual y_m^G and its evaluation $F(y_m^G)$. That is matrix be has *NP* rows and D+1 columns ($1 \le m \le NP$).
3. *while* $G \le NG$

a) *for* $m := 1$ *step* 1 to *NP* *do*
(i) randomly select index $s_m \in \{1, 2, \ldots, D\}$,
(ii) randomly select indexes $r_1, r_2, r_3 \in \{1, \ldots, NP\}$, where $r_l \ne m$ for $1 \le l \le 3$;
$r_1 \ne r_2, r_1 \ne r_3, r_2 \ne r_3$;
(iii) *for* $j := 1$ *step* 1 to *D* *do*
if *rand* $(0,1) \le CR$ or $j = s_m$) *then*
$y_{m,j}^{trial} := y_{r3,j}^G + F\left(y_{r1,j}^G - y_{r2,j}^G\right)$
else
$y_{m,j}^{trial} := y_{m,j}^k$
end if
end for (*j*)
(iv) *if* $f\left(y_m^{trial}\right) \le f\left(y_m^k\right)$ *then* $y_m^{G+1} := y_m^{trial}$
else $y_m^{G+1} := y_m^G$
end if
end for (*m*)

b) store individuals y_m^{G+1} and their evaluations $f\left(y_m^{G+1}\right)$ $(1 \le m \le NP)$ of the new $(G+1)$-st generation in the matrix **B**, $G := G + 1$
end while (*G*).

Output:
The row of matrix **B** that contains the corresponding value $\min\{F\left(y_m^G\right); y_m^G \in \mathbf{B}\}$ represents the best found individual y_{opt}.

Comments

The individual y_{opt} in pseudo-code of CDEA is the final solution of the optimization problem.

One way of possible forming the initial generation ($G = 0$) of individuals y_m^0 is given by relation

$$y_{m,j}^0 := x_{j\min} + \text{rand}(0, 1) \cdot \left(x_{j\max} - x_{j\min}\right). \quad (1)$$

Values $x_{j\min}$ and $x_{j\max}$ are lower and upper limit of variable x_j. The function rand $(0, 1)$ randomly generates a value from a closed interval $< 0, 1 >$.

Counterexample to Global Convergence of CDEA (Premature Convergence)

It is not difficult to find counterexamples to the global convergence of the CDEA. Let us consider for instance the following two graphs of cost functions with the domain in Euclidean space R^2, see Figure 1. Even for the cost function shown in Figure 1 above the probability that the CDEA finds the global minimum of the cost function is less than one. The reason is that the CDEA

can converge in some cases relatively fast to the local minimum missing completely the global minimum. This results in concentrating the individuals in subsequent generations around the local minimum. As soon as the size of the generation falls under some critical value, the generation is too small to produce trial individuals that could hit the region in the vicinity of the global minimum. This situation is called a *premature convergence*. In this case even increasing the number of generations does not lead to increasing the chance to identify the global minimum. Moreover, the probability that the CDEA finds the global minimum falls with the decreasing measure of the global minimum region. The probability of finding the global minimum for the cost function in Figure 1 below is substantially smaller than for the cost function in Figure 1 above. Additionally, by a sufficient reduction of the measure of the global minimum region this probability can be made as close to zero as possible.

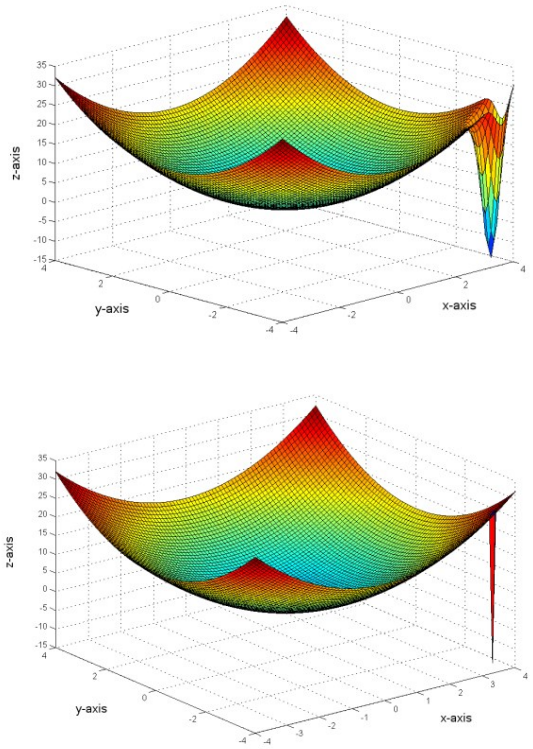


Figure 1: Examples of cost functions with domains in R^2

Numerical Example

We can present a specific cost function to demonstrate the limited ability of CDEA to converge to the global minimum of the cost function. To keep things simple we consider the domain of the cost function as a subset of the two dimensional Euclidean space R^2. We will construct the cost function $F(x_1, x_2)$ as a composition of two simple functions

$$F(x_1, x_2) = F_B(x_1, x_2) + F_M(x_1, x_2). \qquad (2)$$

The term $F_B(x_1, x_2)$ represents the base function. This function is smooth and has one shallow minimum. It can be defined for instance in the following way

$$F_B(x_1, x_2) = x_1^2 + x_2^2,$$

with the domain $D(F_B) = \langle -H, H\rangle \times \langle -H, H\rangle$, where H determines the boundary values of the domain.

The term $F_M(x_1, x_2)$ denotes a modifier function. This function should be relatively steep and with a rather small domain. We use the function $F_M(x_1, x_2)$ to modify the underlying base function $F_B(x_1, x_2)$. The role of the function $F_M(x_1, x_2)$ is to realize the global minimum of the cost function $F(x_1, x_2)$ in relation (2). To be able to construct the function $F_M(x_1, x_2)$ effectively, we introduce another auxiliary function F_P,

$$F_P(x_1, x_2) = x_1^2 + x_2^2 - 1,$$

with the domain $D(F_P) = \{x_1, x_2: x_1^2 + x_2^2 \leq 1\}$. It is obvious that the function $F_P(x_1, x_2)$ is defined exclusively on a unit circle and has values from the closed interval $\langle -1, 0\rangle$. The graph of the function $F_P(x_1, x_2)$ is a circular paraboloid presented in Figure 2.

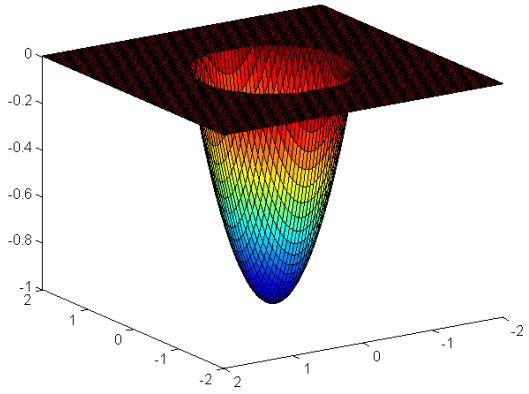


Figure 2: Graph of the auxiliary function $F_P(x_1, x_2)$

The function $F_M(x_1, x_2)$ is then formed as

$$F_M(x_1, x_2) = \lambda_h \cdot F_P\left(\frac{1}{\rho}(x_1 - x_{G1}), \frac{1}{\rho}(x_2 - x_{G2})\right).$$

Here the number λ_h defines the height of the resulting circular paraboloid, ρ denotes the radius of the domain on which the modifier function $F_M(x_1, x_2)$ is defined. Obviously, the modifier function $F_M(x_1, x_2)$ is defined only for points that are closer to the point $[x_{G1}, x_{G2}]$ than the radius of its domain ρ. The coordinates x_{G1}, x_{G2} specify the point, where the modifier function $F_M(x_1, x_2)$ attains its minimum. The overall cost function $F(x_1, x_2)$ is then defined according to the relation (2) by the composite formula

$$F(x_1, x_2) = x_1^2 + x_2^2 + \lambda_h \cdot F_P\left(\frac{1}{\rho}(x_1 - x_{G1}), \frac{1}{\rho}(x_2 - x_{G2})\right) \qquad (3)$$

We have to choose the parameters λ_h, ρ, x_{G1} and x_{G2} in a reasonable way to obtain the required result. It is clear that we can control the dimensions of the modifier function domain by parameter ρ. The point $[x_{G1}, x_{G2}]$ is to be placed relatively close to the boundary of the cost function domain. This means it is relatively far from the point [0, 0] representing the local minimum of the cost function analogously to the cost functions presented in Figure 1. Since the base function $F_B(x_1, x_2)$ is positive definite, it attains a positive value $F_B(x_{G1}, x_{G2})$ at the point $[x_{G1}, x_{G2}]$. This means we have to take the parameter λ_h sufficiently large, so that the global minimum is essentially lower than the local minimum at the point [0, 0].

We performed numerical experiments with following parameters: $D(F_B) = \langle -H, H\rangle \times \langle -H, H\rangle = \langle -4, 4\rangle \times \langle -4, 4\rangle$ with measure $\mu(D(F)) = 8^2 = 64$, $x_{G1} = x_{G2} = 3$. This means that the global minimum of the cost function F is at point [3,3] and local minimum at point [0,0]. Number of individuals in each generation $NP = 200$, number of generation $NG = 160$, value of parameter $F = 0.8$ and value of parameter $CR = 0.9$. We realized 200 numerical experiments with value of parameter $\rho = \frac{1}{10}$ (then $\mu(D(F_M)) = \pi\rho^2 \doteq 0.0314$) and 200 numerical experiments with value of parameter $\rho = \frac{1}{16}$ (then $\mu(D(F_M)) \doteq 0.01223$). The results illustrating the limited ability of CDEA to identify the global minimum are summarized in Table 1.

Table 1: Experimental testing of CDEA

CDEA	Local minimum hits	Global minimum hits	Success rate in %
$\rho = \frac{1}{10}$	163	37	18.5
$\rho = \frac{1}{16}$	185	15	7.5

Based on the values given in Table 1, it can be assumed that the decreasing value of ρ (and thus value of $\mu(D(F_M))$ will significantly decrease the success rate of the CDEA algorithm in finding the global minimum.

MODIFIED DIFFERENTIAL EVOLUTION ALGORITHM

As illustrated in the previous part, CDEA does not in general guarantee the convergence to the global minimum of the cost function. This is caused by the too fast convergence of CDEA to the local minimum (premature convergence) resulting in rapid reduction of the generation size (which means a loss of diversity). The most straightforward way how to limit the premature convergence is to replace some individuals with the highest values of the cost function by random individuals in each generation. Though these random individuals reduce partially the convergence speed, they increase substantially the diversity of the generation. In technical terms, it is necessary to make one simple change in the CDEA scheme. We present only the differences with respect to CDEA. See the pseudocode description of CDEA in chapter "Classic Differential Evolution Algorithm and Global Convergence".

Input:

We add another parameter R that determines the ratio of random individuals in each generation, $R \in < 0, 1 >$, e.g., $R = 0.1$ means that 10% of individuals in each generation are generated randomly.

Computation:

We add another procedure to the part 3., specifically:

c) determine in matrix B the quantity $\lfloor NP.R \rfloor$ of individuals with the highest cost function values and replace these individuals by randomly generated individuals (e.g. by use of relation (1)) from the search space). Note that here the symbol $\lfloor x \rfloor$ denotes the integer part of the real number x.

This modified algorithm will be called the Modified Differential Evolution Algorithm (MDEA). We applied the numerical experiments on MDEA with the same input parameters as in the previous chapter on CDEA. In addition, the value of R parameter is equal to $R = 0.1$. The results are summarized in Table 2.

Table 2: Experimental testing of MDEA

MDEA	Local minimum hits	Global minimum hits	Success rate in %
$\rho = \frac{1}{10}$	34	166	83.0
$\rho = \frac{1}{16}$	70	130	65.0

Another positive feature of the algorithm MDEA is that if we increase the number of generations NG the global minimum will be identified with an increased probability. This probability can come close to 1 for a sufficiently high number G of generations. We call this aspect of the MDEA an *asymptotic global convergence*. We describe this topic in the following chapter.

ASYMPTOTIC GLOBAL CONVERGENCE OF MDEA

In this part we present several theoretical concepts and statements that can be used to prove the asymptotic global convergence of MDEA. More specifically, we will show that when the number of generations $G \to \infty$ then the probability that MDEA identifies the global minimum of the cost function approaches 1.

Optimal Solution Set

We would like to find the minimum of the cost function with the lowest value

$$\min\{F(x): x \in S\}, \quad (4)$$

where S is a measurable search space of a finite measure representing all possible configurations of variables $x = (x_1, x_2, \dots x_n)$. We suppose that the global minimum of function F exists on S. We define a solution set S^* as

$$S^* = \{x^*: F(x^*) = \min\{F(x): x \in S\}\},$$

where x^* represent global minima of the function F. We introduce an optimal solution set S_ε^* as

$$S_\varepsilon^* = \{x \in S: |F(x) - F(x^*)| < \varepsilon\},$$

where $\varepsilon > 0$ is a small positive real number. Denoting by μ the Lebessgue measure, we suppose that for each ε it holds $\mu(S_\varepsilon^*) > 0$.

Convergence in Probability

To examine the global convergence of MDEA we need to introduce a concept of the convergence in probability defined in (Hu et al. 2013).

Definition: Let $\{G(k), k = 1,2, \dots\}$ be a generation sequence created by a differential evolution algorithm to solve optimization task (4). We say that the algorithm converges to the optimal solution set in probability if

$$\lim_{k\to\infty} p\{G(k) \cap S_\varepsilon^* \neq \emptyset\} = 1, \quad (5)$$

where p denotes the probability of an event.

Now, we can use this concept to formulate the following statement.

Proposition: Let us suppose that for each generation $G(k)$ of a differential evolution algorithm there exists at least one individual y such that

$$p\{y \in S_\varepsilon^*\} \geq \alpha > 0,$$

where α is a small positive value. Then the algorithm converges to the optimal set S_ε^* in probability. That is relation (5) holds.

The proof of this proposition is stated in full in (Knobloch et al. 2017).

It holds that for each generation G of the MDEA it is true

$$p\{y \in S_\varepsilon^*\} = \alpha \geq 0, \quad (6)$$

where α is a small positive value. Here $p\{y \in S_\varepsilon^*\}$ denotes the probability that y belongs to S_ε^*. The validity of relation (6) necessarily results from the generation of random individuals in each generation G of MDEA. It follows that MDEA converges to $y \in S_\varepsilon^*$ for any small real positive number ε. This implies the asymptotic global convergence of MDEA. Thus, we know that MDEA converges to the global minimum. The asymptotic convergence of MDEA is proved in detail in (Knobloch et al. 2017), see also (Hu et al. 2013). Probability estimates of reaching the global minimum after performing G generations of MDEA are given in (Knobloch and Mlýnek 2020). These estimates help to decide after how many generations to finish the MDEA calculation.

CONCLUSIONS

CDEA is a universal optimization algorithm that is frequently used in technical projects, economy studies, natural sciences and other important areas of interest. Nevertheless, it has some principal limitations. The main weakness of CDEA is a possible premature convergence of the computing process to a local minimum of the cost function. We demonstrated this fact by means of a simple example.

Identification of this weakness was the starting point for a search of an improved version of the algorithm that would provide better chances regarding the convergence to the global minimum of the cost function. MDEA is a result of these efforts.

MDEA is not prone to the premature convergence because a certain ratio of random individuals in each generation makes it immune to the loss of generation diversity. From the theoretical point of view, we proved that MDEA converges asymptotically to the global minimum of the cost function in probabilistic sense.

The use of MDEA has proved successful to the authors in solving complicated practical optimization problems. For example, it is the task of optimizing the placement of infrared heaters over a metal thin walled mould in the production of artificial leather for the automotive industry (Slush Moulding technology).

The cost function of this optimization problem is a function of many variables (often 300 and more) and has many local minima. Gradient methods, genetic algorithms and also CDEA found only a local minimum of the corresponding cost function (this optimization problem is described in more detail in (Mlýnek and Knobloch 2018) and (Mlýnek et al. 2016). MDEA has also proved successful in optimizing the fibre winding procedures using a fibre-processing head and a non-bearing frame moved by an industrial robot (for more details see (Mlýnek et al. 2020)).

ACKNOWLEDGMENTS

This article was supported by project "Modular platform for autonomous chassis of specialized electric vehicles for freight and equipment transportation", Reg. No. CZ.02.1.01/0.0/0.0/16_025/0007293.

REFERENCES

Affenzeller, M.; Winkler, S.; Wagner, S.; and A. Beham. 2009. "Genetic Algorithms and Genetic Programming." CRC Press, Boca Raton.

Červenka, M. and H. Boudná. 2018. "Visual Guide of F and CR Parameters Influence on Differential Evolution Solution Quality." *Proceedings of 24th International Conference Engineering Mechanics 2018*, Svratka, Czech Republic, 141-144, DOI: 10.21495/91-8-141.

Hu, Z.; S. Xiong; Q. Su; and X. Deng. 2013. "Sufficient Conditions for Global Convergence of Differential Evolution Algorithm." *Journal of Applied Mathematics*, Article ID 139196.

Knobloch, R.; J. Mlýnek; and R. Srb. 2017. "The Classic Differential Evolution Algorithm and Its Convergence Properties". *Applications of Mathematics*, Vol. 62, No. 2, 197-208.

Knobloch, R. and J. Mlýnek. 2020. "Probabilistic Analysis of the Convergence of the Differential Evolution Algorithm." *J. Neural Network World*, Vol. 30, 249-263, DOI: 10.14311/NNW.2020.30.017.

Mlýnek, J.; R. Knobloch; and R. Srb. 2016. "Optimization of a Heat Radiation Intensity and Temperature Field on the Mould Surface." *Proceedings of 30th European Conference on Modelling and Simulation*, Regensburg, Germany, ISBN: 978-0-9932440-2-5, DOI: 10.7148/2016-0425.

Mlýnek, J. and R. Knobloch. 2018. "Model of shell Metal Mould Heating in the Automotive Industry." *Applications of Mathematics*, Vol. 63, No. 2, 111-124.

Mlýnek, J.; M. Petrů; T. Martinec, T. ; and S.S.R. Koloor. 2020. "Fabrication of High-Quality Polymer Composite Frame by a New Method of Fiber Winding Process." *J. Polymers*, Volume 12(5), 30 pages, https://doi.org/10.3390/polym12051037, Open Access.

Price, K.V. 1996. "Differential Evolution - A Fast and Simple Numerical Optimizer." *Proceedings of North American Fuzzy Information Processing. Berkeley*, 524-527.

Price, K.V.; R.M. Storn ; and J.A. Lampien. 2005. "Differential Evolution, A Practical Approach to Global Optimization." Springer-Verlag, Berlin Heidelberg.

Simon, D. 2013. "Evolutionary Optimization Algorithms." John Wiley & Sons, Hoboken, New Jersey.

Storn, R.M. and K.V. Price. 1997. "Differential Evolution - A Simple and Efficient Heuristics for Global Optimization over Continuous Spaces." *Journal of Global Optimization*. Kluwer Academic Publishers, 11, 341-359.

AUTHOR BIOGRAPHIES

ROMAN KNOBLOCH was born in Turnov, the Czech Republic. He finished his studies at the Charles University in Prague, the Faculty of Mathematics and Physics, where he studied physics and teaching of mathematics and physics. His main areas of interest are: modelling of physical phenomena, modern optimization methods and heat and transport phenomena in continuum mechanics. He works as an assistant professor at the Technical University of Liberec where he also graduated his PhD study programme. His e-mail address is roman.knobloch@tul.cz

JAROSLAV MLÝNEK was born in Trnava, Czechoslovakia and went to the Charles University in Prague, where he studied numerical mathematics at the Faculty of Mathematics and Physics and he graduated in 1981. In his work he focuses on the computational problems of heating and thermal losses in components of electrical machines and on optimization procedures. Currently he works as an associate professor at the Technical University of Liberec. His e-mail address is: jaroslav.mlynek@tul.cz

A ROBUST AND ADAPTIVE APPROACH TO CONTROL OF A CONTINUOUS STIRRED TANK REACTOR WITH JACKET COOLING

Roman Prokop, Radek Matušů and Jiří Vojtěšek
Faculty of Applied Informatics
Tomas Bata University in Zlín
nám. T. G. Masaryka 5555, 760 01 Zlín, Czech Republic
E-mail: prokop@utb.cz

KEYWORDS
Robust Control, Adaptive Control, Continuous Stirred Tank Reactor, 2DOF Control Structure, Algebraic Control Design.

ABSTRACT

Continuous Stirred Tank Reactors (CSTR) are one of the main technological plants used in chemical and biochemical industry. These systems are quite complex with many nonlinearities and the conventional linear control with fixed parameters can be questionable or sometimes unacceptable. The solution should be found in so-called "non-traditional" control approaches like adaptive, robust, fuzzy, or artificial intelligent methods. One way is the utilization of self-tuning adaptive schemes, but computations may be quite difficult, clumsy and time-consuming. This paper brings an alternative principle called a robust approach and the comparison of the robust and adaptive control responses. Robust control considers a CSTR model as a linear system with parametric uncertainty, which covers a family of all feasible plants. Then several controllers with fix parameters are designed so that for all possible plants, the acceptable control behavior is obtained. The two-degree-of-freedom (2DOF) structure for the control law was chosen. Both robust and adaptive control is applied to an original nonlinear model of a CSTR. All calculations and simulations of mathematical models and control responses were performed in the Matlab and Simulink environment.

INTRODUCTION

The plants in technological processes and especially in chemical and biochemical industry usually have nonlinear behavior that causes difficulties in the control of such processes. Another unpleasant feature can be found in the complexity of such processes with a lot of variables and properties that result in difficult mathematical descriptions. This negative property should be overcome with the linearization of nonlinear models that introduces simplifications that reduces the intricacy of the system. On the other hand, this simplification can result in inaccurate descriptions of the system. The utilization of adaptive (e.g. self-tuning) schemes brings more difficult, clumsy and time-consuming computations (Åström and Wittenmark 1989). The control design using a hybrid adaptive control principle was used in (Vojtesek et al. 2017) where the originally nonlinear system was represented by an external linear model with recursively identified parameters and the pole-placement method adjustment principle was applied. A practically favored approach to overcome the loss of the model accuracy, compensated by its structure simplicity, consists in the utilization of a model with uncertainty. This idea allows working with the linear time-invariant low order mathematical models also for the case of real systems with complex dynamics or nonlinear behavior. There are several ways how to incorporate the uncertainty into the mathematical model available, see (Barmish 1994; Bhattacharyya et al. 1995). The popular group of uncertain systems is known as the systems with parametric uncertainty, which means the model structure is fixed but its parameters can vary, typically within some prescribed intervals. Then, the natural task is to find a controller, called a robust controller, that ensures the preserving some important closed-loop properties (e.g. stability) for the whole assumed family of controlled plants, see (Grimble 2006).

The system under the consideration is the Continuous Stirred Tank Reactor (CSTR) with the cooling in the jacket. The mathematical model of this system is described by the set of four nonlinear Ordinary Differential Equations (ODE). This set can be solved by standard numerical methods that are implemented in mathematical software such as Matlab, Simulink etc.

The main aim of this paper is in the design a robustly stabilizing controller for the CSTR with the cooling in the jacket, modelled as a system with parametric uncertainty, by means of algebraic approach. The work will put emphasis on the relatively easily tunable and applicable conventional PID controllers. The robust stabilization and control are verified and discussed by a simulation example of nonlinear CSTR.

CONTINUOUS STIRRED TANK REACTOR

The nonlinear controlled system under the consideration is a CSTR display of which can be found in Figure 1. The so-called Van der Vusse reaction described by general scheme:

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$
$$2A \xrightarrow{k_3} D \qquad (1)$$

is performed inside the reactor.

This system can be described by a nonlinear mathematical model derived with the commonly used simplifications that reduce complexity of the system that has a lot of variables and connections. If we introduce these simplifications, the originally very complex system can be described by the set of nonlinear ordinary differential equations – see e.g. (Russell and Denn 1972) or (Vojtesek et al. 2017):

$$\begin{aligned}
\frac{dc_A}{dt} &= \frac{q_r}{V_r}\left(c_{A0} - c_A\right) - k_1 c_A - k_3 c_A^{\,2} \\
\frac{dc_B}{dt} &= -\frac{q_r}{V_r} c_B + k_1 c_A - k_2 c_B \\
\frac{dT_r}{dt} &= \frac{q_r}{V_r}\left(T_{r0} - T_r\right) - \frac{h_r}{\rho_r c_{pr}} + \frac{A_r U}{V_r \rho_r c_{pr}}\left(T_c - T_r\right) \qquad (2) \\
\frac{dT_c}{dt} &= \frac{1}{m_c c_{pc}}\left(Q_c + A_r U\left(T_r - T_c\right)\right), \\
&\text{where } 0 \le c_A,\ 0 \le c_B
\end{aligned}$$

This set is derivate with the help of material and heat balances inside the reactor. Variable t in the set of Ordinary Differential Equations (ODE) (2) denotes the time, c are concentrations, T represents temperatures, c_p is used for specific heat capacities, q_r means volumetric flow rate of the reactant, Q_c is heat removal of the cooling liquid, V are volumes, ρ stands for densities, A_r is the heat exchange surface and U is the heat transfer coefficient. Indexes $(\bullet)_A$ and $(\bullet)_B$ belong to compounds A and B, respectively, $(\bullet)_r$ denotes the reactant mixture, $(\bullet)_c$ cooling liquid and $(\bullet)_0$ are feed (inlet) values.

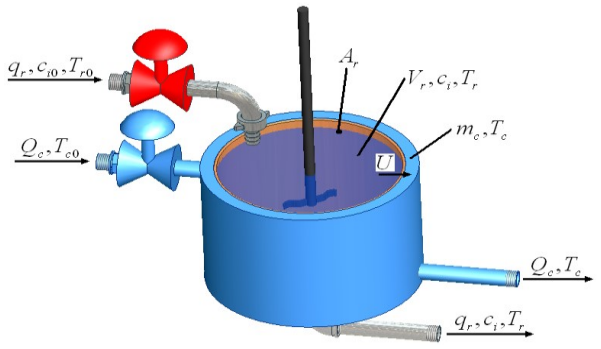


Figure 1: Continuous stirred tank reactor with cooling in the jacket

This reactor belongs to the class of *lumped-parameter nonlinear systems*, see e.g. Ingham et al. (2000). Nonlinearity can be found in reaction rates (k_j), which are described via the Arrhenius law:

$$k_j\left(T_r\right) = k_{0j} \cdot \exp\left(\frac{-E_j}{RT_r}\right), \text{ for } j = 1,2,3 \qquad (3)$$

where k_0 represent pre-exponential factors and E are activation energies.

The reaction heat (h_r) in Eq. (2) is expressed as:

$$h_r = h_1 \cdot k_1 \cdot c_A + h_2 \cdot k_2 \cdot c_B + h_3 \cdot k_3 \cdot c_A^2 \qquad (4)$$

where h_i means reaction enthalpies.

The initial conditions for the set of ODE (2) are

$$c_A\left(0\right) = c_A^s,\ c_B\left(0\right) = c_B^s,\ T_r\left(0\right) = T_r^s,\ T_c\left(0\right) = T_c^s \qquad (5)$$

The mathematical model of the system described by the set of ODE in Eq. (2) shows that this model has four state variables: $c_A(t)$, $c_B(t)$, $T_r(t)$ and $T_c(t)$. From the control point of view, several input variables can be used, e.g. input concentration of compound A, c_{A0}, input temperature of the reactant, T_{r0}, etc. However, the physical viability of these variables is greatly limited from the practical point of view. That is why are simulation studies mainly focused on the volumetric flow rate of the reactant q_r and the heat removal of the cooling liquid Q_c. The change of both quantities can be practically represented for example by the turn of the valve on the inlet pipe, or by the speed of the pump.

Fixed parameters of CSTR are given in Table 1.

Table 1: Parameters of CSTR

Name of the parameter	**Symbol and value of the parameter**
Volume of the reactor	$V_r = 0.01\ m^3$
Density of the reactant	$\rho_r = 934.2\ kg.m^{-3}$
Heat capacity of the reactant	$c_{pr} = 3.01\ kJ.kg^{-1}.K^{-1}$
Weight of the coolant	$m_c = 5\ kg$
Heat capacity of the coolant	$c_{pc} = 2.0\ kJ.kg^{-1}.K^{-1}$
Surface of the cooling jacket	$A_r = 0.215\ m^2$
Heat transfer coefficient	$U = 67.2\ kJ.min^{-1}m^{-2}K^{-1}$
Pre-exponential factor for reaction 1	$k_{01} = 2.145 \cdot 10^{10}\ min^{-1}$
Pre-exponential factor for reaction 2	$k_{02} = 2.145 \cdot 10^{10}\ min^{-1}$
Pre-exponential factor for reaction 3	$k_{03} = 1.5072 \cdot 10^{8}\ min^{-1}.kmol^{-1}$
Activation energy of reaction 1 to R	$E_1/R = 9758.3\ K$
Activation energy of reaction 2 to R	$E_2/R = 9758.3\ K$
Activation energy of reaction 3 to R	$E_3/R = 8560\ K$
Enthalpy of reaction 1	$h_1 = -4200\ kJ.kmol^{-1}$
Enthalpy of reaction 2	$h_2 = 11000\ kJ.kmol^{-1}$
Enthalpy of reaction 3	$h_3 = 41850\ kJ.kmol^{-1}$
Input concentration of compound A	$c_{A0} = 5.1\ kmol.m^{-3}$
Input temperature of the reactant	$T_{r0} = 387.05\ K$

STATIC AND DYNAMIC ANALYSES

Once we have mathematical model of the system, we can make simulation experiments that help with the understanding of the system's behaviour. Also, we can use this knowledge in the design of the controller which will be also described later in the Adaptive control section.

Steady-State Analysis

The steady-state analysis as the first step means that we want to know value of state variables, in our case concentrations c_A, c_B and temperatures T_r, T_c in so called steady-state. The mathematical meaning of this claim is the derivatives with respect to time in the set of ODE (2) are set to zero. It means that the set of ODE (2) is transformed to the set of nonlinear algebraic equations

$$
\begin{aligned}
c_A^s &= \frac{-\left(\frac{q_r}{V_r}+k_1\right) \pm \sqrt{\left(\frac{q_r}{V_r}+k_1\right)^2 - \left(4\cdot k_3 \cdot \left(-\frac{q_r}{V_r} c_{A0}\right)\right)}}{2\cdot k_3}; \\
c_B^s &= \frac{k_1 \cdot c_A^s}{k_2 + \frac{q_r}{V_r}}; \\
T_r^s &= \frac{\frac{q_r}{V_r} T_{r0} - \frac{h_r}{\rho_r \cdot c_{pr}} + \frac{U \cdot A_r}{\rho_r \cdot c_{pr} \cdot V_r} T_c^s}{\frac{q_r}{V_r} + \frac{U \cdot A_r}{\rho_r \cdot c_{pr} \cdot V_r}}; \\
T_c^s &= \frac{Q_c}{U \cdot A_r} + T_r
\end{aligned}
\tag{6}
$$

That can be solved numerically for example with the use of simple iteration method. We can observe the steady-state behaviour for various input variables.
Results for various values of volumetric flow rate of the reactant, q_r, and heat removal of coolant, Q_c, are shown in Figure 2.

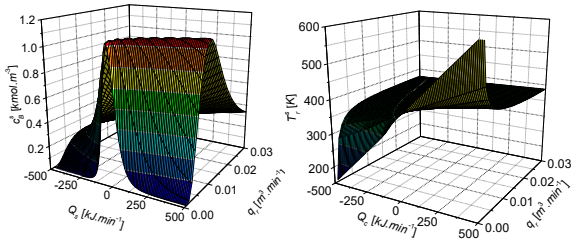


Figure 2: Steady-state analysis for various volumetric flow rate of the reactant, q_r, and heat removal of the cooling, Q_c

We can read from graphs, that this system has strongly nonlinear behaviour. The optimal working point can be represented by the combination of the volumetric flow rate of the reactant $q_r^s = 2.365 \cdot 10^{-3}\ m^3.min^{-1}$ and the heat removal $Q_c^s = -18.56\ kJ.min^{-1}$.

The dynamic analysis and the control is then performed around this working point where steady-state values of state variables are

$$
\begin{aligned}
c_A^s &= 2.1403\ kmol.m^{-3}, \quad c_B^s = 1.0903\ kmol.m^{-3} \\
T_r^s &= 387.34\ K, \qquad\quad T_c^s = 386.06\ K
\end{aligned}
\tag{7}
$$

Dynamic Analysis

Once we have optimal working point from the steady-state analysis, we can continue with the dynamic analysis which means observing of the system's behaviour after the step change of the input variable. In our case, we have chosen the step changes of the coolant's heat removal, ΔQ_c, because this input will be than used as an action value for the control.
Investigated output variables are output concentration of the product B, $c_B(t)$, and output temperature of the coolant, $T_r(t)$. Both values are related to their steady-state values in (7) because we want to display these output from zero and as we can see in (5), these values are initial values in the numerical solution. Input and output variables are then:

$$
\begin{aligned}
u(t) &= \frac{Q_c(t) - Q_c^s}{Q_c^s} \cdot 100\,[\%] \\
y_1(t) &= c_B(t) - c_B^s\left[kmol.m^{-3}\right] \\
y_2(t) &= T_r(t) - T_r^s\,[K]
\end{aligned}
\tag{8}
$$

Mathematically, the dynamic analysis means numerical solution of the set of ODE (2) together with (3) and (4). This numerical solution can be easily performed with build-in functions in Matlab or other mathematical software. Results are shown in Figure 3.

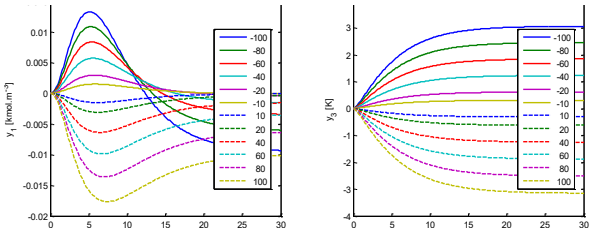


Figure 3: Results of the dynamic analysis for various step changes of input variable $u(t)$

Both courses of output variables $y_1(t)$ and $y_2(t)$ shows nonlinearity of the system which is obvious mainly for the output $y_1(t)$. On the other hand, output $y_1(t)$ can be expressed by the second order transfer function

$$
G(s) = \frac{b(s)}{a(s)} = \frac{b_1 s + b_0}{a_2 s^2 + a_1 s + a_0} \tag{9}
$$

This output will be used as a controlled output in the control section of this paper.

ROBUST CONTROL

Models with Parametric Uncertainty

Systems with parametric uncertainty represent an effective and popular way of considering the uncertainty in the mathematical model of a real plant, see e.g. (Barmish 1994) or (Matušů and Prokop 2013; 2014). The utilization of such models supposes known structure (and order) of the transfer function but not precise knowledge of real parameters, which can be bounded by intervals with minimal and maximal possible values. They can be described by a transfer function:

$$G(s,q)=\frac{b(s,q)}{a(s,q)} \tag{10}$$

where $b(s,q)$ and $a(s,q)$ denote polynomials in s (Laplace transform) with coefficients depending on q, which is a vector of real uncertain parameters. Typically, this vector is confined by some uncertainty bounding set, which is generally a ball in some appropriate norm. The combination of the uncertain system (e.g. transfer function (10)) with an uncertainty bounding set gives the so-called family of systems, see e.g. (Barmish 1994). A special and frequent case of a system with parametric uncertainty is an interval plant. Its parameters vary independently on each other within given bounds, i.e.:

$$G(s,b,a)=\frac{\sum_{i=0}^{m}\left[b_i^-;b_i^+\right]s^i}{\sum_{i=0}^{n}\left[a_i^-;a_i^+\right]s^i} \tag{11}$$

where $b_i^-, b_i^+, a_i^-, a_i^+$ represent lower and upper limits for parameters of numerator and denominator, respectively.

Control Structure and Design

The 2DOF closed-loop control system with separated feedback and feedforward parts of the controller is depicted in Figure 4. The transfer functions $G(s)$, $C_b(s)$, and $C_f(s)$ represent controlled plant, feedback part of the controller, and feedforward part of the controller, respectively and the signals $w(s)$, $n(s)$, and $v(s)$ are reference, load disturbance, and disturbance signal.

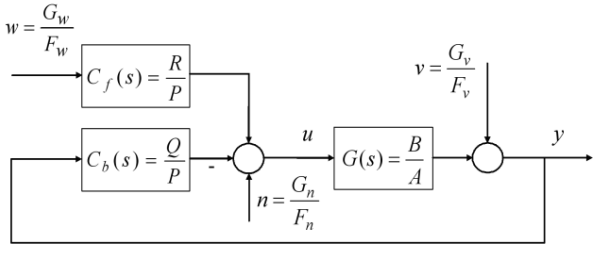


Figure 4: Two-degree-of-freedom control loop

The traditional (one degree of freedom) feedback system is obtained by $R=Q$. However, there are much relevant evidence that the feedforward part brings positive improvements in control responses, see e.g. (Gorez 2003) or (Matušů and Prokop 2013; 2014).

The control synthesis itself is based on the algebraic ideas of Vidyasagar (1985), and Kučera (1993). Subsequently, the specific tuning rules have been developed and analyzed e.g. in (Prokop and Corriou 1997) or (Matušů and Prokop 2013; 2014).

Besides, the controller tuning rules for the case of law order controlled plant under assumption of either purely reference tracking problem or reference tracking and load disturbance rejection together have been already studied e.g. (Kučera 1993) or (Matušů and Prokop 2013; 2014) and so this part presents the important results and then it is applied to the CSTR as a plant with parametric uncertainty.

First, the control design technique supposes the description of linear systems in Fig. 3 by means of the ring of proper and stable rational functions (R_{PS}). The conversion from the ring of polynomials to R_{PS} can be performed very simply – see e.g. (Vidyasagar 1985) or (Prokop and Corriou 1997) according to:

$$G(s)=\frac{b(s)}{a(s)}=\frac{\frac{b(s)}{(s+m)^n}}{\frac{a(s)}{(s+m)^n}}=\frac{B(s)}{A(s)}, \tag{12}$$
$$m>0,\ n=\max\left\{\deg(a),\deg(b)\right\}$$

The parameter $m>0$ will be later used as a controller-tuning knob. The value of the tuning knob has a relevant influence on the control behavior of control responses.

The algebraic analysis (Prokop and Corriou 1997; Matušů and Prokop 2013; 2014) leads to the first Diophantine equation:

$$A(s)P(s)+B(s)Q(s)=1 \tag{13}$$

with a general solution $P(s)=P_0(s)+B(s)T(s)$, $Q(s)=Q_0(s)-A(s)T(s)$, where $T(s)$ is an arbitrary member of (the ring) R_{PS} and the pair $P_0(s)$, $Q_0(s)$ represents any particular solution of (13). Since the feedback part of the controller is responsible not only for stabilization but also for disturbance rejection, the convenient controller from the set of all stabilizing ones can be chosen on the basis of divisibility conditions. The requirement of the reference tracking is obtained by the second Diophantine equation (see Kučera, 1993, Matušů and Prokop, 2013, 2014):

$$F_w(s)Z(s)+B(s)R(s)=1 \tag{14}$$

Robust Stability

The stability of the feedback loop is a crucial requirement in all control applications. Naturally, the feedback loop can be stable when the controlled and/or control plant is unstable. In the case of uncertainty of controlled plants, robust stability means that not only one fixed closed-loop system is stable but also the whole

family of closed-loop control systems is ensured to be stable. Details can be found in e.g. (Ackermann 1993; Barmish 1994; Bhattacharyya et al. 1995; Matušů and Prokop 2011; 2013; 2014). This paper utilizes the robust stability tests based on a universal tool known as the value set concept in combination with the zero exclusion condition – see e.g. (Barmish 1994) or (Matušů and Prokop 2011).

ADAPTIVE CONTROL

The adaptive approach in this work is based on the recursive identification of the linearized model described by the transfer function (9) during the control. The control scheme is very similar to 2DOF control configuration in Figure 4 but block *G* is in this case mathematical model of the controlled system, in our case the set of ODE in (2).

The control synthesis employs pole-placement method together with the spectral factorization. Our previous experiments (for example (Vojtěšek and Dostál 2005; 2016)) have shown, that this method produces sufficient control results.

This control synthesis is based on the solution of the set of Diophantine equations

$$\begin{aligned} a(s)f(s)\tilde{\quad}\quad\quad)q(s) &= d(s) \\ t(s)f_w(s)+b(s)r(s) &= d(s) \end{aligned} \tag{15}$$

where polynomials $a(s)$ and $b(s)$ are polynomials from the transfer function (9) and they are estimated recursively with the Ordinary recursive least-squares method (Bobál et al. 2005). Polynomial $t(s)$ is auxiliary polynomial and unknown controller's polynomials $p(s)$, $q(s)$ and $r(s)$ are computed from (16).

Unknown stable polynomial $d(s)$ on the right side of equations (16) was designed with the use of pole-placement method, e.g. this polynomial is generally

$$d(s) = \prod_{i=1}^{\deg d(s)} (s+s_i) \tag{16}$$

where $s_i = \alpha_i + \omega_i j$ are roots of the polynomial and choice of these roots affects control results. More details about this method can be found for example in (Vojtěšek and Dostál 2005).

SIMULATIONS AND DISCUSSION

A Robust Approach

The CSTR was identified in (Vojtěšek et al. 2017) as a second order system with the transfer function (9) with nominal parameters: $a_2 = 1$, $a_1 = 1.4550$, $a_0 = 0.3072$, $b_1 = -0.0037$, $b_0 = -0.0095$. The intervals for uncertain perturbations were obtained by deeper analysis of the dynamic behavior and they result in the following ones:

$$\begin{aligned} a_0 &= [0.24576;\ 0.36864], \\ a_2 &= [0.8;\ 1.2],\ a_1 = [1.164;\ 1.746], \\ b_1 &= [-0.00296;\ -0.00444],\ b_0 = [-0.0076;\ -0.0114] \end{aligned} \tag{17}$$

Three 2DOF controllers have been designed for the nominal plant and the tuning parameters. The firet one was generated for $m = 0.5$, the second one for $m = 0.8$ and the third one for $m = 1.2$. The feedback and feedforward parts of the controller for the first one is:

$$\begin{aligned} C_b(s) &= \frac{q_2 s^2 + q_1 s + q_0}{s^2 + p_1 s} = \frac{-61.3653 s^2 - 39.7878 s - 6.5789}{s^2 + 0.3179 s} \\ C_f(s) &= \frac{r_2 s^2 + r_1 s + r_0}{s^2 + p_1 s} = \frac{-26.3158 s^2 - 26.3158 s - 6.5789}{s^2 + 0.3179 s} \end{aligned} \tag{18}$$

Figure 5 and Figure 6 show the controlled and control variables for all three tuning parameters. The red lines depict the nominal plant responses and black shadows are responses for the whole uncertain family (17), represented by 3^5=243 members (three values for each interval parameter: minimum, midpoint, and maximum). The load disturbance $n = 10$ was injected in the time $t = 150$ and it is evident that no permanent error is observed.

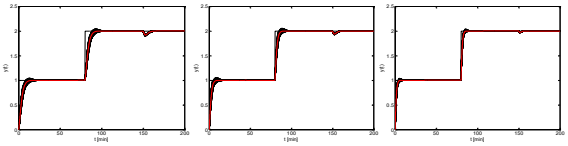


Figure 5: Set of output controlled variables for m=0.5 (left), m=0.8 (middle), and m=1.2 (right)

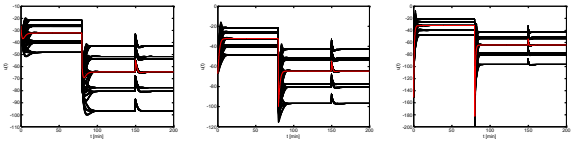


Figure 6: Set of input control variables for m=0.5 (left), m=0.8 (middle), and m=1.2 (right)

Simulation results proved that the fix robust controller could be designed for a wide family of interval systems. The results are shown in Figures 5 and 6 for three values of the tuning parameter m>0. The choice of the tuning parameter m>0 was found empirically and experimentally. Until now, there is no exact theory on how to obtain the optimal value (see e.g. Prokop and Corriou, 1997). The Figure 7 shows the zoomed value sets for all three values of m. All three subfigures from Figure 7 may seem the same for the first sight, but please note the differences in axes ranges. Anyway, they confirm the robust stability of the designed control loops since they are excluded from the critical point (0,0j) and all required preconditions are fulfilled (Barmish 1994).

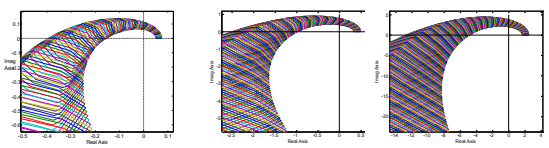


Figure 7: Zoomed value sets for m=0.5 (left), m=0.8 (middle) and m=1.2 (right)

In order to verify the practical usability of the designed controllers, they were applied not only to the linearized model, but also to the original nonlinear model of CSTR. The control results for this nonlinear case are shown and mutually compared in Figure 8.

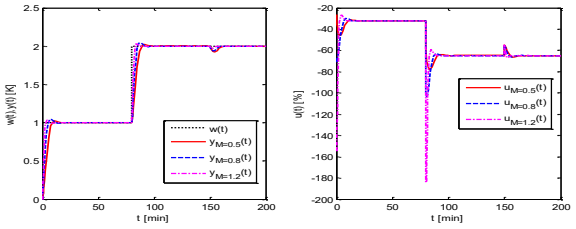


Figure 8: Robust control of the original nonlinear model for three values of m – comparison of the output controlled variables (left) and the input control variables (right)

The control results shown in Figure 8 assumes that there is no limitation of the control signal. On the other hand, Figure 9 provides the control behavior for the same controllers, but with the saturated control signals in the range ±100 %. It can be seen that this saturation affects the signals for m=0.8 and m=1.2. Higher peaks caused by the wind-up effect are observable for m=1.2.

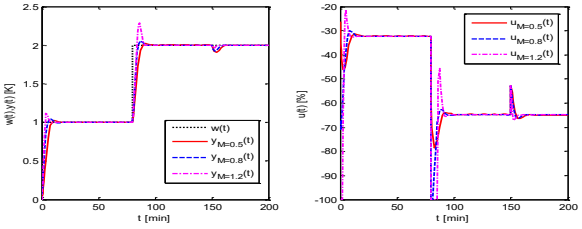


Figure 9: Robust control of the original nonlinear model for three values of m and the saturated control signal (±100 %) – comparison of the output controlled variables (left) and the input control variables (right)

An Adaptive Approach

Three adaptive controllers for 2DOF configuration were tuned, assuming the placement of the closed-loop poles 0.07, 0.1, and 0.2, respectively. Figure 10 shows the control results for the original nonlinear CSTR model, and Figure 11 presents the evolution of the identified parameters during the simulation.

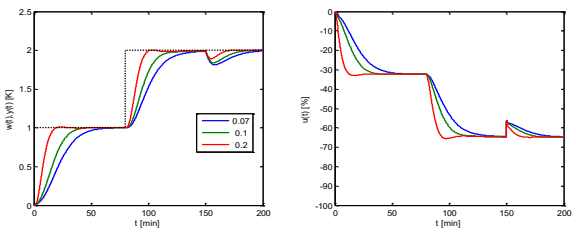


Figure 10: Adaptive control results – output (left) and control (right) signals

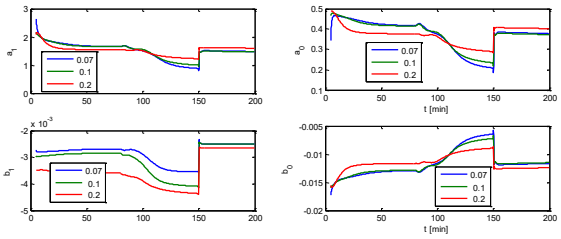


Figure 11: Adaptive control results – identified parameters

Comparison and Discussion

The control performance of both robust and adaptive approaches can be tuned by the parameter m or the proper pole-placement. In all cases, the costs for the rapid control and better disturbance rejection are the higher and more aggressive control signals. For some faster robust controllers, the control signals would have to be restricted for the practical application. The main advantage of the self-tuning controllers is obvious from its name, i.e., after successful initialization, they are able to control the CSTR without knowledge of the model. On the other hand, the main advantage of the off-line tuned robust controllers is their simplicity and reliability, even under prescribed model uncertainty. The comparison of control results for two reasonable choices of tuning parameters, i.e., m=0.5 for the robust controller, and α=0.2 for the adaptive controller, are shown in Figure 12. Anyway, it was shown that both approaches are able to control the CSTR satisfactorily and the final choice of the approach depends on the additional requirements or preferences of a user or control engineer.

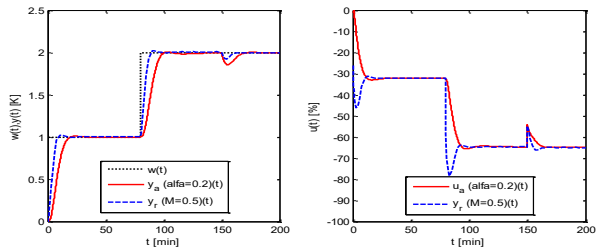


Figure 12: Comparison of selected robust and adaptive controllers – output (left) and control (right) signals

CONCLUSIONS

Modelling and control of CSTR are addressed in the contribution. Two different approaches of control were designed and compared. The first one is an adaptive self-tuning principle based on the recursive identification procedure with polynomial control design. The second control principle utilizes robust control algorithms designed in the ring R_{PS}. The synthesis method itself is based on linearized model with parametric uncertainty and accompanied by the analysis of robust stability. Both approaches use the 2DOF

feedback control structure. As an application, a set of designed robust and adaptive controllers were applied to control of an original nonlinear model of CSTR. The main aim of the control design was energy saving in the industry operation of CSTR. All simulations were performed in the MATLAB and Simulink environment.

REFERENCES

Ackermann, J. 1993. *Robust control – systems with uncertain physical parameters*. Springer-Verlag London, Great Britain.

Åström, K.J., Wittenmark, B., 1989. Adaptive Control. Addison Wesley. Reading. MA, USA. ISBN: 0-201-09720-6.

Barmish, B. R., 1994. *New Tools for Robustness of Linear Systems*, New York, USA: Macmillan.

Bhattacharyya, S.P., Chapellat, H., Keel, L.H., 1995. Robust Control-The Parametric Approach. Prentice Hall.

Bobál, V., Böhm, J., Fessl, J., Macháček, J. 2005. Digital Self-tuning Controllers. Algorithms, Implementation and Applications. Springer 2005.

Gorez, R., 2003. New design relations for 2-DOF PID-like control systems. *Automatica*, vol. 39, no. 5, pp. 901-908,

Grimble, M. J., 2006. Robust Industrial Control Systems. Optimal Design Approach for Polynomial Systems. Prentice Hall. USA. ISBN: 0-470-02073-3

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, Germany. ISBN: 3-527-29776-6

Kučera, V., 1993. Diophantine Equations in Control – A survey. Automatica. 29, 1993, 1361-1375.

Matušů, R., Prokop, R., 2011. Graphical analysis of robust stability for systems with parametric uncertainty: an overview. *Transactions of the Institute of Measurement and Control*, Vol. 33, No. 2, pp. 274-290.

Matušů, R., Prokop, R., 2013. Algebraic Design of Controllers for Two-Degree-of-Freedom Control Structure. *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 7, no. 6, pp. 630-637.

Matušů, R., Prokop, R., 2014: An Algebraic Approach to Two-Degree-of-Freedom Controller Design for Systems with Parametric Uncertainty. *International Journal of Mathematical Models and Methods in Applied Sciences*, Vol. 8, pp. 113-120, ISSN 1998-0140.

Middleton, H., Goodwin, G. C., 2004. Digital Control and Estimation - A Unified Approach. Prentice Hall. Englewood Cliffs, USA. ISBN: 0-13-211798-3

Prokop, R., Matušů, R., Vojtěšek, J., 2019. Robust Control of Continuous Stirred Tank Reactor with Jacket Cooling. *Chemical Engineering Transactions*, Vol. 76, 2019, pp. 787-792, ISSN 2283-9216.

Russell, T., Denn, M. M., 1972. Introduction to Chemical Engineering Analysis. New York: Wiley, USA, xviii, ISBN: 04-717-4545-6.

Vidyasagar, M., 1985. Control system synthesis: A factorization approach, Cambridge, Massachusetts, USA: MIT Press.

Vojtěšek, J., Dostál, P., 2016. Continuous-time vs. discrete-time identification models used for adaptive control of nonlinear process. In: Proceedings - 30th European Conference on Modelling and Simulation, ECMS 2016 [online]. European Council for Modelling and Simulation (ECMS), 2016, s. 320-326.

Vojtěšek, J., Prokop, R., Dostál, P., 2017. Two Degrees-of-Freedom Hybrid Adaptive Approach with Pole-placement Method Used for Control of Isothermal Chemical Reactor. *Chemical Engineering Transactions*, 2017, Vol. 2017, No. 61, pp. p1-p7. ISSN 2283-9216.

AUTHOR BIOGRAPHIES

ROMAN PROKOP was born in Hodonín, Czech Republic in 1952. He graduated in Cybernetics from the Czech Technical University in Prague in 1976. He received post graduate diploma in 1983 from the Slovak Technical University. Since 1995 he has been at Tomas Bata University in Zlín, where he presently holds the position of full professor of the Department of Automation and Control Engineering and a vice-dean of the Faculty of Applied Informatics. His research activities include algebraic methods in control theory, robust and adaptive control, autotuning and optimization techniques. His e-mail address is: prokop@utb.cz.

RADEK MATUŠŮ received the M.S. degree the Faculty of Technology, Tomas Bata University (TBU) in Zlín, in 2002, and the Ph.D. degree from the Faculty of Applied Informatics (FAI), TBU in Zlín, in 2007. He was appointed an Associate Professor of machine and process control at FAI TBU in Zlín, in 2018. He has been holding various research or pedagogical positions at TBU in Zlín, since 2004, where he is currently a Researcher and a Project Manager. His research interests include analysis and synthesis of robust control systems, fractional-order systems, and algebraic methods in control design. He has (co-)authored more than 50 scientific journal papers and over 110 conference contributions. He serves as an Academic Editor or a Reviewer for dozens of scientific journals.

JIRI VOJTESEK was born in Zlin, Czech Republic. 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 an 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@utb.cz.

ROBUST SIMULATION OF IMAGING MASS SPECTROMETRY DATA

Anastasia Sarycheva
Skolkovo Institute of Science and Technology
Bolshoy Boulevard 30, bld. 1,
Moscow 121205, Russia
E-mail: sarycheva.anastasia@gmail.com

Anton Grigoryev
Institute for Information Transmission Problems, RAS
Bolshoy Karetny per. 19,
Moscow, 127051, Russia
E-mail: me@ansgri.com

Evgeny N. Nikolaev
Skolkovo Institute of Science and Technology
Bolshoy Boulevard 30, bld. 1,
Moscow 121205, Russia
E-mail: e.nikolaev@skoltech.ru

Yury Kostyukevich
Skolkovo Institute of Science and Technology
Bolshoy Boulevard 30, bld. 1,
Moscow 121205, Russia
E-mail: y.kostyukevich@skoltech.ru

KEYWORDS

Mass spectrometry imaging, simulation, instrument response function.

ABSTRACT

Mass spectrometry imaging (MSI) with high resolution in mass and space is an analytical method that produces distributions of ions on a sample surface. The algorithms for preprocessing and analysis of the raw data acquired from a mass spectrometer should be evaluated. To do that, the ion composition at every point of the sample should be known. This is possible via the employment of a simulated MSI dataset. In this work, we suggest a pipeline for a robust simulation of MSI datasets that resemble real data with an option to simulate the spectra acquired from any mass spectrometry instrument through the use of the experimental MSI datasets to extract simulation parameters.

INTRODUCTION

High-resolution mass spectrometry is an analytical technique based on the precise measurement of mass-to-charge ratio (*m/z*) of ionized molecules found in a sample and their relative amount. The mass spectrometry (MS) experiment includes the following main steps: sample preparation, ionization, ion separation (employing electric and magnetic fields), ion detection and signal processing. The result of such an analysis is represented as the so-called mass spectrum (see an example of a profile mass-spectrum in Figure 1: in profile mode, a peak is represented by a collection of signals over several MS experiments) where the signal intensities (i.e. the relative number of ions with certain *m/z*) are plotted as y-axis versus corresponding mass-to-charge ratios along x-axis. Mass spectrum is used to determine the compounds of the sample. For each compound information on molecular mass, composition and structure can be derived through the analysis of experimental spectra. This makes mass spectrometry an essential technique utilized in many applied and basic sciences such as Chemistry, Biology, Medicine, Ecology, Forensic science, etc. (De Hoffmann, 2000)

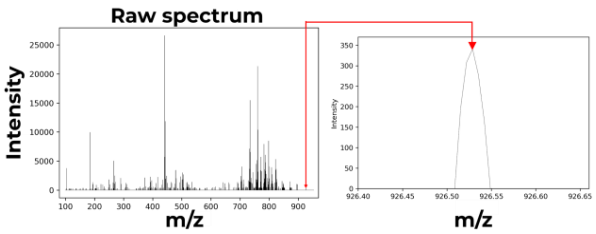


Figure 1: A raw (not preprocessed) profile mass spectrum acquired via MALDI-Orbitrap mass spectrometer (Thermo Scientific Q Exactive Orbitrap) collected from a single region ($35 \times 35\ \mu m^2$) of a mouse full body section. This mass spectrum includes 4934 individual *m/z* with corresponding intensities. Zoomed peak 926.529 *m/z* illustrates a typical peak shape — Gaussian.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

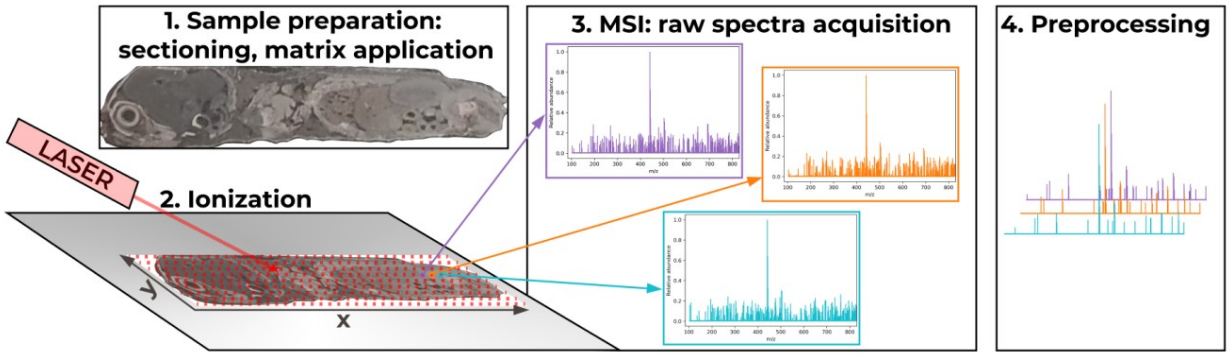


Figure 2: MALDI MSI experiment (sample: mouse full body section).

An MS instrument, mass spectrometer, includes an ion source, a mass analyzer, and a detector (De Hoffmann, 2000). The ion source produces ions of the analyzed sample, the mass analyzer separates these ions according to their *m/z*, and the detector counts the number of ions for every *m/z* bin (the detector aggregates the continuous *m/z* values produced by the analyzer into discrete *m/z* bins, the binning depends on the instrument and its settings).

Mass spectrometry imaging (MSI) is the sequential mass spectrometry analysis of the regions on the surface of the sample. Based on this information, the spatial distributions of the detected ions on the sample's surface are generated. MSI is commonly used in diagnostic applications in the medical and biomedical field (e.g. abnormal regions detection such as tumors, biomarkers search), in medicinal chemistry (medicinal drugs development, research of drugs and their metabolites localization in tissues) (Römpp and Spengler 2013).

There are many ionization techniques developed for MSI, but in this work, we will briefly describe the most popular one: Matrix-Assisted Laser Desorption/Ionization (MALDI) (Baker et al. 2017). The matrix (typically an organic acid) is chosen by the researcher based on the analytes (the compounds of interest to be ionized and detected). The matrix co-crystallizes the analytes, fixing them in place, and facilitates the ionization process. During the desorption/ionization stage, the laser simultaneously vaporizes and ionizes the region it is directed towards, covering the surface of the sample with the given step (raster step). Thus, for each raster (which represents a pixel on the resulting spatial ion distributions) of the sample surface, a mass spectrum (which is a set of detected ions with corresponding signal intensities) is acquired. The MALDI MSI experiment workflow is illustrated in Figure 2. Preprocessing of raw MSI spectra is necessary as the amount of the detected ions for each region is too large for high resolution in mass and space MS instruments (Römpp and Spengler 2013). Preprocessing algorithms should reduce the size of raw MSI dataset, remove noise, eliminate inaccuracies, and make mass spectra from different regions comparable. These algorithms (and/or parameters for them) have to be evaluated, which poses a question of ground truth data in MSI. Due to the complicated nature of the MSI data acquisition, there is no way to get ground truth data for the samples, i.e. it is not possible to know the exact ion composition at each region of the sample in order to compare it to the ion composition revealed by preprocessed raw spectra. Thus, in order to evaluate the preprocessing algorithms, these algorithms are applied to simulated MSI datasets (Palmer 2014; Verbeek 2014; Wijetunge et al. 2015; Guo et al. 2019; Lieb et al. 2020; Booij 2021).

But even after successful preprocessing steps, the number of individual ion distributions is large. Preprocessed MSI data can be treated as a multichannel image (similarly to an optical image taken with a usual camera; such an image has three channels: red, green, and blue, each channel representing the corresponding light wavelength and its intensity in each part of an image), where each channel represents a certain mass-to-charge ratio (*m/z*) and its intensity at each raster of the surface, i.e. each channel is a visualization of a single ion distribution. So preprocessed MSI data are organized in so-called data cubes (see Figure 3).

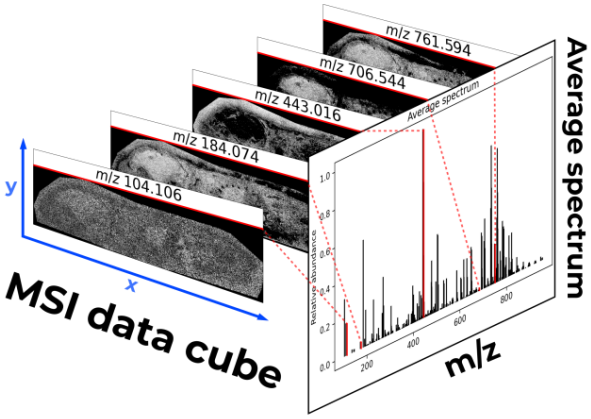


Figure 3: MSI data cube (sample: mouse full body section) which includes spatial distributions of 417 individual ions (*m/z*). Average spectrum is centroid: the signals are displayed as discrete *m/z* with corresponding intensities with zero line widths.

Various feature selection, clustering and visualization algorithms have to be applied to MSI data cubes, in order to perform a sufficient analysis of such data. These algorithms also have to be evaluated, which again requires realistically simulated data (Buchberger et al. 2018; Verbeek et al. 2020, Sarycheva et al., 2020).

While there is a publically available MSI simulation functions for R language (Bemis et al. 2015; Bemis and Harry 2017; Bemis 2020), which will be discussed in detail in the section dedicated to Simulation of imaging mass spectrometry data, they do not offer as much flexibility as the proposed pipeline.

The important feature of the proposed approach is that it allows for simulation of MSI spectra acquired from an individual instrument. This reduces the gap between evaluation of preprocessing and analysis algorithms performances for the experimental MSI dataset acquired by a certain instrument and the simulated dataset for the same instrument.

SIMULATION IN MASS SPECTROMETRY

Mass spectrometry data processing and analysis are not possible without computational methods. These methods should be validated, evaluated and compared to guarantee the credibility of the acquired results as the chosen algorithms and/or parameters have been applied to a MS dataset. To gauge the performance of the algorithms (and/or tune the parameters), the reference results (benchmark) should be determined. While it could be argued that the output of a trusted method (or results amalgamation of multiple trusted methods) might be picked as a reference, such an approach is incorrect for the following reasons. First, the result of the existing methods application might be inaccurate (especially, due to the nature of MS data — there are no flawless tools or data processing pipelines). Moreover, a new method might disagree with the results of the existing ones because it might correct the underlying bias of the latter. Thus, it is not correct to expect the result of a new method to replicate the result of the state-of-the-art methods. Second, such an evaluation of a new method would depend on the set of trusted methods chosen to generate reference. So if the set of trusted methods would be changed, even if the input raw data remains the same, the evaluation conclusions might change, too. Thus, the benchmark should not depend on the compared methods.

Realistic data simulation, where the ground truth is defined, serves as a way to create the reference. For the evaluation of raw data processing, e.g. peak picking, feature detection, raw MS data are simulated (e.g., Schulz-Trieglaff et al. 2008; Bielow et al. 2011; Wijetunge et al. 2015). To evaluate the algorithms for the analysis, the processed MS data are simulated (e.g., Awan and Saeed 2018). The main advantages of simulated data: 1) the ground truth is defined, which is often cannot be achieved in MS experiment due to the competitive ionization processes during the ionization stage, instrument noise, and etc; 2) a lot of datasets can be simulated, represented data from different instruments and from different samples; the acquisition of various experimental MS datasets might be expensive or difficult (Gatto et al. 2016).

The simulation of individual spectra can be based on a mathematical model which describes physics of an instrument, if there are analytical expressions describing the latter. For example, the realistic mass spectra can be simulated for Time of Flight (TOF) mass spectrometers (Coombes et al. 2005): the flight time of a given ion of mass *m* and charge *z* (known for a range of proteins found in biological tissue or fluids) in an electric field is simulated given the parameters describing the virtual MS instrument (Morris et al. 2005) and the amount of detected ions. This simulation model accounts for two factors affecting the mass resolution (the ability of the instrument to provide a mass spectrum where two slightly different masses are distinguishable): the acquisition time resolution of the detector and the distribution of the initial velocities of the ions. The isotope distributions of individual proteins are included in simulation, since proteins mostly consist of the atoms of carbon, oxygen, and nitrogen. To sum up, this simulation approach employs Instrument Response Function (IRF) calculated from physical laws which result in an approximation for a virtual TOF mass spectrometer.

However, a creation of a detailed physical model of mass spectra generation is not possible for every MS instrument. In a simulation tool LC-MSsim (implemented in C++ programing language) for liquid chromatography mass spectrometry (LC–MS) data (Schulz-Trieglaff et al. 2008), the peak shape of an input *m/z* is modelled using a Gaussian distribution. In this simulation, the peak width is chosen by a user in terms of the Full-Width-At-Half-Maximum (FWHM) of a peak in mass spectrum, which is defined as the difference between *m/z* at which the intensity equals half of the maximum intensity of this peak. Since peak shape is modelled as a Gaussian, FWHM of a Gaussian is defined as follows:

$$FWHM_G = 2\sqrt{2ln2}\sigma\,, \qquad (1)$$

where σ is the standard deviation of the Gaussian.

Any real MS dataset includes not only signals caused by the ionized compounds present in the sample, but also noise, which should be accounted for in a simulated spectrum. In LC-MSsim, the FWHM of peaks is used to simulate MS instruments with different mass accuracies (mass accuracy is the difference between measured and actual mass) and resolutions. The inaccuracies in measured peak intensities are simulated by adding Gaussian-distributed noise to peaks. The statistical fluctuations found in MS spectra if the measured intensity of ions with certain *m/z* is very low (i.e. high-frequency noise of low intensity in a mass spectrum), so-called shot noise, is not well defined in MS, yet for Q-TOF and Ion Trap instruments it can be modeled by Poisson distribution (Du et al. 2008). So in LC-MSsim, the number of shot noise signals is sampled from a Poisson distribution, while *m/z* are sampled from Gaussian distribution and intensities of these signals are sampled from Exponential distributions (it was approximated based on the experimental mass spectra). The baseline signal in mass spectra (especially prominent within MALDI MS instruments), which decays with increasing *m/z*, in LC-MSsim is simulated by adding an

exponentially-decaying baseline to a simulated mass spectrum.

A simulation tool MSSimulator (implemented in C++ programing language) for LC–MS and LC–MS/MS (MS/MS is tandem mass spectrometry, where the selected ions, separated by their *m/z* in MS experiment, are split into smaller fragment ions, and then these fragments are also separated and detected by MS experiment) data (Bielow et al. 2011), uses either a truncated Gaussian or Lorentzian distribution for peaks modeling, the width of peaks can be controlled by a user based on the resolution. MSSimulator also provides three models of resolution models in common instruments: resolution is constant in TOF; resolution is degrading linearly with *m/z* in Fourier transform ion cyclotron resonance (FTICR) instruments; resolution is degrading linearly with the square root of *m/z* in Orbitrap mass spectrometers (Makarov et al. 2006).

A simulation tool Mspire-Simulator (implemented in Ruby programing language) for LC–MS data (Noyce et al. 2013) employs IRFs calculated from experimental data for three different instruments, acquired from LTQ-Orbitrap, Orbitrap-Velos, Bruker MicrOTOF-Q mass spectrometers. These default models can be replaced by models provided by the user which would mimic other settings and/or instruments: the simulation parameters can be acquired from LC–MS files using a genetic curve fitting algorithm.

To summarize, simulation is used in MS field as benchmark data to assess various algorithms, since the creation of annotated MS datasets acquired by various instruments with various settings is complicated and expensive, and publically available experimental datasets are scarce (Wijetunge et al. 2015; Gatto et al. 2016; Awan and Saeed 2018).

SIMULATION OF IMAGING MASS SPECTROMETRY DATA

MSI experiments, being a compilation of multiple MS experiments for various points of a sample surface, take more time and are more expensive than the routine MS experiments. Due to the complicated nature of an MSI dataset, the amount of publically available comprehensibility annotated testing datasets is often insufficient (Palmer 2014). For certain methods, the testing datasets might not be available at all. Thus, the simulated MSI data are used for validation and evaluation of MSI data processing and analysis algorithms (Palmer 2014; Verbeek 2014; Guo et al. 2019; Lieb et al. 2020; Booij 2021).

Input data for the MSI simulation is usually the list of ions (it is used if the ionization process is hard to model; otherwise, the list of ions can be predicted from the list of input compounds) with corresponding spatial distributions, and an output MSI dataset is formed according to a statistical model which corresponds to a desired MS instrument.

Verbeek in (Verbeek 2014) describes the simulation approach to datasets creation employed to benchmark MSI analysis algorithms (Booij 2021). An artificial dataset includes areas representing different tissue regions (and thus having distinct spectral composition) which might overlap: the corresponding characteristic spectra are mixed. Each pixel contains *N m/z* bins in a certain mass range m/z_{min} to m/z_{max}. Characteristic spectra contain certain amounts of peaks with various intensities within mass range. Gaussian noise is added to the mass spectrum of each pixel.

Palmer in (Palmer 2014) describes an IRF modeling approach to MSI spectra generation using QqTOF (Quadrupole-time-of-flight mass spectrometer) instrument as the example. IRFs are approximated via fitting mathematical functions to experimental data, which allows for an approximation of any instrument. The continuous *m/z* values are aggregated into discrete mass bins. Binning is defined by the instrument and settings. It is simulated according to the resolution of the desired virtual instrument, and the input list of ions is mapped to the corresponding bins (and their intensities are summed). The binned *m/z* values with corresponding intensities are worked up by IRFs. The latter mimic signal blurring in mass analysers: intensities of neighboring bins affect the input bin intensity (simulated by Gaussian filter moved along *m/z* axis). IRFs also add detection noise to each input bin (e.g., baseline noise for TOF instruments, electronic noise due to detection circuitry's thermal electron motion sampled from Gaussian distribution, shot noise for all counting detectors, chemical noise which adds the detection of randomly distributed ions on the sample surface).

In (Guo et al. 2019), ion distributions with complex morphology are simulated. The ion spatial variation was simulated as follows. The intensity of an ion at each pixel is generated as a sum of the following terms: the mean intensity of morphological component (i.e. a distinct tissue region) of this pixel, the spatial auto-correlation (simulated via the intrinsic conditional auto-regression (ICAR) model: spatial effect is varying around mean spatial effects at neighboring locations drawn from Normal distribution) which reflects similarity or disagreement in ion composition of neighboring pixels, and the random noise (i.e. measurement error).

Dexter in (Dexter 2018) uses multivariate normal distribution for statistical modeling using experimental MSI data, since he demonstrated that the clustered MSI data from the coronal mouse brain (data acquired via MALDI QqTOF instrument) converted to polar coordinates can be approximated by a multivariate normal distribution. Normality testing (the chi squared quantile plots) is performed for the experimental dataset, and if the latter is close to normally distributed, it is used as a reference for simulation. Simulated data are sampled probabilistically from a multivariate normal distribution.

A simulation of HR imaging mass spectrometry data (ims-simulator) scripts for python 2.7 are available at Github(https://github.com/metaspace2020/ims-simulator) as part of Metaspace project (Alexandrov et al. 2019). The input experimental centroided MSI dataset in imzML format (Schramm et al. 2012) is used as a template for the simulation. Other input data include: the

instrument type (two options: FTICR or Orbitrap); the resolving power (instrument's ability to distinguish between two adjacent ions of equal intensity) at *m/z* = 200; database with the list of metabolites (molecules) as well as the list of possible adducts (the adduct ions are formed during ionization process and contain a certain ion along with analyte molecule (M), e.g. hydrogen ion adducts $[M+H]^+$, sodium ion adducts $[M+Na]^+$, potassium ion adducts $[M+K]^+$, etc.) which might be found in the experiment. This information is used to provide false discovery rate (FDR)-controlled metabolite annotation (Palmer et al. 2016) of the input MSI dataset. This annotation (a list of adducts and molecules) is used to simulate a clean (without noise) dataset. Then the basic statistics for the experimental dataset are calculated (sparsity: histogram of *m/z* differences between neighboring *m/z* in each spectrum (i.e. in each pixel); histogram of intensities; minimum intensities for each spectrum). The input dataset's dimensionality is reduced via non negative matrix factorization (NMF): the amount of components is the input amount of desired layers for the simulated dataset (each layer, a simulated tissue, with the spatial distribution represented by an NMF component and with spectral composition represented by a pseudo-spectrum — the loadings of the corresponding NMF component). Noise parameters (median, standard deviation) for each *m/z* value are calculated from the difference between the experimental data cube's ion intensity distribution and the distribution reconstructed by NMF.

The Cardinal, an R package for MSI data processing and analysis (Bemis et al. 2015), provides functions for the simulation of MS and MSI datasets and which were employed for MSI dataset simulation used for the evaluation of peak picking algorithm in (Lieb et al. 2020), based on the documentation (Bemis and Harry 2017) and the corresponding functions in the package. However, some of these functions were deprecated or changed in the newer version of the package, Cardinal 2 (Bemis 2020). It features a function simulateImage() which relies on simulateSpectrum() for the MSI data simulation. The simulation function input: spatial data (Pixel data) features coordinates x, coordinates y, and boolean columns for each spatial region (morphological substructure, reference image masks) specifying whether or not this region is present in (x,y); ions and intensities (Feature data) featuring *m/z* (ions), and columns for each spatial region (morphological substructure) specifying intensities of ions describing the spectral composition of corresponding spatial regions. Minimum and maximum *m/z* values for simulation mass range, as well as step-size for the observed *m/z* values of the profile spectrum can be specified. There are additional parameters introducing noise and variation (virtual instrument parameters): spatial autocorrelation (for spatial covariance calculation), standard deviation giving the run-to-run and/or pixel-to-pixel variance (sampled from Normal distribution), standard deviation for the distribution of the observed peaks, a multiplier for multiplicative variance, standard deviation of the random noise introduced in the spectrum, standard deviation of the mass error in the observed *m/z* values of peaks, mass resolution, maximum intensity of the baseline and its exponential decay, whether output spectra will be in profile or centroided.

While R language is relatively easy to use, the description of Cardinal 2 simulation functions explicitly reads that they are designed for small proof-of-concept examples, and may not scale well to simulating larger datasets(https://rdrr.io/bioc/Cardinal/man/simulateSpectrum.html).

A perfect simulator should be able to produce datasets of any size with defined ground truth, with the option to mimic various instruments. Ideally — any instrument, if the user provides MSI data. This is achievable through the augmentation of simulation with parameters and noise distribution extracted from the experimental datasets.

THE PROPOSED MSI DATA SIMULATION PIPELINE

If the experimental datasets are provided by the user (in imzML format), they can be used to set mass range and *m/z* values binning, approximate noise parameters for the simulated spectra. Thus, the proposed simulation configuration extraction from the experimental datasets includes the following steps: instrument information (i.e. resolution), mass range and statistics extraction. The scheme for such module is illustrated in Figure 4.

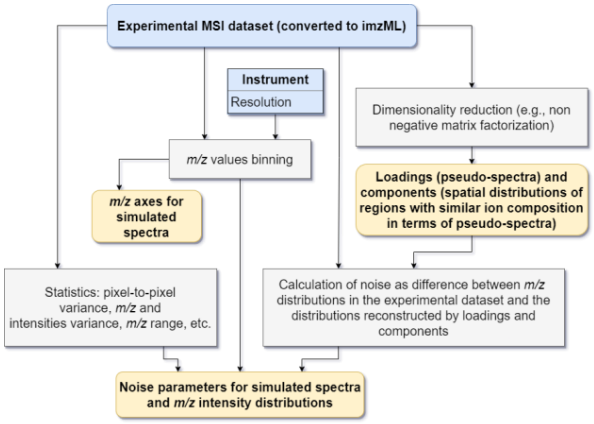


Figure 4: The module for the extraction of simulation parameters from an experimental MSI dataset.

The input morphological components (distinct tissue regions) are provided by the user as separate

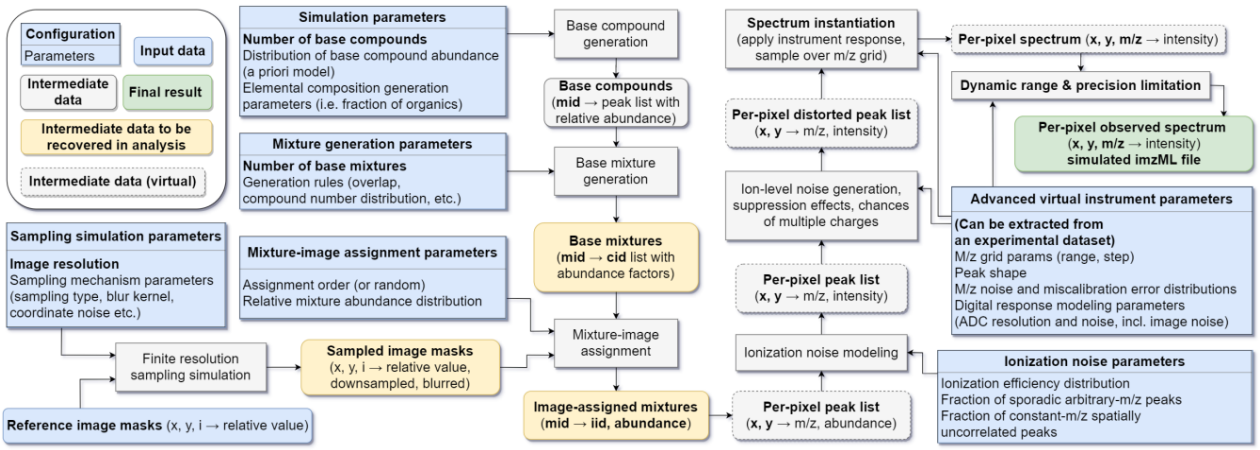


Figure 5: The MSI data simulation pipeline.

grayscale images (could be drawn in with a graphical editor or generated computationally). If the experimental dataset is not provided, the simulation can use preset parameters which correspond to different types of instruments. Parameters can be also set by a user in the configuration file, e.g. analog-to-digital converter resolution, mass range, resolution, etc.

The general steps of the proposed simulation pipeline (parameters can be extracted from the experimental data, see Figure 4) are illustrated in Figure 5.

Let us consider an example of the simulation using the proposed pipeline. We used an experimental MSI dataset of the macaque cerebellum section (22430 single ion distributions 100×80 pixels) as input for the module for extraction of simulation parameters: the m/z values binning was 0.01; the number of ground truth distinct spectral compositions (each corresponds to the simulated tissue) acquired via NMF was set to 6 (Figure 6 A); noise was extracted as the difference between the experimental spectra and the spectra reconstructed by NMF components and loadings. We drew 6 grayscale reference image masks (99×72 pixels each, Figure 6 B) which correspond to 6 overlapping morphological regions with simulated spectral compositions. With added noise, the resulting simulated spectra closely resemble the experimental ones (Figure 6 D). The ground truth data, the exact spectral composition in each spatial location of the simulated imzML, as well as the spectral composition for each reference mask are saved separately.

The main advantages of the proposed pipeline:
1) the ability to produce large artificial datasets in reasonable time; 2) the distinct tissue regions with significantly different spectral composition to be generated (and mixed if overlapped) are provided by the user as simple grayscale images; 3) flexibility of parameters: can be set or extracted from the experimental dataset.

The algorithm is implemented in Python 3.7 and uses the following libraries: imageio, numpy, pandas, pyimzml, scikit-image, sklearn, scipy.

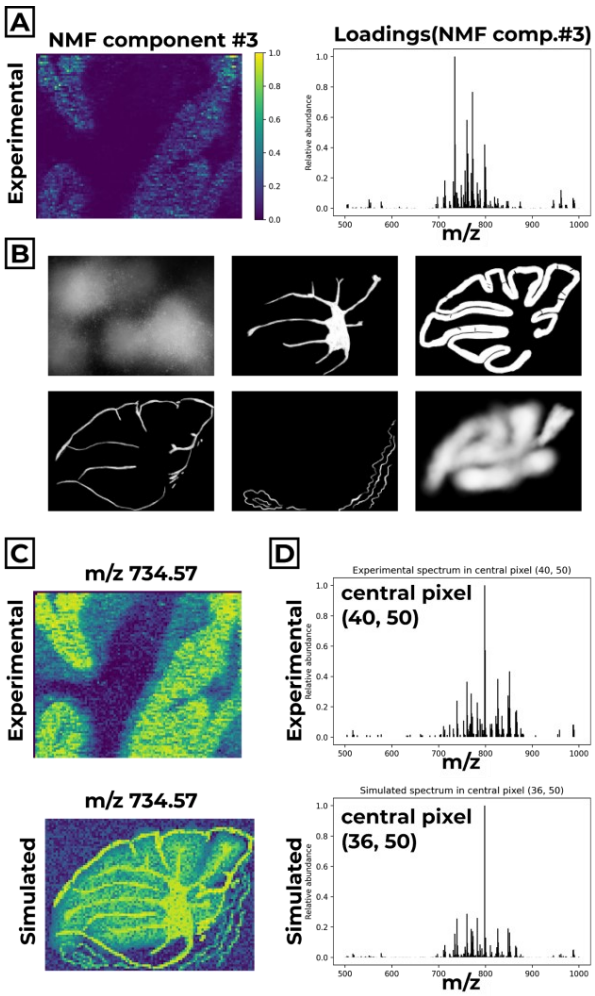


Figure 6: (A) Extraction of simulation parameters: illustration of one of 6 NMF components with corresponding loadings. (B) Input reference image masks (morphological components/tissues to simulate). (C) Comparison of experimental and simulated single ion images. (D) Comparison of experimental spectrum and simulated spectrum in the central pixel of the data cubes.

CONCLUSION

We have found that there is no convenient and universal simulation tool available for large MSI datasets. We proposed a versatile pipeline, which includes the extraction of simulation parameters from the experimental dataset to tailor the simulation to a specific MSI setup. The initial implementation of the proposed pipeline was tested on a high-mass-resolution Orbitrap-based imaging setup as a parameter source with a set of hand-drawn masks as reference images. The implemented algorithm allowed for the generation of large datasets suitable for quantitative testing of algorithms for enhancement, decomposition, and visualization of MSI data.

REFERENCES

Alexandrov, T.; K. Ovchinnikova; A. Palmer; V. Kovalev; A. Tarasov; L. Stuart; ... and S. Shahidi-Latham. 2019. METASPACE: A community-populated knowledge base of spatial metabolomes in health and disease. *BioRxiv*, 539478.

Awan, M.G. and F. Saeed. 2018. MaSS-Simulator: A highly configurable MS/MS simulator for generating test datasets for big data algorithms. *bioRxiv*, 302489.

Baker, T.C.; J. Han; and C.H. Borchers. 2017. Recent advancements in matrix-assisted laser desorption/ionization mass spectrometry imaging. *Current opinion in biotechnology*, 43, 62-69.

Bemis, K.D. and A. Harry. 2017. Cardinal: Analytic tools for mass spectrometry imaging.

Bemis, K.D.; A. Harry; L.S. Eberlin; C. Ferreira; S.M. van de Ven; P. Mallick; M. Stolowitz; and O. Vitek. 2015. Cardinal: an R package for statistical analysis of mass spectrometry-based imaging experiments. *Bioinformatics*, 31(14), 2418-2420.

Bemis, K.A. 2020. Cardinal 2: User guide for mass spectrometry imaging analysis. (http://bioconductor.org/packages/release/bioc/vignettes/Cardinal/inst/doc/Cardinal-2-guide.html#advanced-operations-on-msimagingexperiment)

Bielow, C.; S. Aiche; S. Andreotti; K. Reinert. 2011. MSSimulator: Simulation of mass spectrometry data. *Journal of proteome research*, 10(7), 2922–9.

Booij, T. 2021. Data-Driven Soft Discriminant Maps: Class-aware Linear Feature Extraction in Imaging Mass Spectrometry. (Master thesis, Delft University of Technology)

Buchberger, A.R.; K. DeLaney; J. Johnson; and L. Li. 2018. Mass spectrometry imaging: a review of emerging advancements and future insights. *Analytical chemistry*, 90(1), 240.

Coombes, K.R.; J.M. Koomen; K.A. Baggerly; J.S. Morris; and R. Kobayashi. 2005. Understanding the characteristics of mass spectrometry data through the use of simulation. *Cancer informatics*, 1, 117693510500100103.

De Hoffmann, E. 2000. Mass spectrometry. *Kirk-Othmer Encyclopedia of Chemical Technology.*

Dexter, A. 2018. Developing computational methods for fundamentals and metrology of mass spectrometry imaging (Doctoral dissertation, University of Birmingham).

Du, P.; G. Stolovitzky; P. Horvatovich; R. Bischoff; J. Lim; and F. Suits. 2008. A noise model for mass spectrometry based proteomics. *Bioinformatics*, 24(8), 1070-1077.

Gatto, L.; K.D. Hansen; M.R. Hoopmann; H. Hermjakob; O. Kohlbacher; and A. Beyer. 2016. Testing and validation of computational methods for mass spectrometry. *Journal of proteome research*, 15(3), 809-814.

Guo, D.; K. Bemis; C. Rawlins; J. Agar; and O. Vitek. 2019. Unsupervised segmentation of mass spectrometric ion images characterizes morphology of tissues. *Bioinformatics*, 35(14), i208-i217.

Lieb, F.; T. Boskamp; and H.G. Stark. 2020. Peak detection for MALDI mass spectrometry imaging data using sparse frame multipliers. *Journal of Proteomics*, 225, 103852.

Makarov, A.; E. Denisov; A. Kholomeev; W. Balschun; O. Lange; K. Strupat; and S. Horning. 2006. Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Analytical chemistry*, 78(7), 2113-2120.

Morris, J.S.; K.R. Coombes; J. Koomen; K.A. Baggerly; and R. Kobayashi. 2005. Feature extraction and quantification for mass spectrometry in biomedical applications using the mean spectrum. *Bioinformatics*, 21(9), 1764-1775.

Noyce, A.B.; R. Smith; J. Dalgleish; R.M. Taylor; K.C. Erb; N. Okuda; and J.T. Prince. 2013. Mspire-Simulator: LC-MS shotgun proteomic simulator for creating realistic gold standard data. *Journal of proteome research*, 12(12), 5742-5749.

Palmer, A.D. 2014. Information processing for mass spectrometry imaging (Doctoral dissertation, University of Birmingham).

Palmer, A.; P. Phapale; I. Chernyavsky; R. Lavigne; D. Fay; A. Tarasov; V. Kovalev; J. Fuchser; S. Nikolenko; C. Pineau; and M. Becker. 2017. FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nature methods*, 14(1), 57-60.

Römpp, A. and B. Spengler. 2013. Mass spectrometry imaging with high resolution in mass and space. *Histochemistry and cell biology*, 139(6), 759-783.

Sarycheva, A., Grigoryev, A., Sidorchuk, D., Vladimirov, G., Khaitovich, P., Efimova, O., ... and Kostyukevich, Y. 2020. Structure-Preserving and Perceptually Consistent Approach for Visualization of Mass Spectrometry Imaging Datasets. Analytical Chemistry.

Schramm, T., Z. Hester; I. Klinkert; J.P. Both; R.M. Heeren; A. Brunelle; O. Laprévote; N. Desbenoit; M.F. Robbe; M. Stoeckli; and B. Spengler. 2012. imzML—a common data format for the flexible exchange and processing of mass spectrometry imaging data. *Journal of proteomics*, 75(16), 5106-5110.

Schulz-Trieglaff, O.; N. Pfeifer; C. Grpl; O. Kohlbacher; K. Reinert. 2008. LC-MSsim – a simulation software for liquid chromatography mass spectrometry data. *BMC Bioinformatics*, 9, 423

Verbeeck, N. (2014). Datamining of imaging mass spectrometry data for biomedical tissue exploration. (Doctoral dissertation, KU Leuven).

Verbeeck, N.; R.M. Caprioli; and R. Van de Plas. 2020. Unsupervised machine learning for exploratory data analysis in imaging mass spectrometry. *Mass spectrometry reviews*, 39(3), 245-291.

Wijetunge, C.D.; I. Saeed; B.A. Boughton; U. Roessner; and S.K. Halgamuge. 2015. A new peak detection algorithm for MALDI mass spectrometry data based on a modified Asymmetric Pseudo-Voigt model. *BMC genomics*, 16(12), 1-12.

MAKE-TO-ORDER PRODUCTION PLANNING WITH SEASONAL SUPPLY IN CANNED PINEAPPLE INDUSTRY

Kanapath Plangsrisakul
Tuanjai Somboonwiwat
Chareonchai Khompatraporn
Department of Production Engineering,
King Mongkut's University of Technology Thonburi (KMUTT), Bangkok 10140 Thailand
E-mails: kanapath.002@mail.kmutt.ac.th, tuanjai.som@kmutt.ac.th, charoenchai.kho@kmutt.ac.th

KEYWORDS
Canned Pineapple Industry, Make-to-order Inventory, Multi-products Multi-periods Production Planning, Pineapple Color Ratios, Seasonal Raw Materials.

ABSTRACT

This research studies a make-to-order production planning problem in a canned pineapple industry. Pineapple is a seasonal perishable fruit. Thus, the cost of fresh pineapple which is the main raw material in canned pineapple products is inexpensive during its season because of its abandance. The color of the pineapple also determines the price of the canned pineapple. However, the availability of different colors (referred as "choice" and "standard") is dependent. Specifically, if for a given month the ratio of the choice-color pineapple increases, the ratio of the standard-color pineapple decreases. There are several costs involve such as fresh pineapple cost, can cost, sugar cost, water cost, labor cost, energy cost, and inventory holding cost. This problem is formulated as a mathematical model to maximize the total profit over four-months planning horizon. Two supply uncertainty cases are tested which are low and high ratios of the choice color. The results show that the profit depends on available color ratios of the pineapple. The production planning is best if it matches with the availability of the color ratios. In certain months, some fresh pineapple purchased exceed the need of the production because of the dependency of the two colors. The inventory holding cost also influences the production decision—whether to produce the canned pineapple in earlier months or it is better to produce only the canned pineapple when it is needed to serve the customer orders.

INTRODUCTION

Thailand is the global exporter of canned pineapple with the market share of 37.2% worldwide or USD 338.09 million in value, followed by the Philippines and Indonesia (based on the 2019 statistics) (TRIDGE, 2019). The three countries together cover about 70% of the world's market share (Wattanakul et al., 2020). Pineapple is a seasonal fruit but the demands for canned pineapple exist throughout the year. Therefore, canned pineapple manufactures must produce canned pineapple when the fresh pineapple fruits are abundant to secure a low raw material cost and top fruit quality.

There are variety of canned pineapple products depending upon the fruit colors ("choice" and "standard"), fruit cut (slice, chunk, and tidbit), can sizes, syrup sweetness level, and so forth. The choice color of pineapple refers to a deep dark yellow color of the pineapple meat. The color is preferred by most customers. Canned products made with the choice color pineapple are generally sold at a higher price than the same products made with the standard color fruit. However, the color of the pineapple cannot be identified until the fruit is peeled, but fresh pineapple is sold to the canned manufacturers in bulk and unpeeled. Only monthly ratios of pineapple with choice and standard colors can be estimated. Canned manufacturers must sometimes buy additional fresh pineapple to ensure that there are enough choice color fruits to serve the pre-ordered and future demands. Any leftover fruits after all demands are fulfilled must be processed right away as fresh pineapple is perishable by being canned and stored in the warehouse for future orders. Some leftover is discarded as waste because there is no room available in the warehouse or it is too costly to keep it as a safety stock. Under all these conditions, the objective of a canned pineapple manufacturer is to determine a production plan that maximizes the total profit.

Canned pineapple manufacturers are generally facing a production planning problem under seasonal supply of fresh pineapple. A number of decisions needs to be addressed in the planning, specifically multiple products (cuts and can sizes based on the available colors) to be manufactured over multiple planning periods and under a warehouse capacity constraint. There are also several production related costs involved, adding additional complexity to the problem.

Certain aspects of this production planning problem were studied by Kogan et al. (1996). Their planning was to be responsive to customer demands as much as possible with make-to-order production, while considering make-to-stock products and minimizing inventory and purchasing costs. Soman et al. (2006) tested a conceptual framework for production planning and inventory management in a food industry with a

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

combined make-to-stock and make-to-order production. Chen et al. (2014) examined pricing and production of a combined make-to-stock and make-to-order system. Any demands that could not be immediately satisfied were backlogged or lost. They focused on monotonicity of the optimal control policy and the optimal price. Grillo et al. (2017) formulated a model that aimed to maximize two conflicting objectives, total profit and mean product freshness, of a fruit supply chain.

This paper focus on multi-products multi-periods production planning for make-to-order demands in the canned pineapple industry in which the raw material—fresh pineapple—is perishable and available seasonally. A challenge in this research is to determine the amount of fresh pineapple to purchase while the colors of the pineapple vary each month.

The organization of this paper is as follows. The next section describes the problem in more details, and the mathematical model is formulated. Then a numerical example in which certain data are based on a canned pineapple manufacturer in Thailand is presented, and its results are discussed. Finally, the last section concludes the paper.

PROBLEM FORMULATION

Problem Description

Since pineapple is a seasonal fruit, its acquiring price is cheapest when it is in season. Like many other fresh fruits, pineapple is perishable and must be processed as soon as it is harvested, often by canning or drying. Canned pineapple has a larger market than the dried one because of its longer shelf life and industrial standards are more acceptable worldwide**.**

Canned pineapple products are influenced by at least three factors: the color of the fruit, the cut, and the size of the can. In addition to fresh pineapple cost, several other costs are involved in the production such as can cost, sugar (to make the syrup) cost, water cost, labor cost, energy cost, and inventory holding cost. Warehouse storage availability during multiple planning periods is also a common issue for any canned pineapple manufacturer. Certain demands of canned pineapple are pre-ordered several months ahead of the delivery date to secure the goods at reasonable prices. The production planner of the canned pineapple manufacturer must determine the quantity of the fresh pineapple to purchase as well as the types of canned pineapple products and their quantities to manufacture in each time period in order to maximize the total profits. It is possible that some products are not sold right away but are stored in the warehouse to serve future demands. The products are usually palletized when kept in the warehouse. Each pallet contains a different number of cans depending on the can size.

Mathematical Model

The following mathematical model is a system of equations established to reflect the multi-products multi-periods production planning problem described above. Its objective is to maximize the total profit.

Indices

i	Pineapple color	i =1,2,3,…,I
j	Pineapple cut	j =1,2,3,…,J
k	Can size	k =1,2,3,…,K
t	Month	t =1,2,3,…,T

Parameters

$Profit$	Total profit (baht)
$Revenue$	Total revenue (baht)
$Cost$	Total costs (baht)
$D^t{}_{ijk}$	Demand of canned pineapple with color i cut j can size k in month t (cans)
$PO^t{}_{ijk}$	Price of canned pineapple with color i cut j can size k in month t (baht)
$InvO^t{}_{ijk}$	On-hand inventory of canned pineapple with color i cut j can size k in month t (cans)
CC_k	Cost per can of can size k (baht)
CE_k	Energy cost to manufacture a can of pineapple in can size i (baht/can)
CL_{jk}	Labor cost to manufacture a can of pineapple with cut j in can size k (baht/can)
CP^t	Average cost of fresh pineapple per kilogram in month t (baht/kilogram)
CS_k	Sugar cost to manufacture a can of pineapple in can size k (baht/can)
CW_k	Water cost to manufacture a can of pineapple in can size k (baht/can)
H	Inventory holding cost per pallet per month (baht/pallet/month)
QC_k	Quantity of can size k per pallet (cans/pallet)
WP_k	Weight of fresh pineapple needed to fill a can of size k (kilogram)
δ	Proportion of fresh pineapple by weight that can be canned (percentage)
$GP^t{}_i$	Proportion of pineapple color i available in month t (percentage)
Cap^t	Maximum product quantity that can be manufactured in month t (cans)
QCO^t_{max}	Maximum inventory that can be kept in month t (pallets)

Decision Variables

$X^t{}_{ijk}$	Quantity of canned pineapple product with color i cut j can size k to be produced in month t (cans)
QP^t	Quantity of pineapple that is bought in month t (kilograms)

Objective Function

The objective function is to maximize the total profits. Some costs are included for the completion of the model and may not affect the decision. They are also used as a means to communicate within the case study manufacturer.

Maximize *Profit*

$$Profit = Revenue - Cost \quad (1)$$

where *revenue* and *cost* are described by Equations (2) and (3)

$$Revenue = \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(PO^{t}{}_{ijk} \cdot D^{t}{}_{ijk}) \quad (2)$$

$$Cost = \sum_{t=1}^{T}(CP^{t} \cdot QP^{t}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CC_{k} \cdot X^{t}{}_{ijk}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CE_{k} \cdot X^{t}{}_{ijk}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CL_{jk} \cdot X^{t}{}_{ijk}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CS_{k} \cdot X^{t}{}_{ijk}) + \sum_{t=1}^{T}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(CW_{k} \cdot X^{t}{}_{ijk}) + \sum_{t=1}^{T}\sum_{k=1}^{K}\left(H \cdot \left\lceil \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} InvO^{t}{}_{ijk}}{QC_{k}} \right\rceil\right) \quad (3)$$

where $InvO^{t}{}_{ijk}$ is described by Equations (4).

$$InvO^{t}{}_{ijk} = \left(InvO^{t-1}{}_{ijk} + X^{t}{}_{ijk}\right) - D^{t}{}_{ijk} \quad \forall\, i,j,k,t \quad (4)$$

and let $InvO^{0}{}_{ijk} = 0$ for all *i, j,* and *k.*

Constraints

1. For each color, the raw material usage cannot exceed the available raw material in each month:

$$\sum_{j=1}^{J}\sum_{k=1}^{K}(WP_{k} \cdot X^{t}{}_{ijk}) \le \delta \cdot GP^{t}{}_{i} \cdot QP^{t} \quad \forall\, i,t \quad (5)$$

2. The production cannot exceed the capacity in any month:

$$Cap^{t} \ge \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} X^{t}{}_{ijk} \qquad \forall_{t} \quad (6)$$

3. The stock kept must not exceed the warehouse space availability in each month:

$$\sum_{k=1}^{K}\left\lceil \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} InvO^{t}{}_{ijk}}{QC_{k}} \right\rceil \le QCO^{t}_{max} \quad \forall_{t} \quad (7)$$

4. The monthly pre-ordered demands must be completely fullfilled:

$$InvO^{t-1}{}_{ijk} + X^{t}{}_{ijk} \ge D^{t}{}_{ijk} \quad \forall\, i,j,k,t \quad (8)$$

5. The decision variables and parameters must satisfy the following conditions.

$$InvO^{t-1}{}_{ijk} \ge 0 \quad \forall\, i,j,k,t \quad (9)$$

$$QP^{t} \ge 0 \quad \forall\, t \quad (10)$$

$$X^{t}{}_{ijk}\, \epsilon\, integer \quad \forall\, i,j,k,t \quad (11)$$

$$X^{t}{}_{ijk} \ge 0 \quad \forall\, i,j,k,t \quad (12)$$

NUMERICAL EXAMPLE

In this paper, only eight canned pineapple products are considered with the following variety: color i = 1 (choice), i = 2 (standard); cut j = 1 (slice), j = 2 (tidbit) and can size k = 1 (0.6 kilograms per can), k = 2 (3 kilograms per can). The planning covers four consecutive months (t = 1,2,3,4). The average prices of fresh pineapple in month 1 to 4 are 6.4, 6.5, 6.7, 7.0 baht per kilogram, respectively.

The can cost (CC), energy cost (CE), labor cost (CL), sugar cost (CS), and water cost (CW) of each product are shown in Table 1. A pallet can accommodate up to 20 smaller size cans (k = 1), or 10 of the larger ones (k = 2). Each pallet incurs about 40 baht per month as its inventory holding cost.

Table 1: Production Costs by Product

Product			Production Cost (Bath/Can)				
i	*j*	*k*	CC	CE	CL	CS	CW
1	1	1	1	0.3	0.4	0.2	0.25
1	1	2	6	1.5	1.5	0.6	0.75
1	2	1	1	0.3	0.45	0.2	0.25
1	2	2	6	1.5	2	0.6	0.75
2	1	1	1	0.3	0.4	0.2	0.25
2	1	2	6	1.5	1.5	0.6	0.75
2	2	1	1	0.3	0.45	0.2	0.25
2	2	2	6	1.5	2	0.6	0.75

The price of different canned pineapple products (PO) often varies on a month basis and can be summarized in Table 2. The table also shows the monthly pre-ordered demands for all the canned products, and the ratio of choice-color pineapple estimated based on a monthly basis. This ratio is uncertain and may change year by year. Two scenarios of this ratio are explored. The first one is when the ratio varies in a larger range (LR) than from 0.3-0.7; and the other one is when it varies in a smaller range (SR) from 0.4-0.6.

Once a fresh pineapple is peeled and cored, only about 80% of the original weight is left to be processed and canned. Due to the warehouse space availability, the manufacturer may hold up to 4,500,000 cans in the warehouse in any time period, or an equivalence of 100,00 pallets.

Table 2: Monthly Price per Can and Pre-Ordered Demand by Product

Month	Product			PO (Baht/ Can)	Demand (Can)	Ratio of Choice-color Pineapple	
	i	*j*	*k*			LR	SR
1	1	1	1	16	600,000	0.7	0.6
	1	1	2	25	-		
	1	2	1	16	640,000	0.7	0.6
	1	2	2	59	160,000		
	2	1	1	15	-	0.3	0.4
	2	1	2	24	200,000		
	2	2	1	16	320,000	0.3	0.4
	2	2	2	56	60,000		
2	1	1	1	20	860,000	0.6	0.5
	1	1	2	25	-		
	1	2	1	17	880,000	0.6	0.5
	1	2	2	59	240,000		
	2	1	1	15	-	0.4	0.5
	2	1	2	24	280,000		
	2	2	1	16	440,000	0.4	0.5
	2	2	2	51	80,000		
3	1	1	1	16	248,000	0.4	0.5
	1	1	2	25	328,000		
	1	2	1	17	244,000	0.4	0.5
	1	2	2	59	272,000		
	2	1	1	15	-	0.6	0.5
	2	1	2	24	320,000		
	2	2	1	17	480,000	0.6	0.5
	2	2	2	62	276,000		
4	1	1	1	15	1,000,000	0.3	0.4
	1	1	2	25	-		
	1	2	1	17	960,000	0.3	0.4
	1	2	2	59	180,000		
	2	1	1	15	-	0.7	0.6
	2	1	2	22	640,000		
	2	2	1	14	1,000,000	0.7	0.6
	2	2	2	62	-		

Results

The production planning problem above was solved using Excel Solver. The results are shown Table 3.

From the table, the results show that in both scenarios the production plan tends to take advantage of low average fresh pineapple costs in earlier months by over-manufacturing certain products in the months prior to the delivery date even though inventory holding costs are incurred. Let a triplet (*i*,*j*,*k*) represents a product with color *i*, cut *j* and can size *k*. In the SR scenario for example, the production of product (1,1,1) and product (2,1,2) in month 3 exceed their monthly demand. The excess quantities of these products together with additional production in month 4 are to serve their demands in month 4. Similarly, in the LR scenario the production of product (1,1,1) in month 3 can accommodate both demands for months 3 and 4. Another example is product (1,2,2) in the LR scenario. Its production is accumulated over months 1, 2, and 3 to serve the demand in months 3 and 4. The results of some other canned pineapple products in Table 3 also exhibit similar early production.

Table 3: Production Plan for Large and Small Range of Choice-Color Pineapplr Ratios Compared to Demands

Month	Product			Demand (Can)	Production Plan	
	i	*j*	*k*		LR	SR
1	1	1	1	600,000	600,000	965,599
	1	1	2	-	-	3,534
	1	2	1	640,000	640,000	640,002
	1	2	2	160,000	508,000	161,332
	2	1	1	-	-	-
	2	1	2	200,000	200,000	200,000
	2	2	1	320,000	320,000	320,000
	2	2	2	60,000	60,000	60,000
2	1	1	1	860,000	860,000	494,401
	1	1	2	-	-	7
	1	2	1	880,000	880,000	879,998
	1	2	2	240,000	324,000	238,668
	2	1	1	-	-	-
	2	1	2	280,000	280,000	280,000
	2	2	1	440,000	440,000	440,000
	2	2	2	80,000	80,000	145,554
3	1	1	1	248,000	448,125	1,248,000
	1	1	2	328,000	328,000	324,459
	1	2	1	244,000	244,000	244,000
	1	2	2	272,000	20,000	270,000
	2	1	1	-	-	-
	2	1	2	320,000	338,729	588,413
	2	2	1	480,000	480,000	480,000
	2	2	2	276,000	276,000	210,446
4	1	1	1	1,000,000	799,875	-
	1	1	2	-	-	-
	1	2	1	960,000	960,000	960,000
	1	2	2	180,000	-	180,000
	2	1	1	-	-	-
	2	1	2	640,000	621,271	371,587
	2	2	1	1,000,000	1,000,000	1,000,000
	2	2	2	-	-	-

From production planning results of both scenarios, the quantities of fresh pineapple to purchase each month must be determined to meet production plan as shown in Figure 1. From the figure, the quantities of fresh pineapple to purchase in both scenarios are the highest in month 3. This is because the price of the fresh pineapple increase in month 4, so it is worthwhile to manufacture some excess canned pineapple in month 3 and keep it in stock for a month before selling it in month 4. In general, the quantities of fresh pineapple to purchase in the SR scenario is lower or close to that in the LR one, except for month 3.

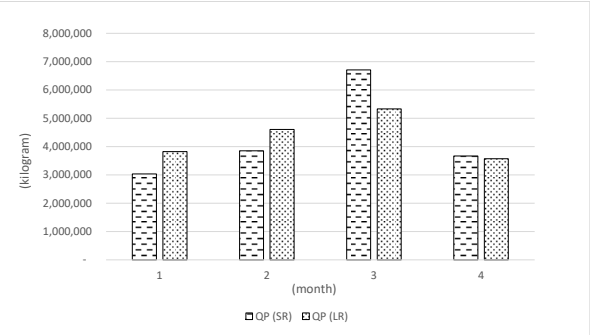


Figure 1: Quantity of fresh pineapple to be purchased each month in both scenarios (in kilogram)

Based on the pre-ordered demands and pre-determined revenue per can of all the canned pineapple products. The total revenue is 242,056,000 million baht. However, the production plans based on the two scenarios offer different total profit because they incur different costs, particular the pineapple cost and the inventory holding cost as shown in Table 4.

Table 4: Pineapple Cost, Inventory Holding Cost, and Profit in The Two Scenarios (In Baht)

Financial Item	Scenario	
	LR	SR
Pineapple Cost	114,571,382	114,449,393
Inventory Holding Cost	4,315,200	4,100,800
Other production Costs	48,799,600	48,799,600
Total Costs	167,686,183	167,349,793
Total Profit	74,369,817	74,706,207

Table 4 reveals that when the variation in the ratio of the choice-color pineapple fluctuates in a smaller range, the total profit is 336,390 baht higher than when this ratio varies in a larger range. This is because the fresh pineapple cost and the inventory holding cost in the SR scenario are lower than those in the LR scenario. A reason that the fresh pineapple (raw material) cost in the LR scenario is higher because in certain months the manufacturer needs to buy excess fresh pineapple so that there is enough choice-color pineapple to satisfy the product demands.

CONCLUSION

This paper studies a production planning problem that is normally experienced by a canned pineapple manufacturer. Pineapple is a seasonal fruit. Its price and availability of choice-color pineapple differ throughout the year. Several production and inventory holding costs are considered in this study. A mathematical model of this problem is formulated with eight different canned products under warehouse storage space limitation and pre-ordered demands over four consecutive months. The objective is to maximize the total profit. Two scenarios are experimented to examine how different ratios of choice-color pineapple availability affect the planning and total profit.

The numerical results show that the production plans in both scenarios tend to over-manufacture some products when the price of the fresh pineapple is low and allow to incur inventory holding cost to keep the excess products until they are later needed to serve the demands. The ratio of choice-color pineapple availability slightly affects the total profit. The scenario when the ratio of choice-color pineapple availability varies within a smaller range offers a lower fresh pineapple cost and inventory holding cost than those of the scenario when this ratio fluctuates within a larger range.

This study focuses on make-to-order production. Due to seasonal availability of the fresh pineapple, a common practice in the canned pineapple industry is to make-to-stock some of the products to serve future demands. This is because the cost of the fresh pineapple which is the main ingredient of the canned pineapple products are the cheapest and its availability is peaked during the pineapple season. The optimal production plans obtained consider various production and inventory holding costs, estimated ratios of choice-color pineapple availability, and warehouse storage limitation to serve pro-ordered demands.

Excess fresh pineapple acquired so that there would be sufficient choice-color products to serve the customers could be used to produce make-to-stock products before it is wested. Future research may suggest make-to-stock quantities of different canned pineapple products based on the forecasted demands.

The results of this research can be used as a guide to canned pineapple production planning when there are changes in fresh pineaaple prices and color ratios. With this guide, a raw material purchasing plan and selling promotions could be set ahead of time.

REFERENCES

Chen, X., Tai, A.H. and Y. Yang. 2014. "Optimal production and pricing policies in a combined make-to-order/make-to-stock system," *International Journal of Production Research*, 52, 7027-45.

Grillo, H., Alemany, M.M.E., Ortiz, A. and V. S. Fuertes-Miquel. 2017. "Mathematical modelling of the order-promising process for fruit supply chains considering the perishability and subtypes of products," *Applied Mathematical Modelling*, 49, 255-78.

Kogan, K., Khmelnitsky, E. and O. Maimon. 1998. "Balancing facilities in aggregate production planning: Make-to-order and make-to-stock environments," *International Journal of Production Research*, 36, 2585-96.

Soman, C. A., van Donk, D.P. and G.J.C. Gaalman. 2007. "Capacitated planning and scheduling for combined make-to-order and make-to-stock production in the food industry: An illustrative case study," *International Journal of Production Economics*, 108, 191-99.

TRIDGE 2019 Overview of global value added pineapple market.https://www.tridge.com/intelligences/canned-pineapple/export (On-line accessed on 1 February 2021).
Wattanakul, T., Nonthapot, S., and T. Watchalaanun. 2020. "Factors determine Thailand's processed pineapple export competitiveness," *International Journal of Managerial Studies and Research*, 8(1), 36-41.

AUTHOR BIOGRAPHIES

KANAPATH PLANGSRISAKUL is a graduate student in Industrial and Manufacturing Systems Engineering program at the Department of Production Engineering, King Mongkut's University of Technology Thonburi, Thailand. His research interests are in production planning and optimization, and data management. His e-mail address is: kanapath.002@mail.kmutt.ac.th.

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 Oregon State University, Corvallis, USA. Her research interests include green supply chain and logistics, business process and applications of operations research. Her e-mail address is: tuanjai.som@kmutt.ac.th.

CHAROENCHAI KHOMPATRAPORN holds a Ph.D. from University of Washington, USA. He is an Associate Professor in the Department of Production Engineering at King Mongkut's University of Technology Thonburi, Thailand. His research interests include supply chain and logistics management, applied operations research, optimization algorithms, and industrial sustainable operations management. He has published in several peer-reviewed journals, and worked closely with both public and private sectors such as hard disk drive manufacturer, automotive industry, and banking. His e-mail address is: charoenchai.kho@kmutt.ac.th

Modelling Player Combat Behaviour for NPC Imitation and Combat Awareness Analysis

Paul Williamson
School of Computing and Mathematics
University of South Wales
CF37 1DL, Pontypridd
15082938@students.southwales.ac.uk

Dr Christopher Tubb
School of Computing and Mathematics
University of South Wales
CF37 1DL, Pontypridd
christopher.tubb@southwales.ac.uk

***Abstract*—NPC [non-player characters] have progressed over the past two decades, they fulfil a number of different roles, each with different problems and development techniques. When fulfilling the role typically reserved for human-players, a problem occurs because they can be identified as NPC by observing their gameplay behaviours. This has negative consequences when deployed in a team-based game where eliminations impact game objectives. This research investigates the key combat characteristics exhibited by players during certain scenarios, analysing the data acquired through experiments to determine where generalised patterns emerge. It also explores the combat awareness of players when NPCs have overly tuned combat skill, and determine how effective standard game industry techniques are for creating believable NPCs.**

Keywords—NPC; Player Modelling; Gaming; Behaviour; Gameplay

I. Introduction

This paper explores the combat behaviour of human players in a FPS [first person shooter] game, and to determine what elements of their combat can be modelled for NPC [non-player character] imitation. It also examines how much combat awareness players have when NPCs are not modelled to imitate human behaviours.

This paper details two experiments, Firstly, an experiment was undertaken to capture how subjects reacted to common FPS combat situations, these range from suddenly appearing targets to varying target size. The second experiment focuses on the combat perception of NPCs when modelled using current standard techniques; those commonly used in the games industry, with emphasis on fast reaction speed and accuracy.

These experiments help show that players exhibit generalised patterns of behaviour that can be modelled in NPCs to imitate human players, this would present a novel approach to NPC development. It also highlights the potential impact of poorly modelled NPCs and when NPCs fulfil a player role, when they do not exhibit the same behaviours as players, it has negative effect on immersion and entertainment.

II. Background

A. Non-Player Character

Non-player characters have been an important aspect of gaming since the birth of the video game industry; they fulfil an array of different roles and are an important entity for most video games. In the FPS genre, enemy NPCs often fulfil one of the following roles;

- Boss: These NPCs have a variety of mechanics and tactics; they should present a challenging experience and have a prominent presence in the title.
- Elite Mobs: These NPCs are strong enemies that require more attention than trash mobs and often have more health and/or strong weapons.
- Trash Mobs: These NPCs are the most frequent type of enemy; they are individually the weakest category of enemy in the title.

The purpose of an NPC in a title can vary, but generally they can only fulfil one of the following stances toward the player at a given time;

- Friendly: When an NPC is friendly, this often means they are an ally and a non-threat.
- Neutral: An NPC with a neutral stance will not attack the player unless attacked or some in-game situation causes a change.
- Enemy: An enemy is an NPC that will attack the player and are seen as a threat.

Finally, there are NPCs that fulfil the role normally reserved for a human-player; these NPCs are expected to provide the same amount of challenge that is experienced when playing against human opponents. These NPCs can be employed in team-based games should one side have more players than the other; thus, evening the teams. Other application for these NPCs are for when the player wants to experience the multiplayer game, but does not necessarily want to play against actual human-players, these type of NPCs are often referred as bots.

B. NPC Modelling Techniques

There are a number of common techniques used in the gaming industry for developing the mechanisms necessary for NPCs. Depending on the action of the NPC will in part determine what type of solution will be required;

- A* Algorithm: Traversing the environment is an important function for an NPC, A* algorithm provides heuristic approach for the path finding.
- Finite-State Machine: Ensuring NPCs are actively pursuing an objective, it is important that NPCs can switch between tasks depending on the current situation. FSM [finite-state machines], provide a good solution as the NPC can only occupy one state at a time, and can transition to a new state when the condition requirement is met.
- Scripting: Scripting is a powerful tool for NPCs because scripts can be used to run exclusively for the NPC, they can be used to obtain external data, such as scanning for

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

enemies. It can also monitor internal data, for example, keeping track of health and ammunition.

- Behaviour Tree: BT [Behaviour Tree] can be used to model NPC behaviour; they are particularly powerful when creating complex behaviour for decision making. BTs can also be influenced through event-driven mechanisms, this makes BTs a very powerful tool which can be visually illustrated and quickly implemented.

It should be noted that there is a number of techniques for NPC development; with the role they will fulfil influencing which techniques will be most suitable.

C. Modelling Player Gameplay

The patterns and combat characteristics of players can be modelled by observing how they react and respond to commonly occurring situations and recording the combat efficiency as the fight unfolds. The core attributes for combat are;

- Reaction Time: This is how long it takes for the player to respond to a situation as it occurs. For example, when an enemy suddenly appears, this can be modelled by timing how long it takes for NPCs to react.
- Accuracy: Combat accuracy is a crucial aspect of combat efficiency because higher the accuracy, the faster they can eliminate the target and it helps determines skill level. Modelling accuracy can be achieved by adjusting the probability to hit the target based on skill and situational data.
- Combat Awareness: While combat awareness is an abstract notion and can be quite vague, it can be modelled by comparing the data variance between difference scenarios which are likely to occur during play. The patterns that emerge could then be used in the model to display situational human-like combat gameplay.

While these combat attributes are important, it is also vital to determine what mechanisms influence these attributes and what affect they have on generalised patterns.

D. What is Gameplay?

The term gameplay can be quite ambiguous and so when trying to model gameplay it presents a problem as to what needs to be modelled. Fabricatore [1] describes gameplay as the actions completed by a player as the game unfolds; this is supported by Sedig et al [2] that suggests gameplay is the emergence of experience from the actions of a given player. For the purposes of this paper, gameplay should be extended to include that of NPCs and therefore define gameplay as;

> 'Actions or decisions taken by an entity, within the parameters of the specific game mechanisms and rules'

This statement suggests gameplay is therefore derived from the mechanisms of a game, and the gameplay emerges when actions are performed with the constraints of the rules.

III. MOTIVATION

While NPCs fulfil a vast variety of roles, when tasked with standing in for a human player it is vital they are capable of imitating the general characteristics of a player. When NPCs are easily identified by an opposing human player, they can become the primary focus for the opposition, especially if the combat behaviour of the NPC performance is low or predictable. When there are consequences attached to dying, such as in a death match where the objective is to get more eliminations than your opposition, having a poor performing NPC can impact the overall enjoyment of the game. This presents a unique problem, having a poorly modelled NPC can impact enjoyment, but also having uneven teams can be a negative and unfair experience.

We believe to have an engaging multiplayer experience where NPCs have the capability of fulfilling player roles; they should ideally be indistinguishable from human participants. This type of NPC needs to be flexible in combat behaviour; similarly to how randomly selected players will have random skill levels. They could also have a positive impact on single player experiences as well, because NPCs modelled on the performance of an individual and scaled to match the ability of the player, would provide a more challenging experience.

IV. RELATED RESEARCH

A. Non-Player Character AI Techniques

While traditional techniques are often used in the gaming industry, research into applying complex AI [artificial intelligence] techniques have been undertaken. With regards to believability and fulfilling a player role, Pfau et al [3] have explored using deep learning to simulate player behaviour by using DPBM [deep player behaviour modelling]. In this research they applied DPBM by generating an action based off its current state description using a neural network where the weighted choice is based of real-time feedback and previously captured human player gameplay. Their results yielded promising progress with a significant selection of subjects finding the NPCs to be undistinguishable from human players.

Research undertaken by Galvin and Madden [4] focused on using RL [reinforcement learning] to gradually train the NPC to improve its skill, cataloguing the skill in intervals as it is progressing, then enabling the NPC to dynamically select a desired skill level in real-time, which they called the skill experience catalogue. When experimenting with combat behaviour, they successfully created a skill timeline with five milestones; each milestone matched the proficiency of a fixed-strategy opponent with varying skill.

B. Believability Identification

Research by Warpefelt [5] stated that NPCs have significantly improved in believability over the past generation; however, further research is required. They developed a matrix called GAM [Game Agent Model]; this builds upon the Carley & Newell fractionation matrix, by categorising NPCs in five levels of social complexity. Their findings show that the higher the complexity the more likely the illusion of believability fails and when an NPC fails to be believable, it can cascade across the GAM.

Hinkkenen et al. [6] created a framework to evaluate the believability of NPCs in a game, using player based perception. They propose that NPCs believability can be reduced to three key aspects, movement, behaviour and animation and by improving these will improve believability. Similar research conducted by Togelius et al. [7] suggest that to evaluate believability in NPCs, it is more accurate to use

an external observer, who are impartial to the game, which form the basis for the Turin test for bots.

Kersjes and Spronck [8] argue that using only techniques such as decision trees and finite-state machines is detrimental to believability and that when adding emotion to NPCs it can create more individualistic and diverse NPCs. This is somewhat supported by Hamdy and King [9] who argues that to imitate human players, NPCs must exhibit traits intrinsic to humans, such as, emotion, interpretation and memory, and to use a MARPO-type architecture for modelling NPCs.

C. NPC Behaviour

Bakkes et al. research [10] into rapid adaptation AI show that NPCs were able to adapt to current situations by gathering information of its domain and exploiting it accordingly in a case-based system. This enables NPCs to adapt its behaviour based on its opponent by accessing previously stored gameplay samples and then producing a suitable behaviour. They show that in strategy situations, their approach proved effective for enabling NPCs to adapt to its current condition and when combined with scaling difficulty able to define the effectiveness of NPCs.

This is supported by Delatorre et al. [11] who highlight the negative impact of overly difficult NPCs and that their study shows players are more immersed when the rules are unknown. This can be interpreted for NPCs as a whole, and so when NPCs are predictable or overly difficult, it significantly affects the player perception and enjoyment.

V. Experiments

In order to accurately model the combat behaviour of human players, a number of experiments were performed to understand some of the parameters of this behaviour. The following section details the combat experiments by having subjects complete a series of scenarios and by playing against NPCs with no dedicated combat modelling. The purpose of the first experiment is to determine if subjects exhibit generalised patterns in their gameplay, and to evaluate the main mechanisms that influence combat behaviour. The second experiment analyses the combat awareness subjects possessed when facing NPCs that were over tuned to have unrealistically fast reactions and accuracy, this will help identify just how observant players are of their opposition and the effects of poorly modelled NPCS.

A. Experimantal Environment

All experiments were developed using Unity3D and all subjects were selected anomalously. Subjects were required to download and run the experiment on their system and use their peripherals. After experiment completion, data was sent and stored on an online database.

B. Combat Experiment

This experiment consists of four scenarios, each with five stages, as subjects progress through the stages the difficulty increasing by tweaking the constraint.

a) Single Target

This scenario consists of targets spawning individually, the objective is for the subject to move their cursor and left click over the target to eliminate it before it self-destructs (figure 1).

Fig. 1. Single Target Example

The initial self-destruct time was set to 2.5 seconds, with a reduction of 0.5 seconds per stage; in the final stage self-destruct was at 0.5 seconds. This experiment was intended to find the base reaction time to respond to a target suddenly appearing at a random location. The distance between the cursor and target was recorded to account for any effect distance has on time to click.

b) Increasing Spawn Rate

This scenario was designed to simulate the effect of new targets suddenly appearing while one or more targets are still active. The initial spawn rate was set to 1.12 seconds, then as the subjects progress, the spawn timer reduced by 0.225 seconds, with the final stage spawn rate at 0.275 seconds. The targets have a flat 2 second self-destruct time, which does not change throughout the scenario (figure 2).

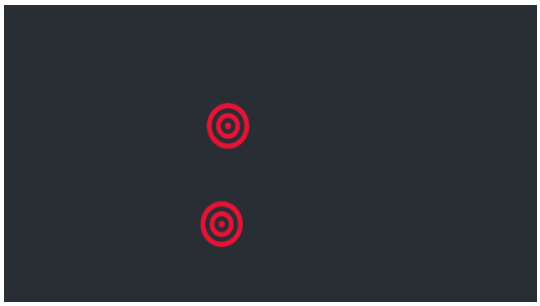


Fig. 2. Increased Spawn Rate Example

The experiment was intended to determine if subjects have targeting sequences and if patterns emerge in their targeting techniques as more targets are present on the screen. This is an important scenario because it is not unusual for new targets to appear during combat.

c) Grouped Targets

This scenario analyses the targeting patterns exhibited when a group of targets spawn at a given time. The initial stage consisted of three target, with an additional two targets added per stage, therefore, the final stage will have eleven targets (figure 3).

Fig. 3. Grouped Targets Example

Targets will self-destruct after 5 seconds, with the time and distance to targets recorded as in previous experiments. Targeting decisions are an essential aspect of combat behaviour, this scenario aims to identify if subjects employ personal techniques to increase efficiency.

d) Varying Size

Size is a crucial aspect of a 3D game, because size can denote distance to the target, this means targets of various sizes are likely to be presented to the player. The initial size is set to 0.6 inches, with the size decreasing by 0.1 inches per stage; the final stage has a size of 0.2 inches (figure 4).

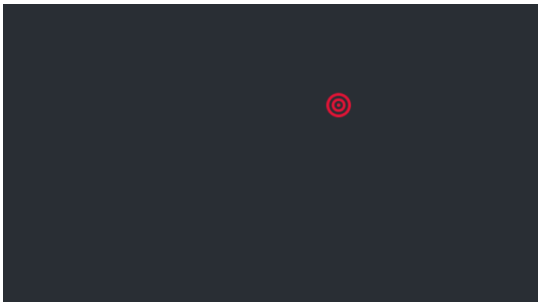


Fig. 4. Size Targets Example

Targets spawn every 2 seconds and have a self-destruct timer of 2 seconds. The scenario will determine if there is a performance drop as the target size decreases, and if size has a detrimental effect to combat efficiency.

C. *Combat Awareness*

This experiment is a general DM [death match] scenario, there are two NPCs and the winning criteria are to obtain five eliminations before the opponents. The map consists of a large reception area with eleven rooms; there is a slight variance in room sizes (figure 5).

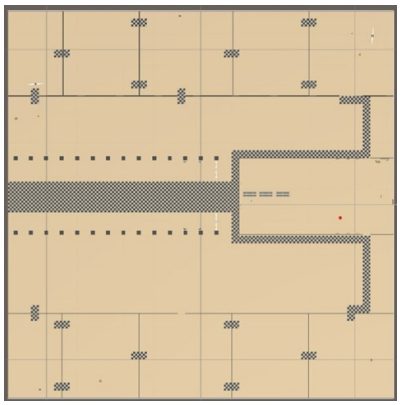


Fig. 5. Increased Spawn Rate Example

There is only one weapon which has one automatic fire mode, jumping has been disabled to simplify the experiment. There are two types of collectables, a medic-pack which increases the players health by 50% and an ammunition-pack which adds one clip worth of ammunition to the players overall stash.

The main purpose of this experiment is to gauge the feedback from subjects with regards to how human-like the NPC combat gameplay appears when using those basic techniques which are commonly employed. Another purpose is to determine to what extent subjects are aware of their opponents combat skill, and what impact this has on overall enjoyment of the experience.

a) NPC

The NPCs were designed to have unrealistic combat skill, they have instant reactions and perfect accuracy, this high degree of performance was chosen to fully evaluate player awareness. NPCs scan for opponents by constantly checking the viewport of the attached camera, it cycles through opponent world position to calculate if the opponent is in front of them, then using 'line casts' to determine if the target can be seen. When the NPC successfully identifies a target, the scanning is paused and the NPC engages in combat, when combat ends the scanning is resumed.

During combat the NPC will only attack when the target is in range and can be seen, they move towards the target when out of range and begin attacking when the range threshold has been crossed. If the target flees out of sight, the NPC will resume scanning and continue to roam the map. Reloading only occurs when the NPCs clip is out of ammunition during combat, or if reloading can be performed when out of combat.

For navigation, a standard slighted weighted A* algorithm was used, with increased cost for walkable nodes close to non-walkable nodes. Figure 6 shows the A* grid (left) and the map (right), the darker the shade of grey the more cost to use the node.

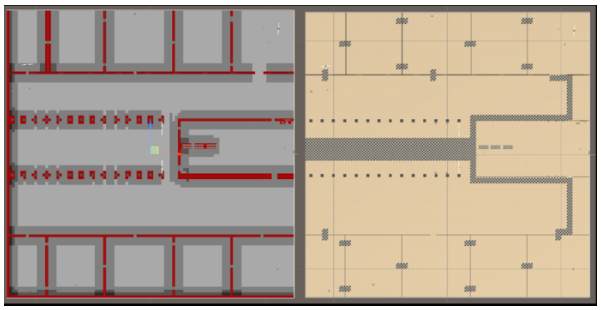


Fig. 6. A* Algorithm Grid and Map

When roaming NPCs create a list of locations they are going to visit, this was done by placing points of interest throughout the map, these points are used to create a list starting with the closest, then closest to the previous item in the list. When all points are in the list the NPCs creates a route to the first item, when a point has been reached, it is removed from the list and a new route is created to the next item. As roaming is the default navigation behaviour, all other activities have higher privilege, when roaming is stopped and then restarted, a new list is created.

Decision-making uses FSM [finite-state machines] with a BT [behaviour tree] philosophy, this was done by giving

activities a priority order and transitions to different states can only occur when the behaviour tree is satisfied. The FSM uses a static methodology, transition conditions to other states are always the same, for example, when the NPCs health is below 50%, they will stop roaming and go to a medic-pack location.

An NPC manager was created to enable combat, navigation and decision-making scripts communicate, and to store crucial data about the NPC, such as health.

The NPC also monitored itself and depending on its current situation can cause transitions to a different state, this method of control enabled the NPC to ensure it is always had a state occupied and to perform subtle tasks, for instance reloading.

b) Collectables

Collectables were achieved by attaching a collision box to a game object, the boundary of the collision box was increased so the player does not need to run through the object but pass by very close. When the players collision box intersects with the collectables collision box, a check is performed to determine if they player can use the item and if so the item is destroyed and correct stats added to the player. The collectable respawns 5 seconds after collection in the same place; all collectables have a set location which does not change.

Collectables will only be taken if they can be used and will not increase the maximum health or ammunition of a player. Therefore, if a player has 1 health point missing and they pick up a health-pack they will only receive 1 health point. If a player passing through a collectable that they cannot use, the collectable is not collected or destroyed.

c) Constraints

A series of constraints were added to simplify the process so more focus could be applied to the main objective of the experiment. Jumping can be a key identifier of human-controlled character but they sometimes jump unnecessarily, and as there are no pitfalls in the experiment, jumping is not required.

There is only one weapon, all players start with an assault rifle which has one automatic fire mode, this was decided as it removes the need to balance weapons and combat strength to ensure a fair fight. No scopes were added, but a crosshair was added to help assist subjects as they were effectively shooting off the hip.

d) Environment

The environment of the map is all static and non-interactive, there are no doors and lighting is ambient only. Bullet holes are shown on walls and ceiling, this provides feedback when shots missed their target.

As it was important a smooth experience would yield more accurate results, it was decided limiting the number of non-player to player potential actions would be beneficial.

VI. Results

The results show there is conclusive evidence to suggest that player combat behaviour can be modelled and when NPC combat is modelled poorly, it can have a negative effect for players.

a) Modelling Player Combat Behaviour

The results from the reaction experiments show that patterns do emerge with regards to player combat behaviour, these behaviours could be modelled to develop human-like NPCs. Figure 7 displays the result of the single target experiment, the average reaction time was 0.73 seconds with a standard deviation of 0.42.

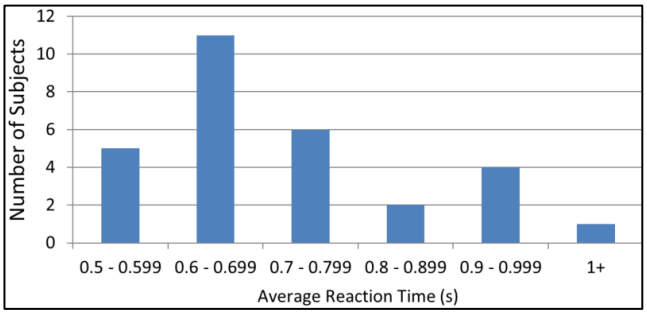


Fig. 7. Single Target Average Reaction Time

While there was some dispersion amongst the individual results, the majority of the successful hits were clustered between 0.6 – 0.8 seconds, this provides a good basis for the average responsiveness for NPC in a model.

When targets were spawning at an increased rate, it had a noticeable effect on reaction time, and suggests that when more targets are in view of the player, it can often have a negative effect on combat efficiency on that player. Figure 8 shows the reaction time for the increased spawn rate, the average reaction time was 0.81 seconds with a standard deviation of 2.83. This is an important behaviour and it could be argued when targets appear during combat, it has a slight effect on reactions because the player needs to factor the new target into their current strategy.

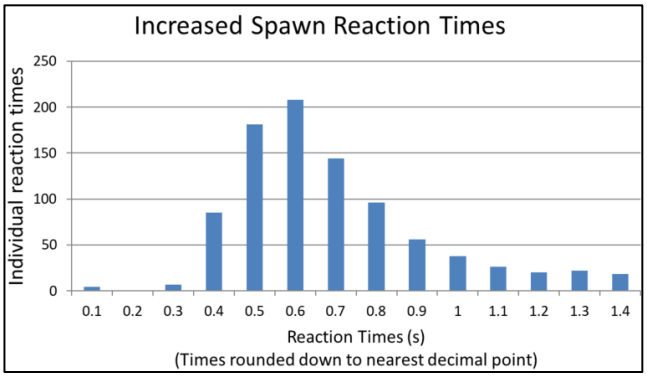


Fig. 8. Increased Spawn Rate Reaction Time

The analysis of the grouped experiment suggest subjects capable of formulating a strategy for attacking the targets, while tactics varied, the majority of the subjects first attacked the targets that spawned to form clusters, and then targeted the solo targets. When comparing the average reaction time between single target and grouped targets, there is a significant difference, there was better efficiency when multiple targets spawn than when they spawned one at a time (figure. 9).

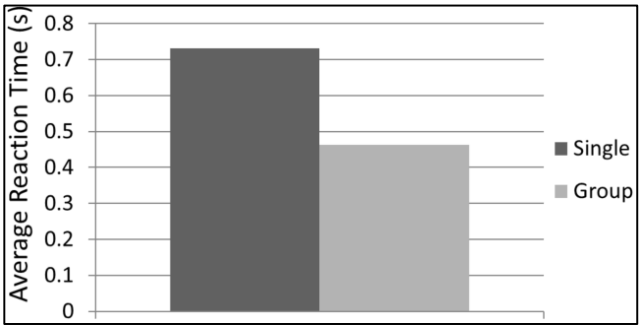


Fig. 9. Comparison Reaction Time

This is an important discovery because it suggests that there is a slight overhead when formulating a response, and that the initial number of targets does not have a noticeable time effect when formulating this response. It is also interesting when comparing with the increased spawning because in that scenario the strategy needed to be frequently updated, this proves that there is some overhead time needed to calculate a strategy.

When evaluating the effect size had on reaction time, there is a noticeable increase in average reaction time as the target got smaller. Figure 10 shows the reactions time of the large target (0.6 inch) and the smallest (0.2 inch), the cluster of results clearly show subjects were able to dispatch the large targets faster than the smaller targets.

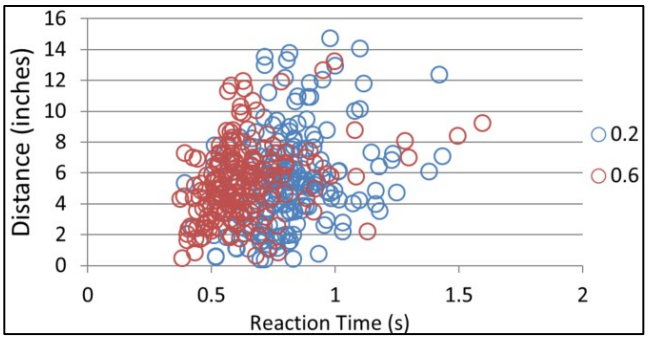


Fig. 10. Size Reaction Time Correlation

The importance of this finding is that it proves precision has an effect on reaction time, and with a difference of 0.19 seconds, which mean it can be modelled. It should also be noted that target distance from cursor did not have a noteworthy effect, and it was precision that the biggest effect. It could also be argued that precision will have an exponential negative effect on reaction time, the smaller the target becomes, the more precise and thus longer it will take to dispatch.

A pattern also emerged when analysing the targeting sequence when multiple targets appeared at separate times, figure 11 highlights the results when three or more targets were present on the screen.

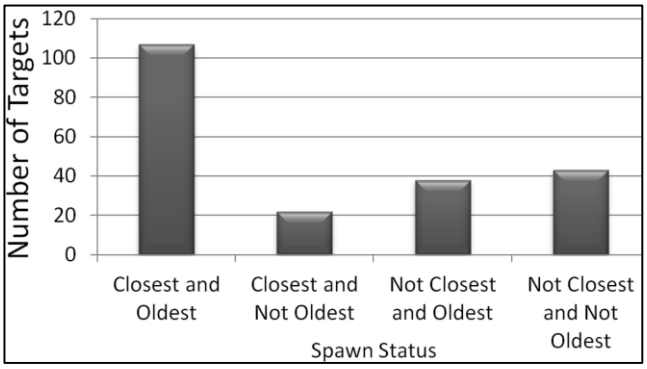


Fig. 11. Three or More Target Sequence

While the majority seem to stick to the strategy of targeting the oldest target, there were some subjects that were target switching as the number of targets grew. This implies flexibility in the combat gameplay for some subjects, and the ability to adapt to changing situations needs to be modelled for NPCs.

Lastly, when analysing the results for missed shots, there was evidence to conclude that when subjects missed a target they responded by rapidly taking more shots. This scatter shot behaviour did not seem to be present during the easier stages, but was more visible in the latter. This further show the adaptiveness of human-players and it is important NPCs are not static in their behaviour but able to employ different tactics when needed.

b) Player Combat Awareness

The results from the combat awareness experiment show a high degree of awareness from subjects, and when NPCs are overly tuned in combat efficiency, it can have a negative effect on enjoyment. Figure 12 shows the general perception of the NPCs in relation to how human-like they appeared.

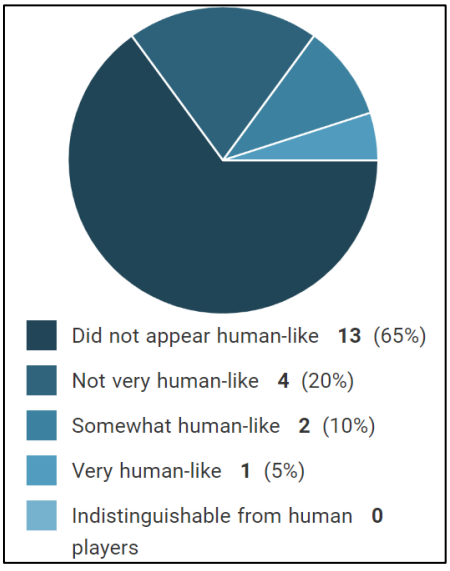


Fig. 12. How human-like did the NPCs appear

The result provides a good indication that when NPCs are not modelled to imitate human-players they can be clearly identified as non-human controlled, despite having the same combat actions available. When evaluating where the combat failed to appear human-like, figure 13 highlights

that both reaction speed and accuracy were selected by most subjects. This is interesting because it proves players do have a good understanding of the combat efficiency of their opposition and when this efficiency is higher than usual, it is detected.

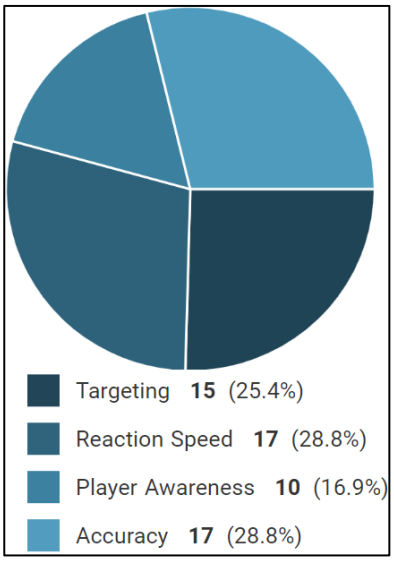


Fig. 13. Combat Awareness Identification

When analysing the feedback of the overly tuned combat efficiency, the general consensus is focused on the fact that the NPCs had perfect aim (figure 14). This supports the claim that players have good perception of combat awareness, and that poorly modelled combat can have a detrimental effect on overall enjoyment.

the aim
Nothing i can think of.
The fact that they never miss even when you strafe the shots
the snapping to me
Their movement system, there was no strategy to it, it seemed to be random
They had perfect aim and response times
Didn't see much movement but was killed as soon as I saw them
the movement around the map was believable
bot accuracy
Bot like accuracy of the NPCs when they are tiny specs on the screen.
Comabat
Movement
Shooting
fast reactions
Everything
fighting

Fig. 14. Combat Awareness Feedback

It should be noted that subjects were aware that they were playing against NPCs, judging by the tone of the feedback, if this combat model was used in a multiplayer title, there could be accusations that cheaters were prevalent in the game, which is expressed in some of the comments. This is a vital discovery because it suggests NPCs that are too skilled have the appearance of cheaters, and the perception of cheaters has a profound negative effect on the game reputation and player entertainment [12].

This survey was necessary because it shows there is a problem when using industry standard techniques for developing NPCs to imitate human players. It shows that players are acutely aware of the combat behaviour of their opponent and when displaying behaviours that are not commonly identified in human players, it influences their opinion of that opponent.

The data acquired from the survey could be used as a guide as to what areas of combat have the biggest impact on believability of the NPC. For example, when modelling the targeting and reaction speed of NPCs, it is imperative that they resemble human players as this perception was identified as a key area where believability failed.

This survey along with the data acquired from the combat experiments could be used to model the core combat attributes for NPCs. For instance, by implementing a skill based attribute, NPCs accuracy and reaction speed could be controlled individually to resemble that of a human player around the same combat skill level.

VII. Conclusion

This paper has shown that there is a quantitative element to generalised player combat efficiency and patterns emerge which can be modelled. Target distance did not have a noticeable effect on reaction time; however, there was an overhead cost for precision which did have a clear effect on reaction time. Subjects exhibited targeting strategies when numerous targets spawned together, and when new targets were spawning before old targets had been dispatched, there was evidence that it had a slight effect on combat efficiency.

Subjects had a clear awareness of the increased NPC combat efficiency; this had a profound negative effect on how human-like the NPC appeared, with the majority of the responses indicating accuracy and reaction time to be the cause. A number of responses also highlighted that the combat efficiency was reminiscent to that of cheating players; this paper believes the impact of overly potent NPC skill could have negative consequences towards overall enjoyment.

VIII. Future Work

While it was clear patterns in combat behaviour emerged, future work needs to be undertaken to model the patterns for NPC development, an analysis completed to determine the effectiveness of the model. Furthermore, combat is one part of gameplay, to provide a complete model; experimentation will be required in navigation and decision-making with the final objective to undertake a Turing-test for bots.

References

[1] Fabricatore, C. ”Video game development and its potential for education and creativity”. In: Creativity – Path of creation. 2007 (Unpublished)

[2] Sedig, K. & Parsons, P. & Haworth, R. “Player–game interaction and cognitive gameplay: A taxonomic framework for the core mechanic of videogames.” Informatics, 2017, 4(1), 4. Multidisciplinary Digital Publishing Institute.

[3] Pfau, J; Smeddinck, J; Bikas, I and Malaka, R. “Bot or not? User Perceptions of Player Substitution with Deep Player Behavior Models.” In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. ACM

[4] Glavin, F; Madden, M. “Skilled Experience Catalogue : A Skill-Balancing Mechanism for Non-Player Characters using Reinforcement Learning.” 2018 IEEE Conf Comput Intell Games 2018:1–8.

[5] Warpefelt, H. “The Non-Player Character: Exploring the Believability of NPC Presentation and Behavior,” Ph.D. dissertation, Stockholm University, 2016.

[6] Hinkkanen, T; Kurhila, J and Pasanen, T. "Framework for evaluating believability of non-player characters in games", Workshop on Artificial Intelligence in Games, pp. 40, 2008.

[7] Togelius, J; Yannakakis, G; Karakovskiy, S; Shaker, N. “Assessing Believability.” In Believable Bots; Hingston, P., Ed.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 215–230

[8] Kersjes, H. and Spronck, P. “Modeling Believable Game Characters.” In IEEE Conference on Computational Intelligence and Games (Santorini, Greece, Sep.20-23), 2016. IEEE, 193-200.

[9] Hamdy, S and King, D. “Affect and Believability in game characters.” GameOn 2017, Carlow Institute of Technology, Carlow, Ireland.

[10] Bakkes, S; Spronck, P and van den Herik, J: “Rapid and reliable adaptation of video game AI.” IEEE Trans. Comput. Intell. AI Games 1(2), 93–104 (2009)

[11] Delatorre, P; León, C; Salguero, A; Palomo-Duarte, M and Gervás, P. "The long path of frustration: A case study with dead by daylight" in Advances in Computational Intelligence, Madrid, Spain: Springer, vol. 10306, pp. 69-680, Jun. 2017.

[12] Duh, H. and Chen, V. “Cheating behaviors in online gaming”. in Lecture Notes in Computer Science, 2009, Vol. 5621. Berlin: Springer, 567–573.

Employment of temporary workers and use of overtime to achieve volume flexibility using master production scheduling: monetary and social implications

Marco Trost
Thorsten Claus
Technische Universität (TU) Dresden
International Institute (IHI) Zittau
Markt 23, 02763 Zittau, Germany
E-mail: Marco.Trost@tu-dresden.de (corresp. author)

Frank Herrmann
Ostbayerische Technische Hochschule Regensburg
Innovation and Competence Centre for Production Logistics and Factory Planning (IPF)
P. O. Box 12 03 27, 93025 Regensburg, Germany

KEYWORDS

master production scheduling; temporary worker; overtime; flexibility cost; social sustainability

ABSTRACT

Flexibility and in particular volume flexibility is an important topic for industrial manufacturing companies. In this context, the harmonization of the available and required capacity is a central task, especially with increasing fluctuations in customer demand. In classical approaches, this is considered only by the use of additional capacities and there are only a few approaches that combine aspects of personnel planning with production planning. Therefore, this article presents a linear optimization model for master production scheduling that includes aspects of personnel requirements planning. It is used to investigate different strategies for the use of overtime and temporary workers in order to achieve different levels of volume flexibility. With regard to the monetary and social impacts, the results indicate that overtime has a stronger influence to achieve volume flexibility than the use of temporary workers. However, both are affected by substantial deficits in human working conditions. But the results also imply a promising potential for improving the social aspects without a significant increase in costs.

INTRODUCTION

Environmental uncerntainty, the increasing variability of products and processes require a high degree of flexibility from industrial production companies (Jain et al. 2013). Thus, companies have to meet shorter delivery times and life cycles, a wider range of products as well as incresead customization (Toni and Tonchia 1998). The wider range of products and the shorter product life cycle have increased the fluctuations in demand (Francas et al. 2011). To overcome these fluctuations, companies have to achieve volume flexibility. For this, balancing available capacity with capacity requirements has become important (Francas et al. 2011). Companies implement a variety of strategies to attain this harmonization. These include the use of overtime and employment of temporary workers (Qin et al. 2015). An intuitive use of these flexibility measures can be disadvantageous, but a reasoned implementation can reduce costs (Hemig et al. 2014).
In conventional approaches of hierarchical production planning, as described e.g. in Herrmann and Manitz (2015), a harmonization of capacities is realized on the levels of Aggregate Production Planning (APP) and Master Production Scheduling (MPS) by pre-production and additional capacities. Further approaches integrate aspects of personnel requirement planning. However, the flexibility measures are not analysed in terms of volume flexibility as well as their monetary implications. For this reason, this paper investigates the employment of temporary workers and the use of overtime hours at the level of MPS. Different levels of the required volume flexibility are modelled by normally distributed customer demands with different standard deviations. The flexibility costs of the different fluctuations in demand are determined, whereby temporary workers and overtime are permitted or not permitted as flexibility strategies.
Additionally, sustainable developments have become increasingly important in research and industry. This is driven by various interest groups like environmental activists as well as government agencies and other factors like a shortage of skilled workers. For labor-intensive processes, the available capacity is primary defined by the number of employees and their utilization. Therefore, measures to achieve volume flexibility particularly influence the human working conditions. Thus, this paper also investigates the impact of flexibility measures on social aspects. These include the employee utilization, deviations from regular working hours, the amount of overtime and the share of temporary workers.
For this, the paper is structured as follows. Section 2 presents a literature review. In section 3 the optimization model is described and the case study as well as the investigated flexibility strategies are introduced in section 4. The results and the discussion are outlined in section 5. Finally, a conclusion is presented in section 6.

LITERATURE REVIEW

Flexibility at industrial manufacturing companies include several dimensions where there is no general agreement on its definition (Saleh et al. 2009; Yu et al. 2015). One reason for this is that each company has its individual understanding of flexibility (Jain et al. 2013). For an overview of the most common flexibility definitions used by different authors, the reader is referred to Jain et al. (2013). In the literature, the flexibility dimensions are assigned to different classes by several authors (e.g. Sethi and Sethi 1990; Koste and Malhotra 1999). Koste and Malhotra (1999) distinguish between the following four

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

levels: individual resource level; shop floor level; plant level and functional level. This paper deals with the increasing fluctuations in demand, which require an appropriate volume flexibility of manufacturing companies. This volume flexibility is related to the plant level (Koste and Malhotra 1999) and has become an important competitive strategy (Jack and Raturi 2002). Volume flexibility is a measure of the ability of a production system to efficiently adapt to changing demands in response to changing socio-economic conditions (Jack and Raturi 2002; Sillekens et al. 2011; Jain et al. 2013).

The plant level of flexibility is reflected within the operational production planning. APP and MPS are the core elements of operational production planning (Günther and Tempelmeier 2020). The aim is to satisfy the fluctuating demand for the finished products with existing resources by defining production programs and determining the utilization of resources (Günther and Tempelmeier 2020). For this medium-term planning horizon, the working time flexibility of employees is of particular importance (Sillekens et al. 2011). The deployment of the workforce is organized, for instance, via shift models and working time accounts, and temporary workers additionally have a significant importance (Sillekens et al. 2011). A general working time flexibility is represented in basic models of APP and MPS by the use of additional capacities. However, concrete aspects such as legal restrictions or the costs of building up or removing resources are not considered. In addition, no distinction is made between concrete measures (e.g. use of temporary workers, use of overtime), which limits an analysis of concrete strategies for achieving volume flexibility.

Therefore, there are a few approaches which present an extension of the basic models in this respect by adding aspects of human resource requirements planning to the operational production planning. Hemig et al. (2014) consider an integrated production and workforce scheduling problem. The focus is on generating a minimum-cost schedule for the production of a forecasted demand, taking into account the application of volume flexibility tools. The problem is modeled as a (nonlinear) mixed-integer program and solved using dynamic programming. Bose et al. (2016) consider the strategic capacity planning in a multi-product, multi-plant configuration under demand uncertainty. A two-stage stochastic programming model is presented to determine capacity and product-plant configuration to maximize expected profit. The model is solved to understand the effect of product-plant configurations on expected profit and investment in capacity. Treber et al. (2016) present an approach for the management of production networks. The focus is on capacity planning and the use of tools that make the workforce more flexible. In particular, aspects of workforce flexibility are mapped via a mathematical optimization model. Furthermore, a function is implemented for accounting errors in the forecasted demand.

While the monetary evaluation of different flexibility strategies is established, impacts regarding social sustainability have not been considered so far. As already mentioned, flexible deployment of employees is necessary to ensure the required volume flexibility. But, this directly influences the human working conditions and is thus an influencing factor with regard to social sustainability. From a survey of works council ermerges, that there are deficitis in working conditions, e.g. the intensity of work, the pressure to perform, the number of overtime hours and the deviations from standard working hours are described as significant problems (Ahlers 2017). Further, the use of temporary workers can lead to social inequalities, as temporary workers are disadvantaged in terms of income and career mobility (Giesecke and Groß 2004). Furthermore, temporary workers seem to be exposed to higher psychological stress due to the uncertain work perspective (Virtanen et al. 2005). According to Nerdinger et al. (2014), the consequences of unhealthy working conditions can be an increased heart rate, frustration or increased errors, which in the long term leads to psychosomatic illnesses, resignation and demotivation. The BKK Health Report 2017, for example, attributes 25% of lost work days to musculoskeletal disorders and 16% to mental illnesses (Knieps and Pfaff 2017). In addition, the DAK Health Report 2018 shows an increase of more than 160% in days lost from work between 1997 and 2017 (Storm 2018). Furthermore, besides these significant consequences for health, the absence of employees also restricts the volume flexibility of companies.

To the best of the authors' knowledge, there are no papers that combining operational production planning with aspects of human resource requirements planning and considering the monetary and social effects of using temporary workers and overtime to ensure volume flexibility.

MODEL FORMULATION

The linear optimization model presented here is based on Trost (2018) and Trost et al. (2020) where a control of work intensity and basic aspects of personnel requirements planning are included. To ensure the volume flexibility, the building and reducing of available capacity (employees) is integrated. The following notation is used:

Sets

$EG = \{1, \dots, EG\}$	set of employee groups, indexed by *eg*
$J = \{1, \dots, J\}$	set of production segments, indexed by *j*
$K = \{1, \dots, K\}$	set of products, indexed by *k*
$T = \{1, \dots, T\}$	set of time periods, indexed by *t*
$Z = \{0, \dots, Z\}$	set of lead-time periods for capacity load, indexed by *z*

Parameters

$Capa_{eg}$	available capacity per period and employee of employee group *eg*
$d_{k,t}$	demand per product *k* in period *t*
$f_{z,j,k}$	capacity load factors for lead-time period *z*, production segment *j* and product *k*
h_k	inventory holding costs per unit and period for product *k*

I_k^{Init}	initial inventory level for product *k*
I_k^{Max}	maximum inventory level for product *k*
m_{eg}^{Cost}	cost rate for hiring an employee from employee group *eg*
n_{eg}^{Cost}	cost rate for layoff an employee from employee group *eg*
R_j^{Max}	maximum permitted employee utilisation per production segment *j*
$Staff_{eg}^{Cost}$	cost rate per employee of employee group *eg*
$Staff_{eg,j}^{Init}$	initial number of employees per employee group *eg* and production segment *j*
$Staff_{eg,j}^{Max}$	maximum number of employees per employee group *eg* in production segment *j*
V	number of periods for overtime balancing
we_{eg}	lead-time periods for hiring employees of employee group *eg*
wf_{eg}	lead-time periods for employee turnover of employee group *eg*

Decision Variables

$a_{j,t}$	available capacity per production segment *j* in period *t*
$b_{j,t}$	capacity requirement per production segment *j* in period *t*
$I_{k,t}$	inventory level per product *k* in period *t*
$m_{eg,j,t}$	number of hired employees of employee group *eg* in production segment *j* and period *t*
$n_{eg,j,t}$	number of layoffs of employee group *eg* in production segment *j* and period *t*
$overtime_{j,t}$	used overtime per production segment *j* and period *t*
$Staff_{eg,j,t}$	number of employees of employee group *eg*, production segment *j* and period *t*
$x_{k,t}$	production quantity per product *k* in period *t*

Objective Function

The objective function minimizes the total costs from inventories, employees, and worker hiring as well as layoff (see equation (1) to equation (6)).

$$Objective\ Function: Minimize\ (TotalCosts) \quad (1)$$

$$TotalCosts = InventoryCosts + StaffingCost + HiringCosts + LayoffCosts \quad (2)$$

$$InventoryCosts = \sum_{t=1}^{T}\sum_{k=1}^{K} h_k \cdot I_{k,t} \quad (3)$$

$$StaffingCosts = \sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{eg=1}^{EG} Staff_{eg}^{Cost} \cdot Staff_{eg,j,t} \quad (4)$$

$$HiringCosts = \sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{eg=1}^{EG} m_{eg}^{Cost} \cdot m_{eg,j,t} \quad (5)$$

$$LayoffCosts = \sum_{t=1}^{T}\sum_{j=1}^{J}\sum_{eg=1}^{EG} n_{eg}^{Cost} \cdot n_{eg,j,t} \quad (6)$$

Constraints

First, as general constraints there are the inventory balance sheet (equation (7)), the definition of the initial and maximum inventory level (equation (8) and equation (9)) and equation (10) determine the capacity requirements.

$$x_{k,t} + I_{k,t-1} - I_{k,t} = d_{k,t} \quad \forall\ 1 \le k \le K;\ \forall\ 1 \le t \le T \quad (7)$$

$$I_{k,t=0} = I_k^{Init} \quad \forall\ 1 \le k \le K \quad (8)$$

$$I_{k,t} \le I_k^{Max} \quad \forall\ 1 \le k \le K;\ \forall\ 1 \le t \le T \quad (9)$$

$$\sum_{z=0}^{Z}\sum_{k=1}^{K} f_{z,j,k} \cdot x_{k,t+z} = b_{j,t} \quad \forall\ 1 \le j \le J;\ \forall\ 1 \le t \le (T - Z) \quad (10)$$

The aspects of human resource requirements planning are modelled as follows. The available capacity is integrated by equation (11), the employee hiring and layoff by the employee balance sheet (equation (12)) and the initial employee level by equation (13). Between the regular and temporary employees are distinguished by different employee groups (EG) and also different lead times for hiring (we_{eg}) and layoffs (wf_{eg}) are modelled. Equation (14) represents that the available number of (skilled) employees is limited on the labour market.

$$\sum_{eg=1}^{EG} Staff_{eg,j,t} \cdot Capa_{eg} = a_{j,t} \quad \forall\ 1 \le j \le J;\ \forall\ 1 \le t \le T \quad (11)$$

$$Staff_{eg,j,t} = Staff_{eg,j,t-1} + m_{eg,j,t-we_{eg}} - n_{eg,j,t-wf_{eg}} \quad \forall\ 1 \le eg \le EG;\ \forall\ 1 \le j \le J;\ \forall\ 1 \le t \le T \quad (12)$$

$$Staff_{eg,j,t=0} = Staff_{eg,j}^{Init} \quad \forall\ 1 \le eg \le EG;\ \forall\ 1 \le j \le J \quad (13)$$

$$Staff_{eg,j,t} \le Staff_{eg,j}^{Max} \quad \forall\ 1 \le eg \le EG;\ \forall\ 1 \le j \le J;\ \forall\ 1 \le t \le T \quad (14)$$

With regard to the consideration of overtime the maximum utilization of employees is modelled by equation (15). Overtime can occur if the maximum utilization (R_j^{Max}) is over 100 %. The control of overtime is achieved by equation (16) to equation (18). However, overtime do not result in additional costs because they have to be compensated within a specific time interval (by equation (17)) which meets legal restrictions. When the maximum utilization is less than 100% the equation (16) to equation (18) are not restrictive.

$$R_j^{Max} \cdot a_{j,t} \geq b_{j,t} \quad \forall\, 1 \leq j \leq J;\ \forall\, 1 \leq t \leq (T-Z) \tag{15}$$

$$b_{j,t} - a_{j,t} = overtime_{j,t} \quad \forall\, 1 \leq j \leq J;\ \forall\, 1 \leq t \leq (T-Z) \tag{16}$$

$$\sum_{t'=t-V}^{t} overtime_{j,t'} \leq 0 \quad \forall\, 1 \leq j \leq J;\ \forall\, 1 \leq t \leq (T-Z) \tag{17}$$

$$\sum_{t'=0-V}^{t=0} overtime_{j,\,t'} = 0 \quad \forall\, 1 \leq j \leq J \tag{18}$$

Finally, the non-negative conditions and the integer conditions are defined in equation (19) and equantion (20).

$$a_{j,t}, b_{j,t}, I_{k,t}, m_{eg,j,t}, n_{eg,j,t}, Staff_{eg,j,t}, x_{k,t} \geq 0 \quad \forall\, 1 \leq eg \leq EG;\ \forall\, 1 \leq j \leq J;\ \forall\, 1 \leq k \leq K;\ \forall\, 1 \leq t \leq T \tag{19}$$

$$Staff_{eg,j,t} \in \{\mathbb{Z}\} \quad \forall\, 1 \leq eg \leq EG;\ \forall\, 1 \leq j \leq J;\ \forall\, 1 \leq t \leq T \tag{20}$$

CASE STUDY AND FLEXIBILITY STRATEGIES

The case study considered here is based on Trost et al. (2019). At first, general parameters are presented in Table 1. The different employee groups (EG) represent the regular ($eg = 1$) and temporary employees ($eg = 2$).

Table 1: General Parameters

Parameter	Value
EG	2
J	2
K	2
W	3
Z	1

Further, Table 2 presents the employee related parameters for regular and temporary workers, which are based on a Saxon railway company with a IG Metal collective agreement for the metal and electrical industry. Due to the experience gap of temporary workers, they have a lower available capacity ($Capa_{eg}$ in seconds) than regular workers, which means that for given capacity load factors a lower productivity is depicted. However, the hiring and layoffs of temporary workers are outsourced to an external service provider, resulting in shorter lead times (we_{eg} and wf_{eg} in periods) and lower cost rates (m_{eg}^{Cost} and n_{eg}^{Cost}). But due to the agency fees, the cost rate per employee and period ($Staff_{eg}^{Cost}$) are higher for temporary workers than for regular workers.

Table 2: Employee parameters per worker class (eg)

Parameter	$eg = 1$	$eg = 2$
$Capa_{eg}$	524 400	393 300
m_{eg}^{Cost}	15 000	1 500
n_{eg}^{Cost}	60 000	100
we_{eg}	3	1
wf_{eg}	3	0
$Staff_{eg}^{Cost}$	3 671	5 435

Finally, Table 3 contains the inventory cost rate (h_k), the maximum inventory level (I_k^{Max} in quantity units) and the capacity load factors ($f_{z,j,k}$ in seconds). Note, that the capacity load only occur in lead-time period $z = 1$.

Table 3: Further parameters

Parameter		$k = 1$	$k = 2$
h_k		115	165
I_k^{Max}		30 000	37 500
$f_{z=1,j,k}$	$j = 1$	3 867	4 092
	$j = 2$	13 976	10 184

In order to investigate several required volume flexibility situations, different customer demands ($d_{k,t}$) are distinguished. A constant demand with 40 000 units of product one and 50 000 units of product two is the initial demand situation. Based on this, three normally distributed demand courses with a coefficient of variation from 5 %, 10 % and 20 % are regarded. To achieve this volume flexibility, the following four strategies are applied:

- **Strategy 1** enables the use of regular employees and temporary workers ($Staff_{eg,j=1}^{Max} = 1\,500$ and $Staff_{eg,j=2}^{Max} = 4\,500$), as well as 20 % overtime hours related to the regular working time ($R_j^{Max} = 1.2$). The overtime hours have to be compensated in accordance with the working time law § 3 (Germany) within an half a year ($V = 5$).
- Within **Strategy 2**, temporary workers cannot be employed ($R_j^{Max} = 1.2$, $Staff_{eg=1,j=1}^{Max} = 1\,500$, $Staff_{eg=1,j=2}^{Max} = 4\,500$ and $Staff_{eg=2,j}^{Max} = 0$).
- For **Strategy 3**, overtime cannot be used ($R_j^{Max} = 1.0$, $Staff_{eg,j=1}^{Max} = 1\,500$ and $Staff_{eg,j=2}^{Max} = 4\,500$).
- Finally, in **Strategy 4** neither overtime nor temporary workers are permitted ($R_j^{Max} = 1.0$, $Staff_{eg=1,j=1}^{Max} = 1\,500$, $Staff_{eg=1,j=2}^{Max} = 4\,500$ and $Staff_{eg=2,j}^{Max} = 0$).

RESULTS AND DISCUSSION

For this investigation, the results of the four strategies considered for achieving volume flexibility are compared. Strategy 1, which allows the use of temporary workers as well as overtime, serves as a benchmark. These strategies are applied to the four demand courses. In order to increase the statistical significance, five random demand series are realized for each normally distributed demand course. In total, 64 different planning problems are considered. Each planning problem regard a planning horizon from 84 month ($T = 84$). However, a 6-month warm up as well as run out phase are taken into account, so that the results from 72 months are analysed ($\hat{T} = 72$). The results were obtained using CPLEX 12.10 on a 3.30 GHz PC with 192 GByte RAM. Each problem could be solved within 12.44 seconds on average.

First, the monetary effects of the different strategies per demand course are considered. Table 4 presents the total costs for Strategy 1 as well as the relative deviations of the other strategies per demand course. Unrelated to the used flexibility strategy, the total costs increase with increasing required volume flexibility. These flexibility costs results from an increased pre-production, an increased number of employed regular as well as temporary workers and increased adjustments in the number of employees. Thus, an increase of all cost components (inventory costs, staffing costs, hiring costs and layoff costs) can be observed.

Table 4: Total costs for strategy 1 and relative deviations of the further strategies for each demand course

	Demand courses			
	Constant	5 % coefficient of variation	10 % coefficient of variation	20 % coefficient of variation
Strategy 1	719 607 417 MU	730 834 953 MU	746 430 404 MU	770 407 813 MU
Strategy 2	+ 0.04 %	+ 0.26 %	+ 1.19 %	+ 0.96 %
Strategy 3	+ 0.03 %	+ 0.87 %	+ 2.24 %	+ 3.68 %
Strategy 4	+ 0.04 %	+ 1.75 %	+ 4.10 %	+ 6.26 %

In more detail, it emerges that for a constant demand course the waiver of temporary workers and/or overtime hours have only a small monetary impact from maximum 0.04 %. Further, within Strategy 2 (no temporary workers) a moderate increase in costs for all demand courses from maximum 1.19 % occur. In comparison to this, Strategy 3 (no overtime) result in higher flexibility costs than Strategy 2. Accordingly, not using overtime is associated with higher flexibility costs than not using temporary workers. It is concluded that the use of overtime has a greater contribution for achieving volume flexibility. The waiver of overtime and temporary workers (Strategy 4) cause increased flexibility costs up to 6.26 %. Thus, the (partial) waiver of the considerd flexibility measures result in negative monetary effects. However, in some cases, the cost increase is low, especially when the required volume flexibility is low. With higher required volume flexibility, there is also a low increase in costs if only the the employment of temporary workers is not permitted.

Concerning the social impact, we consider the before mentioned deficits in human working conditions. For this, Table 5 reports the average worker utilization to assess the work intensity ($\overline{U}_j$) and the amplitude of worker utilization to assess the deviations of standard working hours ($U_j^{Max-Min}$). Table 6 present the use of overtime for Strategy 1 and Strategy 2 by the share of periods with used overtime ($\hat{T}_j^{OT}$) as well as the mean ($\overline{OT}_j$) and maximum (OT_j^{Max}) overtime within these periods related to the regular working time. For Strategy 3 and Strategy 4 overtime is excluded.

Table **7** reports the employment of temporary workers for Strategy 1 and Strategy 3 by the share of periods in which temporary workers are employed ($\hat{T}_j^{TW}$), the average share of temporary workers within these periodes ($\overline{TW}_j$) and the maximum share of temporary workers (TW_j^{Max}). For Strategy 2 and Strategy 4 temporary workers are excluded. All results refer to production segment one, note that production segment two contains analogous effects.

Table 5 to

Table **7** indicate that with higher required volume flexibility a higher deviations from regular working hours, use of overtime and employment of temporary workers occur. However, the average employee utilization decreases with increasing volume flexibility and with the absence of (single) flexibilty measures (Strategy 2 to Strategy 4) further reduction in average utilization occur.

Table 5: Mean utilization ($\overline{U}_j$) and amplitude of utilization ($U_j^{Max-Min}$) for production segment one

	Demand courses							
	Constant		5 % coefficient of variation		10 % coefficient of variation		20 % coefficient of variation	
	$\overline{U}_{j=1}$	$U_{j=1}^{Max-Min}$	$\overline{U}_{j=1}$	$U_{j=1}^{Max-Min}$	$\overline{U}_{j=1}$	$U_{j=1}^{Max-Min}$	$\overline{U}_{j=1}$	$U_{j=1}^{Max-Min}$
Strategy 1	100.00 %	0.14 %	99.04 %	14.28 %	98.26 %	31.75 %	96.76 %	51.06 %
Strategy 2	99.87 %	0.00 %	98.69 %	16.22 %	97.01 %	33.93 %	95.24 %	55.11 %
Strategy 3	99.87 %	0.03 %	98.84 %	7.84 %	97.71 %	17.17 %	95.69 %	29.34 %
Strategy 4	99.87 %	0.00 %	97.65 %	10.32 %	95.06 %	21.64 %	91.63 %	37.41 %

Table 6: Share of overtime periods ($\hat{T}_j^{OT}$) as well as mean ($\overline{OT}_j$) and maximum overtime (OT_j^{Max}) for production segment within these periods and related to the regular working time

	Demand courses											
	Constant			5 % coefficient of variation			10 % coefficient of variation			20 % coefficient of variation		
	$\hat{T}_{j=1}^{OT}$	$\overline{OT}_{j=1}$	$OT_{j=1}^{Max}$	$\hat{T}_{j=1}^{OT}$	$\overline{OT}_{j=1}$	$OT_{j=1}^{Max}$	$\hat{T}_{j=1}^{OT}$	$\overline{OT}_{j=1}$	$OT_{j=1}^{Max}$	$\hat{T}_{j=1}^{OT}$	$\overline{OT}_{j=1}$	$OT_{j=1}^{Max}$
Strategy 1	79.17 %	0.02 %	0.03 %	35.56 %	2.19 %	5.99 %	40.83 %	4.28 %	13.68 %	41.11 %	7.43 %	19.15 %
Strategy 2	0.00 %	0.00 %	0.00 %	33.89 %	2.34 %	6.92 %	33.89 %	4.41 %	13.81 %	36.39 %	7.39 %	19.79 %

Table 7: Share of periods with temporary workers ($\hat{T}_j^{TW}$) as well as mean ($\overline{TW}_j$) and maximum (TW_j^{Max}) share of temporary workers for production segment one within these periods and related to all employees

	Demand courses											
	Constant			5 % coefficient of variation			10 % coefficient of variation			20 % coefficient of variation		
	$\hat{T}_{j=1}^{TW}$	$\overline{TW}_{j=1}$	$TW_{j=1}^{Max}$	$\hat{T}_{j=1}^{TW}$	$\overline{TW}_{j=1}$	$TW_{j=1}^{Max}$	$\hat{T}_{j=1}^{TW}$	$\overline{TW}_{j=1}$	$TW_{j=1}^{Max}$	$\hat{T}_{j=1}^{TW}$	$\overline{TW}_{j=1}$	$TW_{j=1}^{Max}$
Strategy 1	20.83 %	0.15 %	0.15 %	30.56 %	1.82 %	6.63 %	33.33 %	4.08 %	10.57 %	36.11 %	6.01 %	17.65 %
Strategy 3	0.00 %	0.00 %	0.00 %	40.39 %	2.63 %	10.82 %	45.00 %	4.45 %	15.01 %	44.44 %	7.68 %	23.39 %

More in detail, with Strategy 1 and constant demand overtime and temporary workers are used. Because the available capacity from the optimal number of employees are not sufficient to satisfy the demand, a small amount of overtime is required. Since this has to be compensated within 6 months, the employment of a temporary worker is necessary in some periods. By not using (single) flexibility measures (Strategy 2 to Strategy 4) and a constant demand, the optimal number of employees is increased and no additional overtime or temporary workers are required. However, the average employee utilization is lower. With normally distributed demand courses, permitting overtime (strategy 1 and strategy 2) leads to significantly higher fluctuations in working hours and, in some cases, to an extensive use of overtime. If the use of temporary workers is not permitted in this context (strategy 2), working time fluctuations are even higher. For example, the maximum overtime of 19.79 % that occur at Strategy 2 and 20 % demand variation correspond to approximately 7.5 hours per week for a 38-hour work week. Accordingly, the fluctuations in working hours of 55.11 % correspond to varying weekly working hours of approximately 24.6 to 45.5 hours per week in relation to a standard working time of 38 hours per week. With the avoidance of overtime (Strategy 3 and Strategy 4) these fluctuations in working hours decrease. However, if the employment of temporary workers is permitted while overtime is avoided (Strategy 3), an increased employment of temporary workers occur as well. For this, even with low fluctuations in demand, the share of temporary workers is in some cases higher than 10 % and temporary workers are employed in more than 40 % of the periods. Strategy 4 excludes the use of overtime and the employment of temporary workers, which corresponds to a stronger social orientation and result, in some cases, in a significantly lower average employee utilization. In comparison to Strategy 1, which allows overtime and temporary workers, the fluctuations in working hours can be reduced by Strategy 4 as well. However, the reduction is not as strong as in Strategy 3, in which only overtime is avoided.

In summary, the use of overtime and temporary workers can reduce flexibility costs compared to not using these flexibility measures. However, in some cases, the achieved cost savings are small. But the social impact of these measures is negative. The avoidance of single flexibility measures does not lead to a comprehensive improvement of the social aspects, since restrictions of single social characteristics are compensated by the deterioration of other social aspects. Thus, we suggest that it might be more beneficial to limit for certain social aspects, e.g. fluctuations in working hours and the share of temporary workers, than to avoid it. The described monetary effects from avoiding a measure lead us to expect a corresponding potential for social improvements without a significant increase in costs.

CONCLUSION

The article demonstrated the monetary and social impact of the employment of temporary workers and the use of overtime. From the presented literature review it was pointed out that there are only a few approaches that consider production planning and personnel planning simultaneously. Further there are deficits in investigations regarding the impact on the use of overtime and temporary workers to archive volume flexibility. To address this gap, a linear optimization model for MPS was presented. The results indicate that ensuring the necessary volume flexibility impacts the human working conditions. Particularly in the case of higher fluctuations in demand, there are strong deviations from regular working hours, including frequent and, in some cases, extensive use of overtime. In addition, there is a frequent employment of temporary workers and, in some cases, a high share of temporary workers. The monetary impact of avoiding the flexibility measures demonstrate that this not necessarily cause a significant increase in costs even with higher fluctuations in demand. However, the use of overtime

contributes more to achieving volume flexibility than the use of temporary workers, as indicated by higher flexibility costs for overtime avoidance than for temporary workers avoidance.
Finally, the working conditions might be improved by limiting specific social aspects without causing a significant increase in costs. The investigation of suitable limits is left for future work.

REFERENCES

Ahlers, E. (2017): Work and health in German companies. Findings from the WSI works councils survey 2015. In: *WSI Institute of Economic and Social Research* Report No. 33e.

Bose, D.; Chatterjee, A. K.; Barman, S. (2016): Towards dominant flexibility configurations in strategic capacity planning under demand uncertainty. In: *OPSEARCH* 53 (3), pp. 604–619.

Francas, D.; Löhndorf, N.; Minner, S. (2011): Machine and labor flexibility in manufacturing networks. In: *International Journal of Production Economics* 131 (1), pp. 165–174.

Giesecke, J.; Groß, M. (2004): External labour market flexibility and social inequality. In: *European Societies* 6 (3), pp. 347–382.

Günther, H. O.; Tempelmeier, H. (2020): Supply Chain Analytics. Operations Management und Logistik. Norderstedt: Books on Demand.

Hemig, C.; Rieck, J.; Zimmermann, J. (2014): Integrated production and staff planning for heterogeneous, parallel assembly lines: an application in the automotive industry. In: *International Journal of Production Research* 52 (13), pp. 3966–3985.

Herrmann, F.; Manitz, M. (2015): Ein hierarchisches Planungskonzept zur operativen Produktionsplanung und -steuerung. In: Th. Claus, F. Herrmann und M. Manitz (Ed.): Produktionsplanung und -steuerung. Forschungsansätze, Methoden und deren Anwendungen. Wiesbaden: Springer Gabler, pp. 7–22.

Jack, E. P.; Raturi, A. (2002): Sources of volume flexibility and their impact on performance. In: *Journal of Operations Management* 20 (5), pp. 519–548.

Jain, A.; Jain, P. K.; Chan, F. T.S.; Singh, S. (2013): A review on manufacturing flexibility. In: *International Journal of Production Research* 51 (19), pp. 5946–5970.

Knieps, F.; Pfaff, H. (Ed.) (2017): Digitale Arbeit - Digitale Gesundheit. BKK-Gesundheitsreport 2017. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft.

Koste, L. L.; Malhotra, M. K. (1999): A theoretical framework for analyzing the dimensions of manufacturing flexibility. In: *Journal of Operations Management* 18 (1), pp. 75–93.

Nerdinger, F. W.; Blickle, G.; Schaper, N. (2019): Arbeits- und Organisationspsychologie. 4th edition Berlin, Heidelberg: Springer.

Qin, R.; Nembhard, D. A.; Barnes II, W. L. (2015): Workforce flexibility in operations management. In: *Surveys in Operations Research and Management Science* 20 (1), pp. 19–33.

Saleh, J. H.; Mark, G.; Jordan, N. C. (2009): Flexibility: a multi-disciplinary literature review and a research agenda for designing flexible engineering systems. In: *Journal of Engineering Design* 20 (3), pp. 307–323.

Sethi, A. K.; Sethi, S. P. (1990): Flexibility in manufacturing: A survey. In: *Int J Flex Manuf Syst* 2 (4).

Sillekens, T.; Koberstein, A.; Suhl, L. (2011): Aggregate production planning in the automotive industry with special consideration of workforce flexibility. In: *International Journal of Production Research* 49 (17), pp. 5055–5078.

Storm, A. (Ed.) (2018): DAK-Gesundheitsreport 2018. Beiträge zur Gesundheitsökonomie und Versorgungsforschung (21).

Toni, A. de; Tonchia, S. (1998): Manufacturing flexibility: A literature review. In: *International Journal of Production Research* 36 (6), pp. 1587–1617.

Treber, S.; Moser, E.; Lanza, G. (2016): Workforce Flexibility in Production Networks: Mid-Term Capacity Planning Illustrated by an Example of the Automotive Industry. In: *AMR* 1140, pp. 427–434.

Trost, M. (2018): Master production scheduling with integrated aspects of personnel planning and consideration of employee utilization specific processing times. In: *Proceedings of the 32nd ECMS International Conference on Modelling and Simulation. Wilhelmshaven, Germany*, pp. 329–335.

Trost, M.; Claus, Th.; Herrmann, F. (2019): Adapted master production scheduling: Potential for improving human working conditions. In: *Proceedings of the 33rd ECMS International Conference on Modelling and Simulation. Caserta, Italy*, pp. 310–316.

Trost, M.; Claus, Th.; Herrmann, F. (2020): Influence of company sizes in adapted master production scheduling for improving human working conditions. In: *Proceedings of the 34th ECMS International Conference on Modelling and Simulation. Wildau, Germany*, pp. 287–293.

Virtanen, M.; Kivimäki, M.; Joensuu, M.; Virtanen, P.; Elovainio, M.; Vahtera, J. (2005): Temporary employment and health: a review. In: *International journal of epidemiology* 34 (3), pp. 610–622.

Yu, K.; Cadeaux, J.; Luo, B. N. (2015): Operational flexibility: Review and meta-analysis. In: *International Journal of Production Economics* 169, pp. 190–202.

AUTHORS BIOGRAPHY

MARCO TROST is a phd-student and research associate at the professorship for Production and Information Technology at the International Institute (IHI) Zittau, a central academic unit of Technische Universität (TU) Dresden. His e-mail address is: Marco.Trost@tu-dresden.de.

PROFESSOR DR. THORSTEN CLAUS holds the professorship for Production and Information Technology at the International Institute (IHI) Zittau, a central academic unit of Technische Universität (TU) Dresden and he is the director of the International Institute (IHI) Zittau. His e-mail address is: Thorsten.Claus@tu-dresden.de.

PROFESSOR DR. FRANK HERRMANN holds the professorship for operative production planning and control at the OTH Regensburg and he is the head of the Innovation and Competence Centre for Production Logistics and Factory Planning (IPF). His e-mail address is: Frank.Herrmann@oth-regensburg.de

CHANGE DETECTION FOR AREA SURVEILLANCE USING A MOVING CAMERA

Tatsuhisa Watanabe, Tomoharu Nakashima, and Yoshifumi Kusunoki
Graduate School of Humanities and Sustainable System Sciences
Osaka Prefecture University
Gakuen-cho 1-1, Sakai, Osaka 599-8531, Japan
Email: {tatsuhisa.watanabe, tomoharu.nakashima, yoshifumi.kusunoki}@kis.osakafu-u.ac.jp

KEYWORDS

Change detection, Area surveillance, Monocular camera, Autonomous robot.

ABSTRACT

This paper tackles area surveillance with a moving camera by change detection. None of the existing datasets for change detection meets a surveillance scenario where a camera is mounted on a moving platform and pointed in the direction of moving. Thus, this paper creates a new dataset including several challenging points. For this dataset, this paper employs a composable method and proposes some components. To evaluate the proposed components, some corresponding classic methods were also tested on the dataset. As a result, the proposals outperformed them. Moreover, this paper investigated the relationship between the parameters of the components and their performance.

INTRODUCTION

Surveillance systems have been attracting attention because they have a great potential to reduce the workload of monitors. Surveillance systems can be applied to various fields such as urban monitoring, agriculture, and traffic analysis with manifold sensors. In recent years, some sensors have been mounted on Unmanned Aerial Vehicles or Unmanned Ground Vehicles (UGVs) as the development of industrial technologies. This paper aims to automatically observe areas for security using these autonomous vehicles. To do so, this paper overviews some previous methods and datasets in the following paragraphs and proposes a new dataset and method.

Regarding the area monitoring, there are two types of automatic methods: target-limited and target-agnostic. The target-limited methods focus on a specific anomaly including human behaviors (Morais et al. 2019; Singh et al. 2018). While the target-limited methods performed well, they cannot be a complete replacement for humans because they can only detect the expected target. The idea of combining them does not work because one cannot obtain or even list all possible abnormal patterns. The target-agnostic methods assume the available data as a distribution of normal situations and detect samples far from it as anomalies (Chu et al. 2019; Hao et al. 2019). The target-agnostic methods for area surveillance can be roughly divided into two groups based on situation types. One group is for the place where people are NOT supposed to be. Methods of this sort have to be able to detect the emergence or disappearance of anything. The other group is for the place where people appear. In such places, detectors are required to report anomalous behaviors of humans too. The former situation can be solved by change detection (CD) and the latter by anomaly detection. The former is more important in practice because if there are people, they can take action.

Although a large number of studies devised CD methods, all of them employed fixed cameras. While research of this kind plays an important role in surveillance, blind spots of the fixed cameras can arouse a controversy over security. The blind spots can be reduced by mounting cameras on a moving platform such as autonomous vehicles.

There are only four datasets for CD with moving vehicles. One dataset, known as VDAO (Silva et al. 2014), was created inside an offshore facility. A camera on a mobile robotic platform on a straight rail was used. This dataset contains 15 different abnormal objects such as bags. (Sakurada and Okatani 2015) constructed two datasets consisting of panoramic images: TSUNAMI and GSV. TSUNAMI captured scenes of tsunami-damaged areas in Japan. GSV is a collection of images on Google Street View. The last dataset is the so-called VL-CMU-CD dataset (Alcantarilla et al. 2018). It includes pictures taken in the city of Pittsburgh, PA, USA, over a year.

The four CD datasets with moving devices do not suffice for area surveillance. Every dataset but VDAO (i.e. TSUNAMI, GSV, and VL-CMU-CD) was not designed for the field of surveillance. Thus, their change types such as buildings are not anomalous. The VDAO dataset only contains abandoned objects, not humans. Moreover, they do not contain looming motion. Such motion is unavoidable if one employs a moving camera in a narrow place including a hallway. In response to the inadequacy of the existing datasets, a new CD dataset is created with a moving monocular camera on a hallway. Compared to the existing datasets, the proposed one has some challenging points: 1) looming motion, under which everything varies grad-

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

ually in size and position; 2) non-identical trajectory and inconsistent viewing angles, which cause parallax leading to false alarm; 3) illumination change due mainly to different times of the day. The dataset is detailed more in the EXPERIMENTS AND RESULTS section. To tackle the proposed dataset, a composable procedure is employed as in (Carvalho et al. 2019). However, it was proposed for the VDAO dataset, so three new components are designed for the proposed dataset: Video Compression (VC), Temporal Alignment (TA), and Frame Comparison (FC).

To evaluate the proposed TA and FC components, classic TA and FC methods are also tested. (Evangelidis and Bauckhage 2013) proposed a TA method using local descriptors. One of the TA datasets they tackled is similar to the proposed dataset, which contains looming motion. (Carvalho et al. 2019) proposed a structured method for the VDAO dataset and employed Zero-mean Normalized Cross Correlation (ZNCC) for dissimilarity calculation.

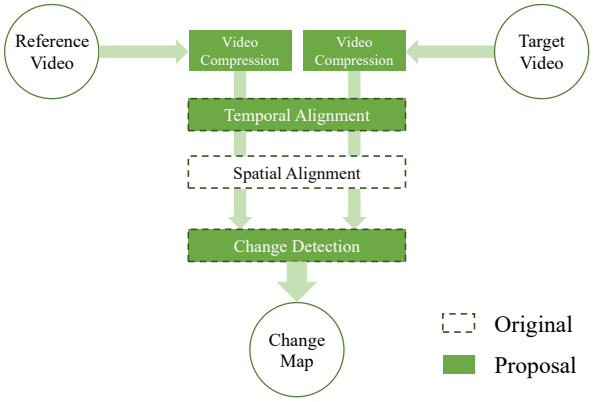
The contributions of this paper are two-fold.

- A dataset has been created with a moving monocular camera. This is the first video-CD dataset that has looming-motion for area surveillance.
- A structured way has been proposed to deal with the proposed dataset.

METHOD

This paper employs a composable procedure as in (Carvalho et al. 2019) with three new components. Figure 1 illustrates its whole procedure, where dotted line boxes indicate that the corresponding part was originally proposed in the literature, and filled-in boxes are the new components that this paper proposes. The input is a pair of videos: reference video and target one. The method detects changes in the target video against the reference one. The first process, VC, is a novel process for reducing the computation cost of the downstream processes. TA synchronizes a compressed version of the target video to that of the reference one. SA performs image registration between each target frame and the matched reference one. FC computes dissimilarity values of each matched frame pair and binarizes them with a threshold.

For each component, visual information of videos needs to be extracted. To do so, famous CNN architectures, called VGG13 and VGG16 (Simonyan and Zisserman 2014), are used. They consist of three types of layers: convolution, max pooling, and fully connected layer. In this paper, all the fully connected layers of VGGs are discarded because the components demand just spatial information, not features for classification. Training a deep network requires a considerable amount of data and time. Therefore, this paper exploits the pre-trained parameters of VGGs on the ImageNet dataset (Deng et al. 2009). For VC and TA, the last pooling layer of VGG16 is replaced with global average pooling (GAP) (Lin et al. 2013). GAP is considered to lose spatial information. However, it seems that this weakness potentially renders a matching method insensitive to parallax or too partial textures. The GAP version of VGG16, noted as VGG16', returns a C dimensional feature map for the input image.



Figures 1: The Whole Procedure of the Proposed Method

Proposed Video Compression (VC)

With the aim of area surveillance, it is not necessary to inspect all frames since several consecutive frames contain almost the same contents. Therefore, VC reduces redundant frames according to the similarity of sequential ones. Let $S^r = \{I^r_1, I^r_2, ..., I^r_{n_r}\}, S^t = \{I^t_1, I^t_2, ..., I^t_{n_t}\}$ be the reference sequence and the target one, respectively. n_r and n_t are the number of the reference frames and that of the target ones, respectively. The aim of this component is to retrieve key frame indices $X^r = \{i^r_1, i^r_2, ..., i^r_{n_{rc}}\}, X^t = \{i^t_1, i^t_2, ..., i^t_{n_{tc}}\}$. n_{rc} and n_{tc} are the number of the reference key frames and that of the target ones, respectively. Unless otherwise noted, all the frames but the key frames are to be disregarded in the ensuing processes. VC is performed in a chronological manner as in Algorithm 1, where $cos\_sim(v_1, v_2)$ is the cosine similarity value of two vectors v_1, v_2. Note that this step processes independently the target video and the reference one.

Algorithm 1 Video Compression

In: S, T_c // S: set of video frames. T_c: threshold.
Out: X // X: set of key frame indices.

array $X \leftarrow \{1\}$
$l \leftarrow 1$
while $l < S.length$ **do**
  for $h = l + 1$ to $S.length$ **do**
    if $cos\_sim(VGG16'(S[l]), VGG16'(S[h])) < T_c$ **then**
      X.append(h)
      break
    end if
  end for
  $l \leftarrow h$
end while
return X

Proposed Temporal Alignment (TA)

TA matches frames in the target video to those in the reference video. This paper proposes a new TA method using deep features. Let $\hat{X} = \{\hat{k}_1, \hat{k}_2, ..., \hat{k}_{n_{tc}}\}$ be the indices of the matched reference frames. $\hat{X}$ is calculated by Equation (1).

$$\hat{k_l} = \arg\max_h \ cos\_sim(VGG16'(I^t_{i^t_l}), VGG16'(I^r_{i^r_h})), \tag{1}$$

where $l = 1, 2, ..., n_{tc}$. To prevent abruption from the previously matched index, a search range is restricted as $h \in \{m^{back}_l, m^{back}_l + 1, ..., m^{front}_l\}$. m^{back}_l, m^{front}_l are calculated by Equations (2), (3), respectively.

$$m^{back}_l = \begin{cases} i^r_1 & \text{if } l = 1, \\ \max(i^r_1,\ \hat{k}_{l-1} - back) & \text{otherwise,} \end{cases} \tag{2}$$

$$m^{front}_l = \begin{cases} front & \text{if } l = 1, \\ \min(i^r_{n_{rc}},\ \hat{k}_{l-1} + front) & \text{otherwise,} \end{cases} \tag{3}$$

where $back$ and $front$ are hyperparameters, discussed in the EXPERIMENTS AND RESULTS section.

Spatial Alignment (SA)

It is impossible to take all videos in an exactly consistent angle or position. Consequently, even if TA perfectly synchronizes the input videos, the matched frame pairs will have different image planes. Thus, SA, or image registration, is performed with Homography transformation. Let $H(I_1, I_2)$ be a transformed image of I_1 to I_2. The outside of an image plane is treated as black (i.e. RGB value (0, 0, 0)), when it gets into the image plane by Homography transformation. This black area would be detected as change in the following stage. Therefore, the counterpart of the target frame is masked. I' signifies a masked version of an image I.

Proposed Frame Comparison (FC)

The last component of the proposed method is FC. FC compares the spatiotemporally aligned frame pairs in the upstream processes and calculates dissimilarity maps. After that, it computes change maps by binarizing the dissimilarity maps with a threshold. Dissimilarity maps are obtained based on the idea by (Kim et al. 2017). They proposed a similarity calculation method with the VGG13 network for template matching. Feature maps of a template image and a target one were extracted from a mid-convolutional-layer of the network. The target feature map was searched with a sliding window in order to determine the patch of the target image most similar to the template image. They used Normalized Cross Correlation (NCC) as a similarity criterion.

A feature map of the m-th target frame and that of the matched reference frame are denoted as $M^t_m = VGG13((I^t_m)')$, $M^r_m = VGG13(H(I^r_{\hat{k}_m}, I^t_m))$, respectively. Note that the target frames and the reference ones, in the setting of this paper, have the same resolution, meaning their feature maps are of the same shape. This fact enables the feature maps to be compared in a position-wise manner as $SM_{m,i,j} = cos\_sim(M^t_{m,i,j}, M^r_{m,i,j})$, where $i \in \{1, 2, ..., H\}$, $j \in \{1, 2, ..., W\}$, and $M^t_{m,i,j}$ and $M^r_{m,i,j}$ are C-dimensional vectors. It is noteworthy that NCC corresponds to cosine similarity when the target pair is two vectors. By following (Kim et al. 2017), one can obtain a similarity map since it was proposed for template matching. Thus, it is converted into a dissimilarity map as $DM_{m,i,j} = 1 - SM_{m,i,j}$.

A multi-scale option is introduced as in (Carvalho et al. 2019) with some modifications. (Carvalho et al. 2019) obtained different-scale maps by resizing an frame. Subsequently, they resized them to the input size and just added them up. This way can be followed, but there are some constraints because of the CNN attribution. Some CNN layers downsample an image. While their processing, they would discard the right-end or bottom-end information of the input image not even considering it due to the filter size or the stride of those layers. VGG13 has five pooling layers, the window size of which is 2×2. The other layers of VGG13 do not affect the output size. For this reason, the resolution of the input should be divisible by $2^5 = 32$ to avoid loss of spatial information. Moreover, the aspect ratio of the proposed dataset is 16:9. Putting these conditions together, two resolution candidates are obtained: 512×288 and 1024×576. This paper extracts features only from the final pooling layer of VGG13. The feature maps from it contain the most abundant peripheral context than those from the preceding layers.

Three weight types are proposed to combine different-scale dissimilarity maps DM^k, $k \in \{1, 2, ..., n_{dm}\}$. n_{dm} denotes the total number of DM. Equation (4) shows how to create a weighted map, $weighted\_DM$.

$$weighted\_DM_{i,j} = \sum_k w^k rDM^k_{i,j}, \tag{4}$$

where rDM^k is the resized DM^k to the input size (W^{org}, H^{org}) with nearest neighbor interpolation. $i \in \{1, 2, ..., W^{org}\}$ and $j \in \{1, 2, ..., H^{org}\}$ are xy-coordinate positions. w^k is the k-th weight. One weight type is MAX as in Equation (5).

$$w^k = \frac{\max(DM^k)}{\sum_{l=1}^{n_{dm}} \max(DM^l)}. \tag{5}$$

A change is more detectable by a suited-scale map than the other scale maps. Thus, using a certain-scale map probably results in higher dissimilarity values for the corresponding-scale change than the other scale maps. Based on this idea, the MAX weight type is designed not to miss changes. Another is EQUAL as in Equation (6).

$$w^k = \frac{1}{n_{dm}}. \tag{6}$$

In EQUAL, all weights have the same value. The third weight type is LARGE as in Equation 7. Assume the

size of DM^1 be the smallest of the maps DM_1, DM_2, ..., $DM_{n_{dm}}$ and set w^1 as the possible maximum weight.

$$w^k = \begin{cases} \dfrac{1}{1 + \sum_{l=2}^{n_{dm}} \max(DM^l)} & \text{if } k = 1, \\ \dfrac{\max(DM^k)}{1 + \sum_{l=2}^{n_{dm}} \max(DM^l)} & \text{otherwise.} \end{cases} \quad (7)$$

With the LARGE weight, the smallest dissimilarity map is assigned with the possible maximum weight. In other words, the weights of the other scale maps would have smaller weights than the case of the other weight types. The smallest map contains the spatially roughest information, meaning it includes less environmental effects such as parallax than the other maps. Thus, the LARGE weight is expected to mitigate environmental effects causing false alarm. Once the weighted map is obtained by Equation (4), a change map can be calculated by binarizing the weight map. The threshold value is discussed in the EXPERIMENTS AND RESULTS section.

EXPERIMENTS AND RESULTS

Dataset

The existing datasets do not contain anomalous changes. Therefore, a new dataset has been created by recording some looming-motion videos with a radio-controlled vehicle. The vehicle ran on a straight corridor at a speed of 0.5 meters per second, and never moved backward. Its trajectories were not identical, and the viewing angle of the vehicle was inconsistent. The proposed dataset includes two sets of a reference and six target videos about 1.5 minutes long each, so the total number of the videos is fourteen. The two sets were captured at different times of the day: day and night. Table 1 shows what kind of changes the target videos contain. Each of the target videos was temporally, spatially aligned to the time-wise corresponding reference video. For TA assessment, each target frame was temporally aligned to a reference one at hand. Subsequently, dissimilarity maps were calculated by comparing each of the aligned frame pairs. To evaluate FC performance, each change was labeled with a bounding box. The proposed dataset is challenging due mainly to parallax or strong illumination change.

Parameter setting

The proposed method has some adjustable parameters. T_c was set to 0.995. For TA, $back$ was fixed to zero since the camera never moved backward in the dataset. Preferable values for *front* were roughly searched for by grid search with a set of values (3, 5, 7, 10). Consequently, this paper chose *front*= 7 for the day targets and *front*= 3 for the night ones. As referred to in the METHOD section, a multi-scale option was employed, and the input images were resized to two scales: 512×288 and 1024×576. For ZNCC, this paper followed (Carvalho et al. 2019) and prepared scales: 20×11, 40×22, 80×45, and 160×90. Besides, another scale 320×180 was also tested for a deeper survey. The window size of ZNCC was set to five.

Table 1: Change Types in the Proposed Dataset

time	data name	included change type
day	fallen	person (fallen)
		bottle
	standing	person (standing)
		umbrella (dropped)
	walking	person (walking)
		umbrella (leaning)
		door
	stacked	two boxes (stacked)
	separate	two boxes (separate)
	bag	bag
night	fallen	person (fallen)
		bottle
		shoe
	standing	person (standing)
		umbrella (leaning)
		shoe
	walking	person (walking)
		umbrella (dropped)
		shoe
	stacked	two boxes (stacked)
	separate	two boxes (separate)
	bag	bag

Experiment

The proposed TA method was compared with (Evangelidis and Bauckhage 2013). To give a quantitative comparison, this paper followed (Diego et al. 2013). They set a ground-truth interval $[l_t, u_t]$ for each index t of a sequence and calculated TA errors as in Equation (8).

$$err(t, \hat{k}_t) = \begin{cases} 0 & \text{if } l_t \le \hat{k}_t \le u_t, \\ \min(|l_t - \hat{k}_t|, |u_t - \hat{k}_t|) & \text{otherwise,} \end{cases} \quad (8)$$

where $\hat{k}_t$ is the t-th index matched by a TA method and $t \in \{1, 2, ..., n_{rc}\}$. The one-by-one ground truth GT_t, which the proposed dataset included, was expanded by one on the negative and positive sides (i.e. $l_t = \max(1,\ GT_t - 1)$, $u_t = \min(n_{rc},\ GT_t + 1)$). Table 2 shows rates of frames with equal or less than each error. TA with VGG16' provided better results than (Evangelidis and Bauckhage 2013) in all the videos. In case of $err = 0$, the error gaps are at least 8.9% (day/separate) and at most 54.5% (day/fallen). Even when comparing the $err = 0$ results by VGG16' and the $err \le 2$ results by (Evangelidis and Bauckhage 2013), most of the former results are better. Comparing the day with the night, both methods rather struggled to align the day sequences. Looking at $err = 0$, the gaps between day/Average and night/Average are 14.9% (VGG16') and 39.6% (Evangelidis and Bauckhage 2013). This is because sunlight through windows formed different shapes on the wall and floor and affected the surrounding brightness, making the day videos of the proposed dataset more challenging. This sunlight worsened (Evangelidis and Bauckhage 2013) more strongly than VGG16' because the former method using local de-

scriptors unfortunately captured the sunlight changing its form. On the other hand, such local changes were invisible for VGG16' thanks to its GAP.

With the TA results by VGG16', the FC performance of the proposed pipeline was evaluated by the area under the receiver-operator curve (AUC). VGG13 and ZNCC with the three weight types were compared as shown in Table 3. This result only exhibits the best combination of scales: [512×288, 1024×576] for VGG and [20×11, 40×22, 80×45] for ZNCC. Table 3 indicates three notable points. Firstly, VGG13 outperformed ZNCC in all of the scenarios. Secondly, both methods performed the worst on day/walking and night/stacked for each time. What deteriorated them is investigated in the following paragraph. Thirdly, the weight type provided just a marginal difference. This is because if just a single pixel in one of the different-scale maps has a high value, it pushes up the weight of that map. Therefore, all weights ended up getting almost the same value.

Table 4 shows AUC scores of the FC methods with and without the proposed TA results. Only the LARGE weight type was shown as it performed the best. At day/walking and night/stacked in Table 4, large gaps can be seen. This suggests the failure of TA led to the terrible FC performance. This suggestion was confirmed by counting detectable pixels, which were non-masked ones in the METHOD section. The inner rate columns in Table 4 show each detectable pixel rate (%). Non-change areas are the outside of bounding boxes, and change areas are the inside. As one can see, the change inner rates for day/walking and night/stacked are obviously low, meaning a large part of the change areas was regarded as unchanging. Therefore, Table 4 proves TA plays a pivotal role in CD.

This paper expanded on how scale sizes affected results. Table 5 shows AUC scores for VGG13 and ZNCC with single scales. Smaller resolutions provided better scores because the AUC was a pixel-wise criterion. That is, a method tuned for larger objects contributes to the score more than smaller ones. Also, this tendency can be found in the best set of scales for ZNCC ([20×11, 40×22, 80×45]). Another notable point is that, comparing the AUC scores in Table 5 with the AUC scores using the TA ground truth in Table 4, the combination of multi scales enhanced the detection ability. In addition to the AUC, this paper looked into the relationship of scales and dissimilarity values for each object. This paper calculated the median of dissimilarity values belonging to each change type or the background and then a ratio of each object median to the background one. If a ratio is less than 1.0, the corresponding change is indistinguishable from the background. The higher it is, the more detectable the change is. Note that the looming motion in the videos significantly varies the size of changes. Thus, Table 6 only shows "person" and "bottle", a large and a relatively small change, in {day, night}/fallen. One can see the tendency of larger resolutions spotting smaller changes and vice versa.

Finally, to discuss the FC performance for each change type, a ratio of each object median to the background one was computed with VGG13 ([512×288, 1024×576]) and ZNCC ([20×11, 40×22, 80×45]) as shown in Table 7. The weight type was fixed to LARGE as in Table 4. Table 7 indicates some characteristics of ZNCC and VGG13. ZNCC shows distinctively strong and weak points. It failed to detect the smallest change, "bottle". Moreover, the value for a relatively small object "umbrella" is significantly smaller than the other changes except for "door". Note that although "shoe" might sound small, it appears close to the vehicle trajectories. Thus, "shoe" looks big in the proposed dataset. On the other hand, VGG13 successfully detected "bottle". Its ratio is actually close to 1.0, but this result seems reasonable because "bottle" is not only small but also unobtrusive in the proposed dataset. For the other changes including "umbrella", VGG13 almost impartially spotted them. This implies that VGG13 does not largely depend on the input scales. This is because its convolutional layers acquire surrounding information.

CONCLUSION

This paper aims to automatically monitor areas for security using a moving camera instead of humans. None of the existing CD datasets was designed for such a purpose. Thus a new dataset for area surveillance has been built with a UGV. Subsequently, this paper has introduced a structured method and devised three components for it: VC, TA, and FC. For FC, three ways to combine different-scale maps have also been proposed. To perform an evaluation, the proposed TA and CD methods were compared with classic methods. Through the experiments, this paper showed the effectiveness of the proposed method in area surveillance using a moving camera.

There are some limitations in the proposed method. First, the proposed CD method cannot detect changes in a target frame if the matched reference frame does not contain the spatially corresponding region. In terms of false positive, there were some times the method falsely detected objects as changes due to difference in viewing angle or position and the sunlight. Second, the FC performance strongly depends on the preceding procedure: TA and SA, as shown in Table 4. Third, if changes appear in a dominant part of an image, TA and SA would provide a poor result. Finally, the reference video has to contain the whole scenes of the target video. This limits a range of applications.

A piece of the future work is to improve the proposed method by overcoming the limitation. It is necessary to research how to make methods robust to environments. Evaluation-wise, this paper performed an evaluation with the pixel-wise AUC. As aforementioned, it tended to give better scores to a method tuned for larger changes. This tendency is not appropriate for surveillance. For this reason, a new frame-level evaluation should be consid-

Table 2: Frame Rates (%) with Equal or Less than Each Error for TA

		VGG16'			georgios		
time	data	$err = 0$	$err \leq 1$	$err \leq 2$	$err = 0$	$err \leq 1$	$err \leq 2$
day	fallen	68.1	80.1	84.3	13.6	20.4	26.7
	standing	52.3	70.5	76.7	8.0	11.9	14.8
	walking	55.7	73.4	83.7	18.2	26.1	34.5
	stacked	78.7	88.1	94.1	38.6	46.0	49.0
	separate	78.8	90.2	90.7	69.9	85.0	89.1
	bag	67.7	80.8	83.8	18.2	23.7	27.8
	Average	66.9	80.5	85.6	27.8	35.5	40.3
night	fallen	91.3	96.9	100.0	81.6	87.2	91.3
	standing	95.3	97.9	98.4	68.6	79.1	84.3
	walking	98.5	100.0	100.0	85.8	92.9	93.9
	stacked	79.7	91.9	97.7	54.7	65.7	70.9
	separate	68.0	74.6	80.1	27.6	40.3	51.9
	bag	97.8	100.0	100.0	86.0	91.6	98.3
	Average	81.8	93.6	96.0	67.4	76.1	81.8

Table 3: AUC Scores with the Proposed TA Component for FC

		VGG13			ZNCC		
time	data	MAX	EQUAL	LARGE	MAX	EQUAL	LARGE
day	fallen	0.870	**0.871**	**0.871**	0.745	0.739	0.738
	standing	**0.842**	**0.842**	**0.842**	0.731	0.730	0.729
	walking	0.686	0.688	**0.689**	0.654	0.661	0.663
	stacked	0.860	0.862	**0.863**	0.708	0.709	0.704
	separate	0.813	0.817	**0.818**	0.765	0.784	0.786
	bag	**0.926**	0.925	0.925	0.812	0.803	0.801
	overall	0.833	0.834	**0.835**	0.736	0.738	0.737
night	fallen	0.913	0.914	**0.914**	0.834	0.837	0.838
	standing	0.915	**0.916**	**0.916**	0.821	0.820	0.820
	walking	0.917	0.921	**0.922**	0.864	0.877	0.879
	stacked	**0.584**	0.583	0.583	0.512	0.513	0.515
	separate	**0.899**	**0.899**	**0.899**	0.837	0.840	0.841
	bag	0.901	0.903	**0.904**	0.808	0.812	0.811
	overall	0.855	**0.856**	**0.856**	0.779	0.783	0.784
Average		0.844	**0.845**	**0.845**	0.758	0.760	0.760

Table 4: Change/Non-Change Inner Rates (%) and AUC Scores with and without the Proposed TA Component

		TA_VGG16'				TA_GROUND_TRUTH			
time	data	non-change inner rate	change inner rate	VGG13	ZNCC	non-change inner rate	change inner rate	VGG13	ZNCC
day	fallen	83.2	97.1	0.871	0.738	86.9	98.6	**0.908**	0.770
	standing	80.2	92.9	0.842	0.729	85.4	90.8	**0.869**	0.754
	walking	79.4	69.8	0.689	0.663	82.0	85.9	**0.831**	0.738
	stacked	86.7	100.0	**0.863**	0.704	88.3	100.0	0.853	0.691
	separate	92.3	86.2	**0.818**	0.786	93.6	81.1	0.798	0.786
	bag	81.4	97.9	0.925	0.801	83.5	100.0	**0.966**	0.862
night	fallen	92.6	98.5	**0.914**	0.838	92.5	97.9	**0.914**	0.842
	standing	87.9	99.6	0.916	0.820	88.1	98.1	**0.929**	0.841
	walking	91.2	95.1	0.922	0.879	90.5	97.3	**0.936**	0.889
	stacked	86.7	64.9	0.583	0.515	88.5	98.1	**0.861**	0.727
	separate	82.5	98.6	0.899	0.841	87.3	99.4	**0.940**	0.869
	bag	93.5	93.4	0.904	0.811	94.2	93.4	**0.911**	0.815
Average				0.845	0.760			**0.893**	0.799

Table 5: AUC Scores for Different Scales with the TA Ground Truth

		VGG13		ZNCC				
time	data	512×288	1024×576	20×11	40×22	80×45	160×90	320×180
day	fallen	0.910	0.880	0.703	0.766	0.764	0.738	0.699
	standing	0.865	0.841	0.706	0.741	0.719	0.683	0.631
	walking	0.848	0.790	0.731	0.725	0.678	0.618	0.571
	stacked	0.837	0.831	0.572	0.689	0.717	0.719	0.697
	separate	0.815	0.757	0.834	0.758	0.677	0.607	0.550
	bag	0.952	0.949	0.762	0.842	0.863	0.824	0.754
	overall	0.871	0.841	0.718	0.753	0.736	0.698	0.650
night	fallen	0.908	0.890	0.807	0.830	0.818	0.768	0.710
	standing	0.919	0.905	0.787	0.835	0.804	0.704	0.609
	walking	0.936	0.895	0.891	0.865	0.819	0.740	0.663
	stacked	0.833	0.854	0.686	0.709	0.726	0.728	0.699
	separate	0.936	0.918	0.847	0.856	0.834	0.796	0.747
	bag	0.909	0.880	0.793	0.815	0.793	0.707	0.599
	overall	0.907	0.890	0.802	0.818	0.799	0.740	0.671
Average		0.889	0.866	0.760	0.786	0.768	0.719	0.661

Table 6: The Ratio of Each Median of Two Objects to the Background One for Each Scale

		VGG13		ZNCC				
time	change	512×288	1024×576	20×11	40×22	80×45	160×90	320×180
day	person	3.39	2.76	1.90	4.36	4.80	3.58	2.54
	bottle	2.78	2.64	1.20	1.09	3.40	2.54	2.14
night	person	4.95	3.13	13.50	20.00	13.67	5.41	2.64
	bottle	0.92	1.64	0.50	0.33	0.33	1.29	1.15

Table 7: The Ratio of Each Object Median to the Background One

	person	bottle	umbrella	door	box	bag	shoe
VGG13	3.23	1.46	2.82	2.80	3.30	3.38	2.91
ZNCC	6.10	0.70	2.70	3.20	6.40	5.50	4.40

ered. Finally, the proposed dataset contains little variation. Thus, it is required to record videos in different seasons or weather. On top of that, other places such as curves should be included.

REFERENCES

Alcantarilla, P. F.; S. Stent; G. Ros; R. Arroyo; and R. Gherardi. 2018. "Street-view change detection with deconvolutional networks", *Journal of Autonomous Robots*, Vol. 42 (May), 1301-1322.

Carvalho, G. H. F. de; L. A. Thomaz; A. F. da Silva; E. A. B. da Silva; and S. L. Netto. 2019. "Anomaly Detection with a Moving Camera using Multiscale Video Analysis", *Journal of Multidimensional Systems and Signal Processing*, Vol. 30, Issue 1 (January), 311-342.

Chu, Wenqing; H. Xue; C. Yao; and D. Cai. 2019. "Sparse Coding Guided Spatiotemporal Feature Learning for Abnormal Event Detection in Large Videos", *IEEE Transactions on Multimedia*, Vol. 21, Issue 1, 246-255.

Deng, J.; W. Dong; R. Socher; L.-J. Li; K. Li; and L. Fei-Fei. 2009. "ImageNet: A large-scale hierarchical image database", *Proceedings of* 2009 *IEEE Conference on Computer Vision and Pattern Recognition*, 20-25.

Diego, F.; J. Serrat; and A. M. López. 2013. "Joint Spatio-Temporal Alignment of Sequences", *Journal of IEEE Transactions on Multimedia*, Vol. 15, No. 6 (October), 1377-1387.

Evangelidis, G. D. and C. Bauckhage. 2013. "Efficient Subframe Video Alignment using Short Descriptors", *Journal of IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 10 (October), 2371-2386.

Hao, Y.; Z.-J. Xu; Y. Liu; J. Wang; and J.-L. Fan. 2019. "Effective Crowd Anomaly Detection through Spatio-temporal Texture Analysis", *Journal of Automation and Computing*, Vol. 16, Issue 1 (February), 27-39.

Kim, J.; J. Kim; S. Choi; M. A. Hasa; and C. Kim. 2017. "Robust Template Matching using Scale-Adaptive Deep Convolutional Features", *Proceedings of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 708-711.

Lin, M.; Q. Chen; and S. Yan. 2014. "Network In Network", *Proceedings of International Conference on Learning Representations*, 10 pages.

Morais R.; V. Le; T. Tran; B. Saha; M. Mansour; and S. Venkatesh. 2019. "Learning Regularity in Skeleton Trajectories for Anomaly Detection in Videos", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), 11996-12004.

Sakurada, K. and T. Okatani. 2015. "Change detection from a street image pair using CNN features and superpixel segmentation", *Proceedings of British Machine Vision Conference*, 12 pages.

Silva, A. F. da; L. A. Thomaz; G. Carvalho; M. T. Nakahata; E. Jardim; J. F. L. de Oliveira; E. A. B. da Silva; S. L. Netto; G. Freitas; and R. R. Costa. 2014. "An Annotated Video Database for Abandoned-Object Detection in a Cluttered Environment", *Proceedings of* 2014 *International Telecommunications Synposium*, 5 pages.

Simonyan, K. and A. Zisserman. 2014. "Very deep convolutional networks for large-scale image recognition", *Preceedings of International Conference on Learning Representations*, 14 pages.

Singh, A.; D. Patil; and S. N. Omkar. 2018. "Eye in the Sky: Real-Time Drone Surveillance System (DSS) for Violent Individuals Identification Using ScatterNet Hybrid Deep Learning Network", *Proceedings of* 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (*CVPRW*), 1710-1718.

PLANNING OF SUSTAINABLE ENERGY SYSTEMS FOR RESIDENTIAL AREAS USING AN OPEN SOURCE OPTIMIZATION TOOL AND OPEN DATA RESSOURCES

Heiko Driever, Ursel Thomßen, Marc Hanfeld
University of Applied Sciences Emden/Leer, Faculty of Business Studies,
Constantia Platz 4, 26723 Emden
E-Mail: Heiko.Driever@HS-Emden-Leer.de

KEYWORDS
Multi Modal Energy System Simulation and Optimization, Open Science, Helmholtz Open Source Framework for INtegrated Energy System Assessment (FINE), Sector Coupling, Time Variance, Open Source, Open Data, Energy System Model, District Optimization.

ABSTRACT

Global warming and CO_2 emission reduction targets mandate a closer look to energy system planning on different levels. In this work we model a typical residential area that will be built and has to be equipped with a cost optimized, decentral energy system with a high degree of energy autarky and integration of renewable sources. We sketch out the optimization problem and show that the optimization can be done using open data and an open source tool, the Helmholtz Framework for Integrated Energy System Assessment tool, FINE, exclusively.
The energy system model chooses from a predetermined set of technologies, takes into account a temporal discretization approach for energy demands as well as for energy production capacities and considers decentralized sector coupling options. As a result, we get a cost-optimized energy system structure as a base for energy system design.

INTRODUCTION

"CO_2 emission reduction and increasing volatile renewable energy production mandate stronger energy sector coupling and the use of energy storage" (Ripp and Steinke 2019) and the investigation of decentral power supply.
In support, we model, optimize and assess the energy system of a residential area including a typical, time-discrete consumption structure as well as a multi modal, energy system. Main optimization targets are Goals Seven (Affordable and Clean Energy) and Eleven (Sustainable Cities and Communities) of the United Nations' Sustainable Development Goals (UN SDG, https://sdgs.un.org/goals) - in addition to the minimization of costs for the applied technologies and the referring commodities.
Along with the trend to open science (Hilpert et al. 2018), we use open source data for demands, applied technologies and commodities in conjunction with the Helmholtz open source *Framework for INtegrated Energy System Assessment* (FINE, Welder et al. 2018b). On top of being available without additional cost, open source data and software ensure that our model is reproducible and can easily be developed further by interested communities. Furthermore, the open source modelling software FINE ensures high quality of code and functions as it will be constantly scrutinized by a top qualified scientific community – especially within the highly reputed Helmholtz Association (Balter 2015).
This kind of optimization tool, and further developments of it, can help with the choice of sectors, technologies and connections to incorporate for any kind of residential or industrial area. Thereby it can be a support to anybody who intends to design, plan or implement an energy network (Lund et al. 2017). Interested parties might include project developers, building contractors, planning offices, local authorities, institutional investors or credit institutions.
Existing studies on sector coupling modelling include for instance considerations of the integration of hydrogen into energy models (Welder et al. 2019), national energy systems (Welder et al. 2018; Welder et al. 2019; Ball et al. 2007) or appraisals of different modelling approaches (Hilpert et al. 2018).
In this work we enhance existing research with a use case, in which we show, that the open source optimization tool FINE is well suited to design and optimize a multi modal energy system for a real-world residential area. And it can be done using open source tools and open data, exclusively.

PROBLEM STATEMENT

The core of our decision problem is a residential area for which we want to configure an energy system that takes into account all of the framework conditions below. That means that from a given portfolio of technologies for energy production and supply, a technology mix should be chosen, that, considering the framework conditions, leads to minimized energy system costs and CO_2 emissions [capital expenditure (CAPEX) and operational expenditure (OPEX)]. Thus, the output of our energy system planning delivers recommendations for a future design of the system.
In rough terms our model follows a call for bids for the design of a residential area on a four-hectare commercial fallow land area in the city center of Brake in northwestern Germany. So, we will also be able to discuss our findings with the mayor of Brake to provide some input concerning the energy system of the final design of the area.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

FINE helps to capture the time-dependence of energy demands, fluctuation of renewable energy production and sector coupling options.

For our use case, we model a typical residential area (composition see Figure 1) including the following framework conditions:

- 65 residential units (houses & apartment blocks)
- Integration of public and commercial infrastructure (Figure 1 and Table 1)
- Decentralized energy supply and a high degree of self-sufficiency (in terms of energy)
- Integration of renewable energy sources and components (technology portfolio see Figure 2)
- Individual traffic with a high share of electro mobility
- Creation of a heat compound system as an isolated solution

All energy demands in this work are stated in yearly numbers.

Objective Function

Target of the optimization process is to minimize the total costs (CAPEX and OPEX) of the energy system (Welder et al. 2018a), to maximize revenues from feeding surplus electricity into the public electricity grid, and to minimize induced CO_2 emissions (covered in the OPEX).

Min → Costs over Lifetime

$$= CAPEX + \sum_{t=1}^{n} OPEX_t - \sum_{t=1}^{n} Revenues_t$$

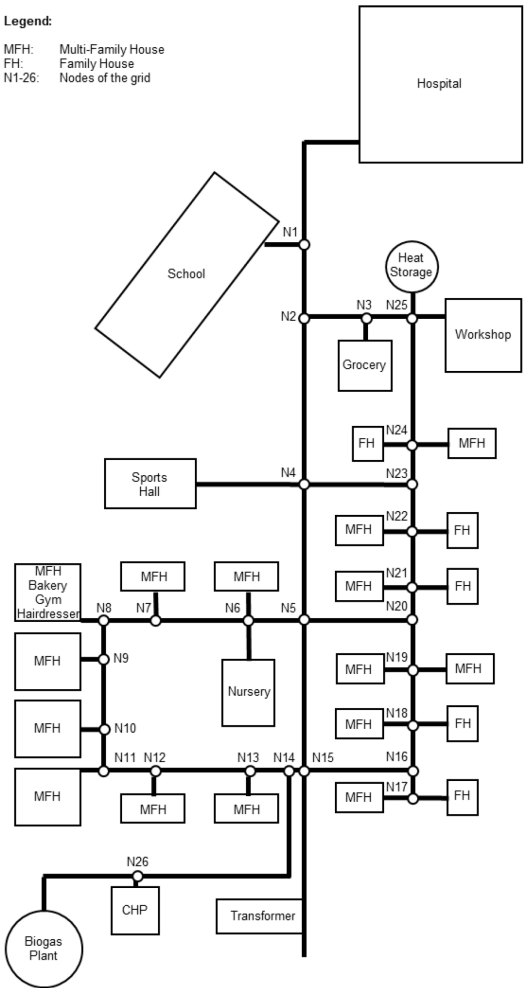


Figure 1: Schema of Buildings, Components and Grid

Table 1: Demands and Capacities of Residential Area Infrastructure per Building Type

Infrastructure classification	Building Type	No. of Buildings	No. of House holds	No. of Residents /Unit	Electricity Demand [kWh]	Heat Demand [kWh]	Roof Surface for Solar Thermal/ Photovoltaic System [m²]	PV Potential Capacity [kWp]	Solar thermal Capacity [kW]	P2H Capacity [kW]	Max. Battery Storage Capacity [kWh]	Max. Heat Storage Capacity [kWh]
Residential	Multi-Family House	3	8	3	114,880	108,000	300	48.0	103.8	255.0	150.0	295.3
	Multi-Family House	1	6	3	29,091	27,000	100	16.0	34.6	85.0	50.0	78.4
	Multi-Family House	4	4	3	76,587	72,000	400	64.0	138.4	340.0	120.0	196.8
	Multi-Family House	7	2	3	69,608	100,800	560	89.6	193.7	595.0	140.0	176.4
	Family House	5	1	4	24,755	45,000	250	40.0	86.5	425.0	50.0	84.5
Public	School	-	-	-	25,000	262,500	500	80.0	173.0	85.0	30.0	-
	Nursery	-	-	-	7,500	78,750	100	16.0	34.6	85.0	-	-
	Sports Hall	-	-	-	26,250	70,000	320	51.2	110.7	85.0	-	-
	Workshop	-	-	-	9,000	33,700	450	72.0	155.7	85.0	-	-
	Hospital	-	-	-	-	-	600	96.0	207.6	-	-	-
Commercial	Grocery Store	-	-	-	6,075	10,125	200	32.0	69.2	85.0	-	-
	Bakery (Int. in MFH 6 HH)	-	-	-	150,000	54,000	-	-	-	-	-	-
	Gym (Int. in MFH 6 HH)	-	-	-	19,440	16,200	-	-	-	-	-	-
	Hairdresser (Int. in MFH 6 HH)	-	-	-	5,265	12,555	-	-	-	-	-	-
Total		**20**	**21**	-	**563,451**	**890,630**	**3,780**	**604.8**	**1307.7**	**2125.0**	**540.0**	**831.4**

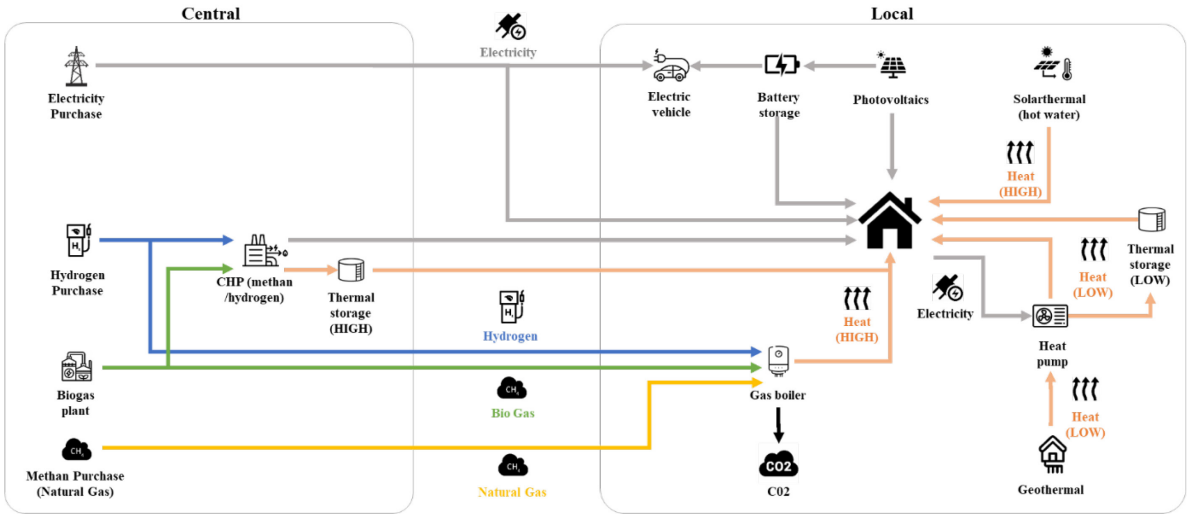


Figure 2: Schematic Diagram of Technology Composition and Interdependencies

In order to evaluate the competitiveness of the resulting energy system we also simulate the referring Levelized Cost of Energy (LCoE) for power and heating of the resulting system.

$$\text{LCOE} = \frac{\text{SUM OF COSTS OVER LIFETIME}}{\text{SUM OF ENERGY PRODUCED OVER LIFETIME}}$$

$$= \frac{\sum_{t=1}^{n} \frac{I_t + M_t + F_t}{(1+i)^t}}{\sum_{t=1}^{n} W_t}$$

I_t: Investment expenditures in the year t
M_t: Operations and maintenance expenditures in the year t
F_t: Fuel expenditures in the year t
W_t: Energy generated in the year t
i: Interest rate
n: Expected lifetime of the energy system (Panos 2017).

Whereas F comprises methane and hydrogen.

Constraints:

- All energy production, except from natural gas, must be from renewable sources
- For purchased electricity from local utility we assume eco-electricity contracts
- Overproduction can be sold
- CO_2 emissions are controlled via penalty costs

Assumptions:

- Commodity prices keep constant
- Power- and heat demands are time-varying
- All data input is deterministic
- CAPEX and OPEX for cables (electricity) and pipelines (heat, H_2, natural gas) are not considered. Our model is confined to identify the necessary performance levels

The planning horizon for our model conforms to the predicted technological life-spans for most of technologies for choice, which is 20 years.
Thus, we can characterize the optimization that we solve, as dynamic decision making (DDM, Hotaling et al. 2017) for a defined period of time with predetermined options (Figure 2) and deterministic planning. The stochastics of exogenous parameters to the system are not considered in our study (see section Solution Approach).

SOLUTION APPROACH

Modelling- and Optimization Tool

For solving our problem, we use the Framework of INtegrated Energy System Assessment (FINE), that was developed by the Helmholtz Energy Computing Initiative (HCEI, Welder et al. 2018a and 2018b). For download, documentation, tutorials and examples please refer to https://github.com/FZJ-IEK3-VSA/FINE.)
FINE is an open source, Python based framework that offers algorithms, component and commodity libraries, variables, and data import features to model, optimize and assess energy systems. It can tackle the challenge of representing the spatial, the temporal as well as the sector-coupling dimensions of multimodal energy systems. Load- and feed-in profiles can be considered in full temporal resolution. The component libraries comprise state of the art options for all parts of the system: Source, conversion-, sink-, storage- as well as transmission component and commodity choices.
We adapt the Urban District Optimization (UDO) workflow, that is recommended by HCEI, to our needs. (https://github.com/FZJ-IEK3-VSA/FINE/blob/master/examples/District_Optimization/Urban_District_Optimization_Workflow.ipyn):
First, based on the structure of the residential area, the energy demands (electricity and heat) have to be

determined. In the second step, the possible technology portfolio with the relevant techno-economic parameters has to be identified. In the third step, all data requirements for determining the load profiles are to be established and the data must be processed. The technology portfolio and the load profiles are input information to the fourth step, the modelling of the energy system model in FINE (Example see Figure 3).

```
esM.add(fn.Source(esM=esM,
                  name='H2Purchase',
                  commodity='hydrogen',
                  hasCapacityVariable=False,
                  operationRateMax=data['H2 Purchase, operationRateMax'],
                  commodityCost=0.285))

esM.add(fn.Source(esM=esM,
                  name='PV',
                  commodity='electricityPV',
                  hasCapacityVariable=True,
                  hasIsBuiltBinaryVariable=True,
                  operationRateMax=data['PV, operationRateMax'],
                  capacityMax=data['PV, capacityMax'],
                  sharedPotentialID ='roofArea',
                  interestRate = 0.04,
                  investIfBuilt=1000,
                  investPerCapacity=1400,
                  opexPerCapacity = 14,
                  bigM = 700))
```

Figure 3: Example of Commodity Sources Definition

In this step the energy system is modeled using the subclasses sources, sinks, conversions, storages and transmissions. After modelling the energy system, it has to be optimized to get a recommendation for the energy system structure of the residential area. For this optimization we use the FINE standard solver Gurobi in the current version 9.1.1.

We were able to define all choices of components and commodities for our residential area model as well as the associated cost- (CAPEX and OPEX), capacity and demand variables in FINE. It is also possible to define a presetting that includes all of the framework conditions mentioned above, including time series for the time-dependence of energy demands and production.

Our model includes the following energy system components (Figure 2):

- Sources: Photovoltaics, solar thermal collectors, biogas plant, electricity purchase, hydrogen purchase, natural gas purchase
- Conversion: Geothermal, gas boilers, combined heat and power plant (CHP), P2H
- Storages: Battery storages, thermal storages (high, low)
- Transmission (commodities): Electricity, hydrogen, natural gas, biogas, heat (high)
- Sinks: Electricity demand, heat demand, electricity sales PV, electricity sales CHP

Collection, Preparation and Processing of Data

The accuracy and results in energy system modelling depend on the availability, selection and preprocessing of input data. For our model we need electricity and heat demands, load profiles, generation capacities as well as techno-economic parameters. Temporal and spatial scales determine level of detail and resolution requirements.

Our residential area consists of 26 buildings, a biogas plant, a CHP as well as a central heat storage in a spatially distributed network with 26 nodes and a transformer (Figure 2).

The investigated annual electricity and heat demands of the residential, public and commercial units as mentioned in Table 1, depend on the number of residents and households per building, the number and usage of electrical vehicles (e-vehicles) and individual user behaviour for hot water demand (Worm and Rathert 2015; Mailach and Oschatz 2016; Stadtwerke Gießen AG, n.d.)

In our model, the temporal resolution is one hour which results in 175,200 time steps for twenty years. The basis for the simulation of the loads and generation profiles is the historical weather data by the German Weather Service (DWD) with mean temperatures for the period of 2010 to 2020 on hourly level for the Bremen region.

The annual demands are distributed temporally as well as according to the load profiles for electricity and heat demand. The load profiles we use are provided by the project DemandRegio (Gotzens et al. 2020; https://github.com/DemandRegioTeam/), as well as by the Open Source Load Profile Generator (Pflugradt 2016; https://www.loadprofilegenerator.de/) for e-vehicles and hot water profiles.

For the electricity demands the following profiles are applied:

Table 2: Applied Load Profiles Buildings

Building Type	Load Profile	
Private Households	H0	Household dynamized
School, Nursery, Workshop	G1	Commerce in General
Sportshall, Gym	G2	Businesses with heavy to predominant consupmtion in the evening hours
Grocery Store	G3	Commerce Continous
Bakery, Hairdresser	G4	Sales Outlet/Barber Store

The profiles for electricity demand consider the seasonality in general and a daily factor in case of profile private households (H0).

The e-vehicle profiles are also generated using the Load Profile Generator and distributed among the residential buildings:

Table 3: Applied Load Profiles E-Vehicles

No.	Load Profiles E-Vehicles		Charging Power [kW]
3	CHS01	Family, 2 children, single family home, 2 Cars	3.5
9	CHR02	Couple, 30-64 years, with work, multi-family house, 30 km commuting distance	11
5	CHR51	Retired, >65 years, multi-family house, 5 km commuting distance	11
4	CHR07	Single, with work, muliti-family house, 30km commuting distance	22

Furthermore, we use the Technical University of Munich (TUM) Sigmoid Function (BDEW/VKU/GEODE-Leitfaden 2018). This function uses the annual heat demand, above mentioned historical mean temperature and the

parameters in its specification 33 (medium heating demand) to calculate the hourly heat demand per building type.
Depending on the building type, the following heat profiles are used: SpaceHeating-EFH for Family Houses; SpaceHeating-MFH for Multi-Family Houses; GKO: Local authorities for School, Nursery and Sports Hall; MK: Metal and automotive for Workshop; BA: Bakery for Bakery; HA: Retail and wholesale for Grocery Store; BD: Other operational service for Gym and Hairdresser.
Hot water profiles are generated using the Load Profile Generator for a 3-person household and a 4-person household.
For the photovoltaics (PV) and solar thermal energy generation profiles the above-mentioned historical weather data with direct and diffuse solar radiation data is used to calculate the global solar radiation on an hourly basis. The maximum installable capacity for PV and solar thermal collectors is determined by the available roof area and has to be split between both. The module efficiency factor for PV is set to 0.16 and the solar constant to 1,000 W/m². For solar thermal the overall equipment effectiveness is set to 0.3 and tilt factor of 1.1. FINE then calculates the hourly electrical (PV) and thermal (solar thermal) yield with the help of global solar radiation, installable capacity and set P-Ratio of 0.85 (PV, Quaschning 2013).
The area that could be used for geothermal collectors, theoretically would be limited by the existing area. For simplification purposes, we assume it as unrestricted so that all necessary heating capacities can be installed.
For the central high heat storage, we allow for a capacity of 35,000 kWh and a lower limit of 2,500 kWh. For the local heat storages, aggregated from low and high heat, maximum capacities are specified in Table 1.
The battery storages capacity for each building can be extracted from Table 1.
Maximum capacities of local Power-to-Heat (P2H) comprise 85 kWh at every residential infrastructure, the School, Sport Hall, Nursery, Workshop and Grocery Store each. A central P2H plant encompasses a maximum of 2,000 kWh and a lower limit of 500 kWh at the central heat storage location. CHP capacities comprise one 500 kWh unit with a lower limit of 100 kWh and one 1,500 kWh unit with a lower limit of 500 kWh at the CHP location as well as two 500 kWh units with a lower limit of 100 kWh, each, at the Hospital location.
We assume the cost structure of the technology portfolio (Tables 4 and 5) based on the following studies: Fattler et al. 2019; Lauinger et al. 2016; Lindberg et al. 2016a; Lindberg et al. 2016b; Mayer et al. 2015; Samweber and Schifflechler. 2017; Stenzel et al. 2019; Sterchle et al. 2016; Streblow and Ansorge 2017; Bundesnetzagentur 2021; Kraft-Wärme-Kopplungsgesetz - KWKG 2020.
Other techno-economic parameters include the following:

Table 4: Other Techno-Economic Parameters

Interest Rate	4%
Household Electricity Price	0.2986 €/kWh
Household Natural Gas Price	0.0615 €/kWh
PV Feed-In Tarif	0.08 €/kWh
CHP Feed-In Tarif	0.073 €/kWh
Purchase Price Biogas	0.12 €/kWh
Purchase Price Hydrogen	0.285 €/kWh

For power and gas from local utility, we assume 0.33 and 0.2 kg/kWh CO_2 emission into the atmosphere. In our model we impede these emissions by applying penalty costs of 1,000€/kg CO_2.
We assume that the distribution grids for electricity, natural gas, biogas and hydrogen are already in place. Therefore, the costs for the distribution grids are not part of the optimization.
Kannengiesser et al. 2019 suggest that a clustering of 20 typical time periods would be the most appropriate trade-off between accuracy and computional load, in their setting. Still, due to restrictions of time and computing capacities, we applied a clustering of 5 typical periods. For that same reason a mixed integer programming gap (MIPGap) tolerance of 0.0005 is set.

Table 5: Technology Portfolio - Cost Structure

Installation	Technologies	$CAPEX_{Cap}$	$CAPEX_{Fix}$	$OPEX_{Cap}$	Efficiency
Local	**PV**	1,400 €/kWp	1,000 €	1.0% of $CAPEX_{Cap}$	-
	Solar Thermal	1,400 €/kWp	1,000 €	1.0% of $CAPEX_{Cap}$	-
	Geothermal Heat Pump	1,700 €/kWth	-	1.3% of $CAPEX_{Cap}$	-
	Condensing Boiler	200 €/kWth	5,600 €	5.0% of $CAPEX_{Cap}$	95.0%
	PH2	350 €/kWth	-	2.0% of $CAPEX_{Cap}$	
	Battery Storage	1,300 €/kWel	2,000 €	-	0.01%/h
	Heat Thermal Storage	55 €/kWth	-	-	0.1%/h
Central	**CHP**	720 €/kWel	-	4.0% of $CAPEX_{Cap}$	45%el/40%th
	Battery Storage	1,000 €/kWel	-	-	0.01%/h
	Heat Thermal Storage	18 €/kWth	-	-	0.1%/h

RESULTS

With regard to the overall effort for the simulation, we can state that the biggest share goes into research and preparation of the required data. The work needed for the modelling depends on the size of the problem, e.g. the number and variation of buildings, the temporal resolution (number of time steps) and the time span that is calculated. As FINE is a Python based software, we consider the coding process for the simulation as manageable. The pure calculation time of one simulation run depends on the size of the problem as well as on the configuration of the available hardware. For a detailed introduction please refer to Welder et al. 2020. All in all, FINE proved well suitable for the task.

As a result of the optimization, FINE recommends the following mix of technologies for installation. The power demand is covered by photovoltaic and CHP. In addition, a central electricity battery storage with a capacity of approx. 26 kWh is suggested. The power system results in total costs of approximately 2 Mio. EUR for 20 years. The heat demand of the locations is covered by CHP capacities. Furthermore, the residential houses will be supplied by local geothermal capacities. In the location 'School' a condensing boiler plus a Power-2-Heat (P2H) capacity is suggested. In addition, thermal storages are planned as a local solution. Biogas and (green) hydrogen are used as fuels for the CHP. Biogas is also used in the condensing boiler. The total costs of the heating system amount to approximately 2.54 Mio. EUR. The aggregated results of the energy system optimization are shown in Table 6. Elements of the technology portfolio that have not been considered by the optimization are not mentioned in the table.

The levelized costs of electricity for our system amount to 0.15 EUR/kWh. The levelized costs of heat amount to 0.12 EUR/kWh. (The LCoE are calculated with an interest rate of zero; costs of distribution are not considered in the model).

Based on data of the Federal Office of Statistics (Statistisches Bundesamt 2021), the average costs of electricity sum up to approximately 0.3 EUR/kWh in 2020; for district heating approximately 0.1 EUR/kWh. That means, in the FINE optimized system about 0.15 EUR/kWh could be spend on electricity cables as distribution capacities.

Considering the results of our study, it seems possible to build and operate a competitive and sustainable energy system with zero CO2-emissions.

Table 6: Optimization Results - Recommended Energy System Structure and Cost Information

PV Generation - Hospital		**Unit**	
PV capacity	86	kW	
PV generation	88,516	kWh/a	
Heat Geothermal	(All residential buildings receive geothermal energy)		
Geothermal capacity	78	kW	
Geothermal generation	390,438	kWh	
Fuels			
Biogas	1,400,012	kWh/a	
Hydrogen	59,349	kWh/a	
CHP	**Electric Capacity [kW]**	**Electricity Generation [kWh]**	**Heat Generation [kWh]**
Central CHP	100	228,591	203,192
Local CHP - Hospital	125	352,326	313,178
Boiler	**Capacity [kW]**	**Heat Generation [kWh]**	
Boiler - School	51	160,013	
P2H	**Capacity [kW]**	**Heat Generation [kWh]**	
Local PH2 - School	2	2,970	
Storage	**Capacity [kWh]**		
Central Battery (Medium)	26		
Local Thermal High Heat	423		
Local Thermal Low Heat	112		
CAPEX/OPEX	**CAPEX [EUR]**	**OPEX [EUR/a]**	
PV	122,033	1,210	
Geothermal	132,291	1,712	
Central CHP	72,219	2,909	
Local CHP - Hospital	89,781	3,616	
Boiler	15,884	514	
P2H	793	16	
Central Battery (Medium)	26,245	-	
Local Thermal High Heat	23,261	-	
Local Thermal Low Heat	6,137	-	
Fuel Costs	**[EUR/a]**		
Biogas	168,001		
Hydrogen	16,914		
Levelized Costs of Heat	**Total Costs [EUR]**	**EUR/kWh**	
CHP	1,614,928		
Geothermal	459,677		
Boiler	430,411		
P2H	10,029		
Storages	29,398		
Total costs	**2,544,442**	**0.1189**	
Levelized Costs of Electricity	**Total Costs [EUR]**	**EUR/kWh**	
CHP	1,825,903		
PV	146,240		
Storages	26,245		
Total costs	**1,998,387**	**0.1502**	
Electricity Demand	**665,408**		

CONCLUSIONS

Our main finding is that FINE is an appropriate tool to model and optimize energy systems based on open access energy data. The consideration of time-dependent demand and supply fluctuation as well as sector coupling options are features that will be needed to handle future energy system management. Based on the problem statement of creating a sustainable energy system for a residential area, we were able to show that specifically in terms of LCoE our result appears to be competitive with current (conventional) energy systems.
In future research and applications, the results of the study can be transferred to similar energy system planning problems like other residential areas, industrial compounds or village structures.

ACKNOWLEDGEMENTS

The authors would like to thank Leon de Vries, Falko Orzessek and Flemming Stötzer, student assistants at the University of Applied Sciences Emden/Leer, as well as André Wessels, Kelly Kummerow and Hannah Stalleicken, research assistants at the University of Applied Sciences Emden/Leer, for discussions, data research and technical support.
Furthermore, our acknowledgements go to the Helmholtz Energy Computing Initiative for providing FINE with all its capacities and features that we were able to use open source – as well as to all authors who provide open tutorials, examples and documentation on GitHub.

REFERENCES

Ball, M.; M. Wietschel; O. Rentz. 2007. „Integration of a hydrogen economy into the German energy system: an optimising modelling approach." *International Journal of Hydrogen Energy 2007*. Volume 32, Issues 10-11, 2007, 1355-1368.

Balter, Ben. 2015. 6 motivations for consuming or publishing open source software. Retrieved from https://opensource.com/life/15/12/why-open-source. (Accessed 2020, November 17).

BDEW/VKU/GEODE-Leitfaden. 2018. *Abwicklung von Standardlastprofilen Gas*. BDEW Bundesverband der Energie- und Wasserwirtschaft e. V., Verband kommunaler Unternehmen e. V. (VKU) sowie von GEODE – Groupement Européen des entreprises et Organismes de Distribution d'Énergie, EWIV. Berlin.

Bundesnetzagentur 2017, Haushaltskundenpreis Strom und Gas/Entwicklungen Beschaffungskosten, Netzentgelte und EEG-Umlage (Stichtag 1. April 2017), 2017. Available online: https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Energie/Unternehmen_Institutionen/DatenaustauschUndMonitoring/Monitoring/Monitoring2017_Kapitel/E_Einzelhandel2017.pdf?__blob=publicationFile&v=1 (accessed on 1 July 2019).

Bundesnetzagentur 2021, EEG-Registerdaten und – Fördersätze. Retrieved from https://www.bundesnetzagentur.de/DE/Sachgebiete/ElektrizitaetundGas/Unternehmen_Institutionen/ErneuerbareEnergien/ZahlenDatenInformationen/EEG_Registerdaten/EEG_RegDaten_FoerdSaetze.html. (2021, January 31)

Caglayan, D. G.; H. U. Heinrichs; J. Linssen; M. Robinius; Detlef Stolten. 2019. "Impact of different weather years on the design of hydrogen supply pathways for transport needs." *International Journal of Hydrogen Energy*, Volume 44, Issue 47, 4 October 2019. 25442-25456.

Fattler, S.; J. Conrad, A. Regett. 2019. *Dynamis Datenanhang. Dynamische und intersektorale Maßnahmenbewertung zur kosteneffizienten Dekarbonisierung des Energiesystems*. Forschungsstelle für Energiewirtschaft e.V. (FfE). Munich.

Fluri, V. 2018. *Wirtschaftlichkeit von zukunftsfähigen Geschäftsmodellen dezentraler Stromspeicher*. Dissertationsschrift. Fraunhofer ISE. Fraunhofer Verlag, Freiburg/BRSG.

Gotzens, F.; Gillessen, B.; Burges, S. Hennings, W.; Müller-Kirchenbauer, J.; Seim, S.; Verwiebe, P.; T. Schmid; F. Jetter: T. Limmer. 2020. DemandRegio. Harmonisierung und Entwicklung von Verfahren zur regionalen und zeitlichen Auflösung von Energienachfragen. Abschlussbericht. Berlin, Jülich, München.

Hilpert, S.; C. Kaldemeyer; U. Krien; S. Günther; C. Wingenbach; G. Plessmann. 2018. "The Open Energy Modelling Framework (oemof) - A new approach to facilitate open science in energy system modelling." *Energy Strategy Reviews 2018*, Volume 22, 16-25.

Hotaling, J.; P. Fakhari; J. R. Busemeyer. 2015. "Dynamic Decision Making." In *International Encyclopedia of the Social and Behavioral Sciences 2015, J.D. Wright* (Eds.). Elsevier, Oxford, 709-714.

https://github.com/FZJ-IEK3-VSA/FINE/blob/master/examples/District_Optimization/Urban_District_Optimization_Workflow.ipynb. (Accessed on 2020, November 17).

Intergovernmental Panel on Climate Change (IPCC). 2018. *Summary* for Policymakers. In: *Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.* World Meteorological Organization, Geneva, Switzerland.

Kannengiesser, T.; M. Hoffmann, L. Kotzur; P. Stenzel; P. Markewitz; F. Schuetz; K. Peters; S. Nykamp; D. Stolten; M. Robinius. 2019. "Reducing Computational Load for Mixed Integer Linear Programming: An Example for a District and an Island Energy System." Energies 2019, 12, 2825.

Kraft-Wärme-Kopplungsgesetz (KWKG) vom 21. Dezember 2015 (BGBl. I S. 2498), last change by Artikel 17 des Gesetzes vom 21. Dezember 2020 (BGBl. I S. 3138)

Lauinger, D.; P. Caliandro; J. Van herle; D. Kuhn. 2016. A linear programming approach to the optimization of residential energy systems. *J. Energy Storage 2016*, 7, 24–37.

Lindberg, K.B.; G. Doorman; D. Fischer; M. Korpås; A. Ånestad; I. Sartori. 2016a. Methodology for optimal energy system design of Zero Energy Buildings using mixed-integer linear programming. *Energy Build. 2016*, 127, 194-205.

Lindberg, K.B.; G. Doorman; D. Fischer; M. Korpås; A. Ånestad; I. Sartori. 2016b. Cost-optimal energy system design in Zero Energy Buildings with resulting grid impact: A case study of a German multi-family house. *Energy Build 2016*, 127, 830–845.

Lund, H.; F. Arler; P. A. Østergaard; F. Hvelplund; D. Connolly; B. V. Mathiesen and P. Karnøe. 2017. "Simulation versus Optimisation: Theoretical Positions in Energy System Modelling." *Energies* 10, no. 7, 840.

Mailach, B.; B. Oschatz. 2016. *BDEW-Heizkostenvergleich Neubau 2016*. ITG Institut für Technischen Gebäudeausrüstung Dresden Forschung und Anwendung GmbH, Dresden.

Mayer, J. N.; S. Philipps; N. S. Hussein; T. Schlegl; C. Senkpiel. 2015. *Current and Future Cost of Photovoltaics. Long-term Scenarios for Market Development, System Prices and LCOE of Utility-Scale PV Systems*. Study on behalf of Agora Energiewende. Fraunhofer ISE.

Panos, Konstantin. 1980. *Praxisbuch Energiewirtschaft*. Springer Vierweg, Berlin.

Pfenninger, S.; L. Hirth.; I. Schlecht; E. Schmid; F. Wiese; T. Brown; C. Wingenbach. 2018. „Opening the black box of energy modelling: Strategies and lessons learned." *Energy Strategy Reviews 2018*, Volume 19, 63-71.

Pflugradt, N.D. 2016. Modellierung von Wasser und Energieverbräuchen in Haushalten. Dissertationsschrift. Technische Universität Chemnitz.

Quaschning, V. 2013. *Erneuerbare Energien und Klimaschutz: Hintergründe - Techniken und Planung - Ökonomie und Ökologie - Energiewende*. Hanser, München.

Ripp, C. and F. Steinke. 2019. „Modeling Time-dependent CO2 Intensities in Multi-modal Energy Systems with Storage." https://arxiv.org/abs/1806.04003

Samweber, F.; C. Schifflechler. 2017. Kostenanalyse Wärmespeicher bis 10.000 l Speichergröße. Retrieved from https://www.ffe.de/publikationen/veroeffentlichungen/659-kostenanalyse-waermespeicher-bis-10-000-l-speichergroesse. (2017, January 10).

Stadtwerke Gießen AG. No Date. Informationen für Sporthallen, Sportplätze und Co., Stadtwerke Gießen AG. Retrieved from https://gc-giessen.stadtwerke-ssl.de/gcGips/static/Mandanten/Giessen/SWG-Broschuere_Sportplaetze.pdf (Accessed on 2020, November 17).

Statistisches Bundesamt (Destatis). 2021. Daten zur Energiepreisentwicklung – Lange Reihen von Januar 2005 bis Dezember 2020. 29. January 2021. Retrieved from https://www.destatis.de/DE/Themen/Wirtschaft/Preise/Publikationen/Energiepreise/energiepreisentwicklung-pdf-5619001.pdf?__blob=publicationFile (Accessed on 2021, February 5).

Stenzel, P.; J. Linssen; M. Robinius; D. Stolten; V. Gottke; H. Teschner; A. Velten; F. Schäfer. 2019. Energiespeicher. *BWK: das Energie-Fachmagazin 2019*. 71. 33-48.

Sterchele, P.; D. Kalz; A. Palzer. 2016. Technisch-ökonomische Analyse von Maßnahmen und Potentialen zur energetischen Sanierung im Wohngebäudesektor heute und für das Jahr 2050. *Bauphysik 2016*, 38, 193–211.

Streblow, R.; K. Ansorge. 2017. Genetischer Algorithmus zur kombinatorischen Optimierung von Gebäudehülle und Anlagentechnik. *Gebäude-Energiewende 2017*, Arbeitspapier 7, Berlin, Germany.

Welder, L.; T. Groß; J. Linssen; Jochen; M. Robiniu; D. Stolten. 2020. "An Introduction to FINE Part I – Installing Software and Simple Model Runs." Retrieved from https://raw.githubusercontent.com/FZJ-IEK3-VSA/FINE/master/examples/Tutorial/FINE_Tutorial_Part1.pdf (2020, November 17).

Welder, L.; J. Linssen; M. Robinius; D. Stolten. 2018a. "FINE—Framework for Integrated Energy System Assessment." Retrieved from https://github.com/FZJ-IEK3-VSA/FINE (Accessed on 2019, 1 July).

Welder L; D. S. Ryberg; L. Kotzur; T. Grube; M. Robinius; D. Stolten. 2018b. „Spatio-temporal optimization of a future energy system for power-to-hydrogen applications in Germany." *Energy 2018,* 158, 1130–1149.

Welder, L.; P. Stenzel; N. Ebersbach; P. Markewitz, M; Robinius; D. Stolten. 2019. "Design and evaluation of hydrogen electricity reconversion pathways in national energy systems using spatially and temporally resolved energy system optimization*." In International Journal of Hydrogen Energy 2019*, Volume 44, Issue 19, 12 April 2019. 9594-9607.

Worm; Rathert. 2015. Bekanntmachung der Regeln für Energieverbrauchswerte und der Vergleichswerte im Nichtwohngebäudebestand, Bundesanzeiger, BAnz AT 21.05.2015 B3, Berlin.

AUTHOR BIOGRAPHIES

HEIKO DRIEVER is a research assistant in the department of Business Studies at the University of Applied Sciences Emden/Leer. He earned his diploma in business administration at the University of Applied Sciences Emden/Leer.

URSEL THOMSSEN is a research assistant in the department of Business Studies at the University of Applied Sciences Emden/Leer. She earned her diploma degree in area studies China combined with business administration at the University of Cologne.

MARC HANFELD is professor for Energy Management in the department of Business Studies at the University of Applied Sciences Emden/Leer.

CAPACITY LOSS ESTIMATION FOR LI-ION BATTERIES BASED ON A SEMI-EMPIRICAL MODEL

*Mohammed Rabah, Eero Immonen, and Sajad Shahsavari
Computational Engineering and Analysis (COMEA)
Turku University of Applied Sciences
20520 Turku, Finland
Email: *mohamed.rabah@turkuamk.fi

Mohammad-Hashem Haghbayan
Department of Future Technologies
University of Turku (UTU)
20500 Turku, Finland

Kirill Murashko
Department of Environmental and Biologic Science
University of Eastern Finland
70211 Kuopio, Finland

Paula Immonen
Laboratory of Electrical Engineering
LUT University
53850 Lappeenranta, Finland

Keywords

SEM (Semi-Empirical Model), LIBs (Li-ion batteries), C_{loss} (Capacity Loss), Cycling Aging, Calendar Aging.

ABSTRACT

Understanding battery capacity degradation is instrumental for designing modern electric vehicles. In this paper, a Semi-Empirical Model for predicting the Capacity Loss of Lithium-ion batteries during Cycling and Calendar Aging is developed. In order to predict the Capacity Loss with a high accuracy, battery operation data from different test conditions and different Lithium-ion batteries chemistries were obtained from literature for parameter optimization (fitting). The obtained models were then compared to experimental data for validation. Our results show that the average error between the estimated Capacity Loss and measured Capacity Loss is less than 1.5% during Cycling Aging, and less than 2% during Calendar Aging. An electric mining dumper, with simulated duty cycle data, is considered as an application example.

INTRODUCTION

The transport sector is one of the largest global emitters of carbon dioxide (CO_2), accounting for about 22% of the total emission (Kluschke et al., 2019). Electric cars have proven to be an efficient way of reducing these emissions in *passenger transport*, but electrification of *heavy-duty vehicles* (e.g. trucks, forest harvesters and mining dumpers) is more challenging, mainly due to limitations in battery technology. Among others, Liimatainen et al. (2019) concluded that battery electric trucks have not been a viable option to replace traditional diesel-powered ones because of the high energy requirements and low energy density of batteries. Furthermore, the absence of charging facilities in off-road conditions may render electrification of forest harvesters impractical. On the other hand, heavy-duty vehicles are typically tailored for a specific application niche, and their production batches are much smaller than those for passenger cars, which means that application-specific design optimization is both necessary and can also have a significant effect on the vehicle performance — and thus on business profitability. At the heart of this design optimization is an understanding of the performance of lithium-ion (Li-ion) batteries.

One major difference between internal combustion engine vehicles (ICEV) and battery electric vehicles (BEV) is that the energy system in the latter degrades during use. While the performance of a diesel engine remains largely unaffected by repeated refuels and use, this is not the case for Li-ion batteries (LIBs): The capacity of LIBs decreases in both repeated cycling and storage. Moreover, the LIB degradation process depends on the battery chemistry and the way (or path) of usage. Dubarry and Devie (2018) concluded that the LIB cell temperature history had the strongest impact on degradation followed by the C-rate (i.e. charge/discharge current) and the state of charge (SoC). Also, they found that LIBs lose capacity faster at low SoCs during calendar aging and under small SoC swings while under cycling.

It is obvious from the above discussion that design optimization for heavy-duty battery electric vehicles (HDBEV) must address battery degradation. The upshot is that, in contrast to passenger cars, since a HDBEV is designed for a specific application, the typical usage conditions — including temperature, SoC and C-rate — can often be estimated with more accuracy during design. Battery system design optimization for HDBEVs thus requires parametric mathematical models of battery aging, estimated from real-world cycling and storage tests. The purpose of this article is to address this concern.

In this article, a Semi-Empirical Model (SEM) is proposed for estimating the capacity loss (C_{loss}) for different LIB chemistries during cycling and calendar aging. The model is developed based on the effect of four different parameters, namely temperature, time, depth of discharge, and C-rate current. The model is able to estimate the C_{loss} of LIBs with a high accuracy and low computation complexity compared to the other models. This model can be used for optimizing LIB systems for different chemistries throughout their lifetime

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

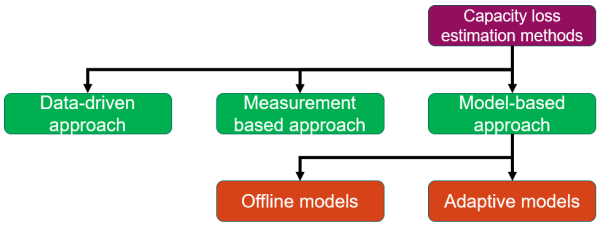


Figure 1: Capacity Loss Estimation methods.

in realistic operation conditions. The contribution of this article is elaborated in the next section.

RELATED WORK

LIB degradation is usually measured by C_{loss}. It reflects the ability of the LIB to store and to supply energy relative to its initial conditions, considering the energy and power requirements of the application (Berecibar et al., 2016). Many different methods for the C_{loss} estimation were presented in recent years and in general they can be divided into three approaches, which are shown in Figure 1.

The measurement based methods by themselves cannot be used for the estimation of the C_{loss} during design of the battery system as these methods requires the analysis of the measured data during operation of the LIB.

The data-driven approach includes methods, where the previously collected information about operation of the LIB is used to find the dependency between the collected information and C_{loss}. Different algorithms such as Support Vector Regression algorithm (Liu et al., 2020), Fuzzy logic (Yang et al., 2020) and Neural Networks (Naha et al., 2020) were investigated for C_{loss} estimation. These methods allows to estimate the C_{loss} with good accuracy in case of the availability of a sufficient amount of previously collected data. However, these methods are computationally intensive and are sensitive to the size and quality of the experimental dataset applied during the training, which are not always available.

The model-based approach (Cacciato et al., 2016) may yield better results for the C_{loss} estimation in case of the low amount of the experimental data. It can be done by applying a model, which should describe the processes occurring in the LIBs. Such models can be used offline and they directly provide the required information on the C_{loss} — or the required information can be obtained during comparison of the calculated and measured data in real time. If a very high accuracy for the C_{loss} estimate is needed, the second type of the battery model, so called adaptive model (Cen & Kubiak, 2020), is often recommended in the literature. The adjustment of the adaptive model parameters and C_{loss} estimation after comparison with measured data can be done by using such algorithms as Kalman Filter, Particle filter, Sliding mode observer etc. As it was reported by (Andre et al., 2013), the use of the adaptive models allows to estimate C_{loss} and State of Charge (SoC) simultaneously with estimation error under 1%.

Despite the high accuracy of the C_{loss} estimation, a high computational complexity may limit the applicability of the adaptive models during battery system design where high accuracy is not always necessary. In this case, offline models such as SEM (Singh et al., 2019) may be more useful as they may estimate the C_{loss} with acceptable uncertainty in case of the lack of the experimental data and they have low computation complexity. In the SEM approach, one attempts to identify a (simple) parametric function that describes the capacity reduction, through parametric optimization.

The applicability of the SEM approach for the C_{loss} estimation during storage (Grolleau et al., 2014) and cycling (Bocca et al., 2015) were widely shown for different LIBs. However, the presented models were usually verified for the same LIBs, from which the SEMs were created and the use of the presented algorithms for the creation of the SEMs for other type of the LIBs is not well discussed. Therefore, the research work described in the present article focused on the analysis of the applicability of the commonly used approach for the creation of the SEMs of different LIBs at different operation conditions.

SEM SPECIFICATIONS

In this article the most commonly used models, which can estimate the capacity loss C_{loss} of the LIBs during Calendar Aging C_{loss}^{cal} and Cycling Aging C_{loss}^{cyc} are analysed. These semi-empirical models were previously used for the modeling of the C_{loss} in LFP cells (Wang et al., 2011), NMC cells (Schmalstieg, Käbitz, Ecker, & Sauer, 2014), NCA cells (Petit, Prada, & Sauvant-Moynot, 2016) and they are briefly described below.

Calendar Aging

The two main factors that affects the Calendar Aging are the T and SoC. The general equation for the Calendar Aging estimation can be presented as:

$$C_{loss}^{cal} = B(SoC) \cdot e^{-\frac{E}{R(T-T_{ref})}} \cdot t^z, \qquad (1)$$

where B is the pre-exponential factor that depends on SoC, T is the temperature expressed in K, T_{ref} is the reference temperature also expressed in K and is equal to 298.15, R is the gas constant, E is the activation energy of a reaction, expressed in J/mol, t is the time in days, and z is a constant. The pre-exponential factor B can be presented as:

$$B = a_1 \cdot SoC + a_2 \qquad (2)$$

Where a_1 and a_2 are fitting constants. Equation 1 can be used to estimate the C_{loss} of the LIB during long period storage.

Cycling Aging

For Cycling Aging, the C_{loss} is mainly affected by current I, T and number of cycles N. Furthermore, other parameters do have a margin effect depending on the temperature of the LIB, e.g. depth of discharge DoD,

and the rated capacity. The general equation used to estimate the C_{loss}^{cyc} is as follow:

$$C_{loss}^{cyc} = B_{cyc}(I) \cdot e^{-\frac{E+\alpha\cdot|I|}{R(T-T_{ref})}} \cdot A_h^{z_{cyc}} \quad (3)$$

Where B_{cyc} is a pre-exponential factor which depends on cycling current I, α and z_{cyc} are the fitting coefficients, and A_h is the full used capacity that can be obtained using the following equation:

$$A_h = FCE \cdot C_r = N \cdot DoD \cdot C_r \quad (4)$$

Where FCE is the full cycle equivalent, C_r is the rated capacity.

MODEL IDENTIFICATION

The process of model identification is divided into two parts; (a) Data Selection and Fitting, and (b) Model Validation. These two parts are illustrated in Figure 2.

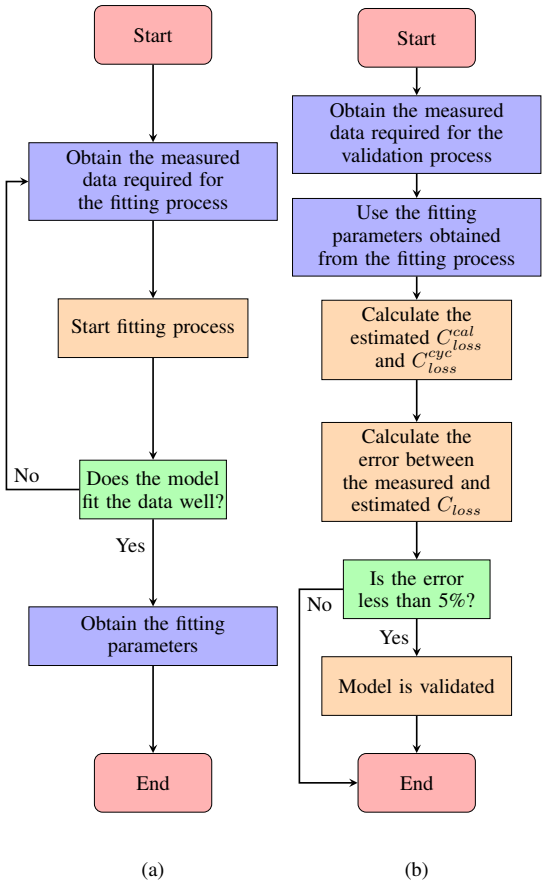


Figure 2: SEM flowchart process; (a) Data Selection and Fitting process, (b) Model Validation process.

Data Selection and Fitting

Figure 2 (a) shows the Flow Chart of the data selection and fitting process. As shown in this Figure, measured data from different references, specifically the C_{loss}^{cal}, C_{loss}^{cyc}, FCE and $Time$ are obtained. Afterwards the fitting process is applied to estimate the value of the fitting parameters. The algorithm used for obtaining the fitting parameters is shown in Algorithm 1. In this algorithm, $Mes\_C_{loss}^{cyc}$ and $Mes\_C_{loss}^{cal}$ are the measured C_{loss} data needed for fitting the model, while $Est\_C_{loss}^{cyc}$ and $Est\_C_{loss}^{cal}$ are the estimated C_{loss} output. In this process, the `fitting` function is using the `fminsearch` from MATLAB which uses a simplex search method (Lagarias et al., 1998) to obtain the estimated data output. In order to decrease the overall error between the measured and the estimated output, the fitting function is proceed in different scenarios, e.g. during a fixed temperature, fixed $DoD = 1 - SoC$, etc., for 1000 iteration in each. For this, several fitting parameters are calculated according to each situation. Once the estimated output reaches the local minimum, the average error is calculated between the estimated output and the measured data. In this case, it's assumed that 5% is when the model does fit the data well. Afterwards, function `fitpar` which uses the `polyfit` from MATLAB is used to generate the fitting parameters based on least square regression. Once these values are calculated, the model validation process is started as shown in Figure 2 (b).

Algorithm 1 Data Selection and Fitting process

Input: $Mes\_C_{loss}^{cyc}$, $Mes\_C_{loss}^{cal}$
Output: $Est\_C_{loss}^{cyc}$, $Est\_C_{loss}^{cal}$,R,a_1,a_2,z,α,z_{cyc},B_{cyc}

Body:

measured $\leftarrow \{\mathbf{Mes\_C_{loss}^{cyc}}, \mathbf{Mes\_C_{loss}^{cal}}\}$
estimated $\leftarrow \{\mathbf{Est\_C_{loss}^{cyc}}, \mathbf{Est\_C_{loss}^{cal}}\}$
for i $\leftarrow 1$ **to** 1000 **do**
  estimated $\leftarrow fitting(\mathbf{measured})$
  if estimated $==$ **localminimum then**
    Stop
  end
end
error $\leftarrow avg.(\frac{\mathbf{estimated-measured}}{\mathbf{measured}})$
if error < 0.05 **then**
  $\{\mathbf{E}, \mathbf{R}, \mathbf{a_1}, \mathbf{a_2}, \mathbf{z}, \alpha, \mathbf{z_{cyc}}, \mathbf{B_{cyc}}\} \leftarrow fitpar(\mathbf{estimated})$
end

Algorithm 2 Model Validation process

Input: $NewMeasuredData$
Output: C_{loss}^{cyc}, C_{loss}^{cal}
Parameters: R,a_1,a_2,z,α,z_{cyc},B_{cyc},T_{ref}

Body:

$\{\mathbf{SoC}, \mathbf{T}, \mathbf{I}, \mathbf{t}, \mathbf{N}, \mathbf{period}, \mathbf{C_r}\} \leftarrow \mathbf{NewMeasuredData}$
`// Calculate` C_{loss} `for Cycling Aging`
$\mathbf{cyc.DoD}, \mathbf{cyc.SoC}, \mathbf{ch1}, \mathbf{ch2} \leftarrow$ rainflow$(\mathbf{SoC})$;
$\mathbf{cyc.I} \leftarrow avg.\,\mathbf{I}\, from\, \mathbf{ch1}\, \mathbf{to}\, \mathbf{ch2}$;
$\mathbf{cyc.T} \leftarrow avg.\,\mathbf{T}\, from\, \mathbf{ch1}\, \mathbf{to}\, \mathbf{ch2}$;
$\mathbf{C_{loss}^{cyc}} \leftarrow B_{cyc} \cdot e^{-\frac{E+\alpha\cdot|cyc.I|}{R(cyc.T-T_{ref})}} \cdot (NC_r(cyc.DoD))^{z_{cyc}}$
`// Calculate` C_{loss} `for Calendar Aging`
if I $== 0$ **then**
  $\mathbf{cal.SoC} = avg.(\mathbf{SoC})$;
  $\mathbf{cal.T} = avg.(\mathbf{T})$;
  $\mathbf{cal.t} = \mathbf{period} * 24 * 3600$;
  $\mathbf{B(SoC)} = \mathbf{a_1} \cdot \mathbf{cal.SoC} + \mathbf{a_2}$;
  $\mathbf{C_{loss}^{cal}} \leftarrow B(SoC) \cdot e^{-\frac{E}{R(cal.T-T_{ref})}} \cdot cal.t^z$
end

Model Validation

In this process, the estimated fitting parameters values are input into Equations (1)-(4), to calculate the C_{loss} during Cycling/Calendar aging, and the output

is compared to a known reference, where the error is calculated between the known C_{loss} and the estimated one. If the error is low (here defined as below 5%), this shows that the model is validated and can be used to estimate the C_{loss} of an experimental LIB.

The method used for calculating the C_{loss} is shown in Algorithm 2. The information about state of charge SoC, temperature T, current I, rated capacity C_r and time between operation cycles is necessary for the C_{loss} calculation t. The loss of the capacity during cycling is calculated from information about N, cycle start time $ch1$ and end time $ch2$, average values of $cyc.DoD$ and $cyc.SoC$ that are calculated from SoC curve by using `Rainflow` algorithm in MATLAB (ASTM E1049-85, 2005). Afterwards, average values of temperature $cyc.T$ and current $cyc.I$. are calculated during a cycle to be used in Equation 3. The calendar C_{loss} is calculated by considering the average state of charge $cal.SoC$ and average temperature $cal.T$ of each long enough period of time when LIBs are not used where there is no current usage during this period.

RESULTS

To test the feasibility of the proposed model, several LIB chemistries should be evaluated. In this work, two different chemistries of LIBs have been chosen; Lithium Iron Phosphate (LFP) and Lithium-Titanate Oxide (LTO). These chemistries are among the primary candidates for modern HDBEV systems.

Lithium-titanate battery (LTO)

Data Selection and Fitting

The measured data that is used for the fitting process is acquired from (Dubarry & Devie, 2018). In his work, he studied the effect of temperature 'T', SoC swing range '$\triangle SoC$', and C-rate 'C' on the battery cells to measure its C_{loss} during 1400 ($1C$ rate) to 4200 ($3C$ rate) full cycle equivalent as demonstrated in Figure 3(a). Likewise, he studied the effect of 'T' and SoC through 61 weeks of Calendar Aging as illustrated in Figure 3(b).

Once the measured data from Figure 6 is extracted, the fitting process is started. The fitting function for the Cycling Aging is proceed in four different scenarios: during a fixed $DoD(40\%)$, fixed temperature (25°C), and fixed C-rate. Figure 4 shows the estimated C_{loss} (dashed line) compared to the measured one (solid line). The average error between the estimated C_{loss} and measured C_{loss} is found to be 0.63% at 50% of FCE and 0.54 at 100% of FCE, except for one point that shows an error of 0.72% during the 4200 FCE that can be found in Figure 4(d) (45/0.7/3).

In the Calendar Aging fitting process, the model has an average error of 0.8% between the estimated C_{loss} (dashed line) and the measured one (solid line), except for the condition T = (55°C), SoC = 5%, as the model does have an average error of 1.4% during this condition.

The results from the literature shows that during Calendar Aging, the LTO tends to degrade faster while the SoC is low compared to higher SoC, in addition to the effect of T. For Cycling Aging, the increase in T and C-rate has significant effect on the LTO chemistry, and the degradation rate is faster when smaller SoC swings $\triangle SoC$ is applied.

Model Validation

In order to test the performance of the proposed model, it needs to be validated and compared to another known measured C_{loss} during cycling and Calendar Aging. For the Cycling Aging, the input data needed for the C_{loss} algorithm is extracted from (Baure & Dubarry, 2020), and the output is compared to the C_{loss} from the same reference. Table 1 and Table 2 first columns show a summary of the extracted data that is required for the Cycling Aging C_{loss}. Furthermore, both tables show the measured C_{loss} in 25/35°C, estimated C_{loss} and the error during 2500 equivalent cycles.

As shown in Table 1, the model does have an average error of 0.46% during Cycling Aging in 25°C, while it does have an average error of 1.39% in 35°C as shown in Table 2.

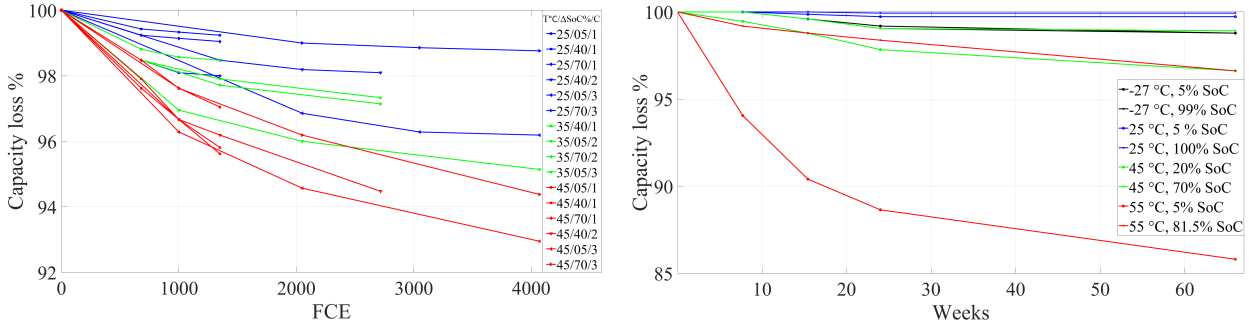
Table 1: Data validation during Cycling Aging measured in 25°C

Data cycling measured in 25 °C	Measured Capacity Loss %	Estimated Capacity Loss %	Error %
Median SoC = 15%, △SoC = 5%	0.37	0.69	0.32
Median SoC = 50%, △SoC = 5%	0.33	0.69	0.36
Median SoC = 85%, △SoC = 5%	0.33	0.69	0.36
Median SoC = 15%, △SoC = 45%	0.42	1.07	0.65
Median SoC = 50%, △SoC = 45%	0.40	1.07	0.67
Median SoC = 85%, △SoC = 45%	0.48	1.07	0.59
Median SoC = 50%, △SoC = 75%	0.41	0.71	0.30

Table 2: Data validation during Cycling Aging measured in 35°C

Data cycling measured in 35 °C	Measured Capacity Loss %	Estimated Capacity Loss %	Error %
Median SoC = 15%, △SoC = 5%	0.54	1.60	1.06
Median SoC = 50%, △SoC = 5%	0.58	1.60	1.02
Median SoC = 85%, △SoC = 5%	0.54	1.60	1.06
Median SoC = 15%, △SoC = 45%	0.62	2.48	1.86
Median SoC = 50%, △SoC = 45%	0.59	2.48	1.89
Median SoC = 85%, △SoC = 45%	0.70	2.48	1.78
Median SoC = 50%, △SoC = 75%	0.58	1.62	1.04

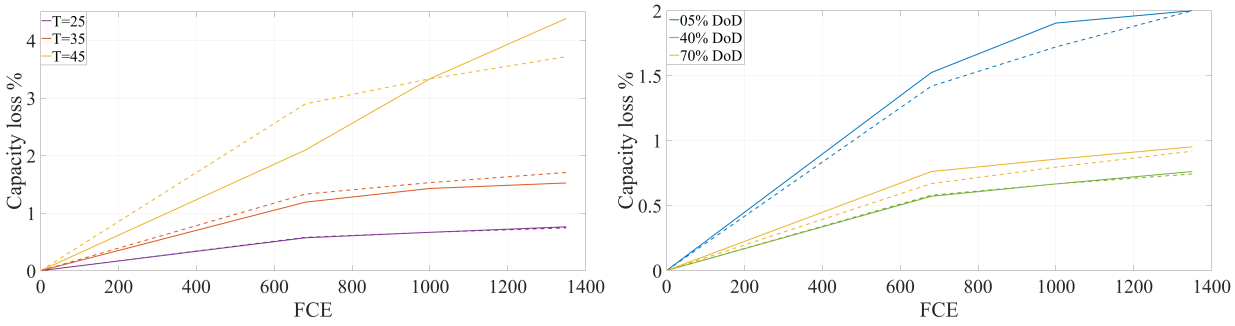
For the Calendar Aging, the required data is extracted and compared to the measured C_{loss} in (Dubarry et al., 2018). Table 3 shows a summary of the extracted data from this reference. To estimate the C_{loss} per 1 month,



(a) Capacity Loss as a function of FCE

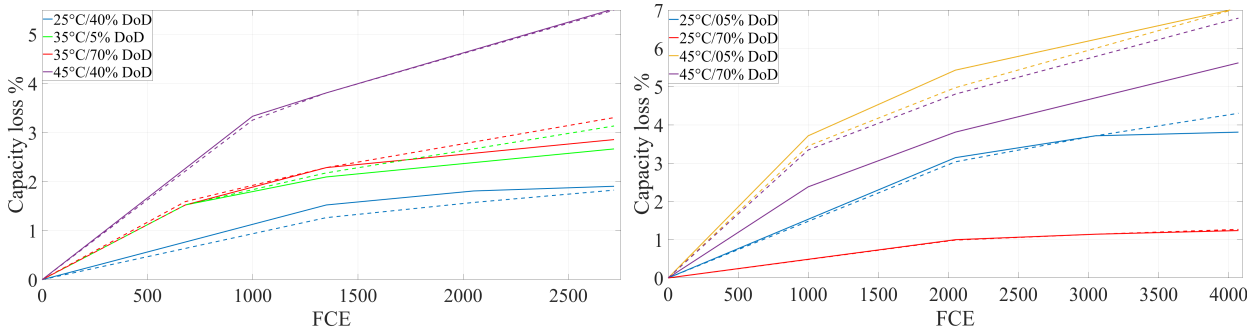
(b) Capacity Loss as a function of storage time

Figure 3: LTO measured Capacity Loss during Cycling and Calendar Aging (Dubarry & Devie, 2018).



(a) Capacity Loss estimated output during a fixed $DoD(40\%)$

(b) Capacity Loss estimated output during a fixed $T(25C)$



(c) Capacity Loss estimated output during a fixed $C-rate(2C)$

(d) Capacity Loss estimated output during a fixed $C-rate(3C)$

Figure 4: LTO estimated Capacity Loss during Cycling Aging.

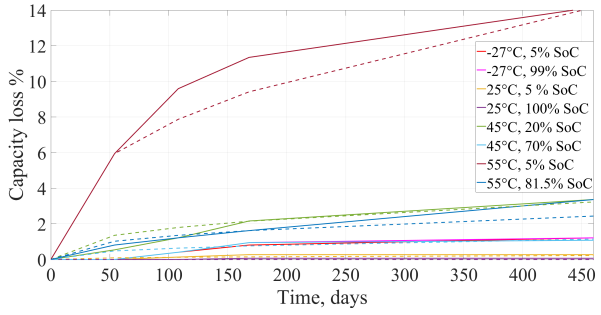


Figure 5: LTO estimated Capacity Loss during Calendar Aging.

`fitlm` from MATLAB is used to fit a linear regression model to obtain this value. Afterwards, the number of days is calculated and used as an input to the developed model. Table 3, 4th column shows the error between the estimated C_{loss} and the measured one. The model does have an average error of 1.98% in Calendar Aging.

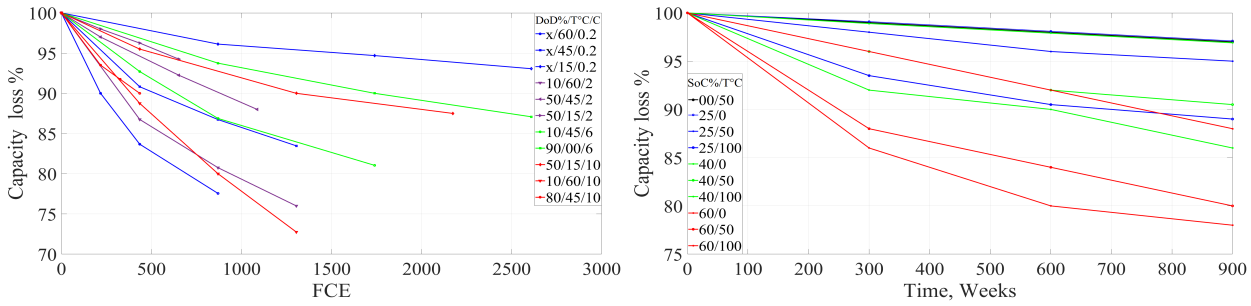
Table 3: Data validation during Calendar Aging

Temperature °C	SoC %	Capacity Loss (%/month)	Error %
-27	5	0.28	2.03
-27	99	0.28	1.95
25	50	0.05	0.67
25	100	0.05	0.68
45	20	0.76	2.74
45	70	0.24	1.76
55	5	3.97	3.37
55	81.5	0.73	2.67

Lithium iron phosphate battery (LFP)

Data Selection and Fitting

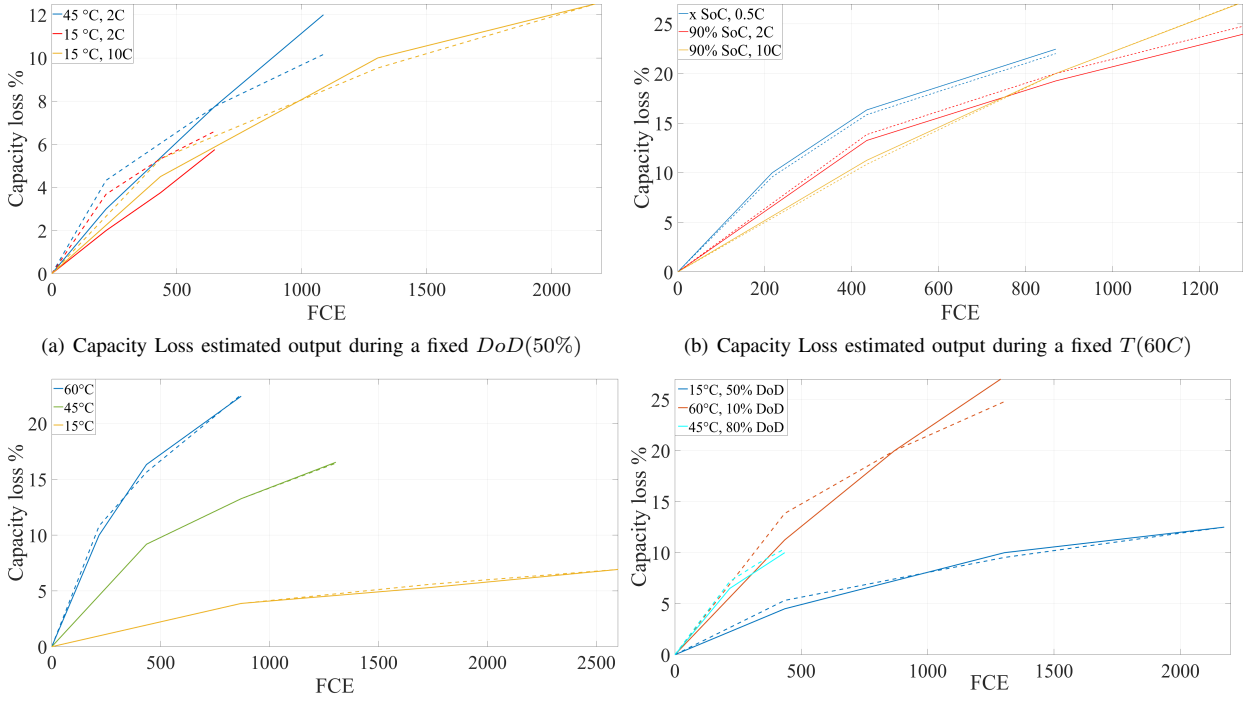
The measured data required for the fitting process during Cycling Aging is obtained from (Wang et al.,



(a) Capacity Loss as a function of FCE (Wang et al., 2011)

(b) Capacity Loss as a function of storage time (Naumann et al., 2018)

Figure 6: LFP measured Capacity Loss during Cycling and Calendar Aging.



(a) Capacity Loss estimated output during a fixed $DoD(50\%)$

(b) Capacity Loss estimated output during a fixed $T(60C)$

(c) Capacity Loss estimated output during a fixed $C - rate(0.5C)$

(d) Capacity Loss estimated output during a fixed $C - rate(10C)$

Figure 7: LFP Estimated Capacity Loss during Cycling Aging.

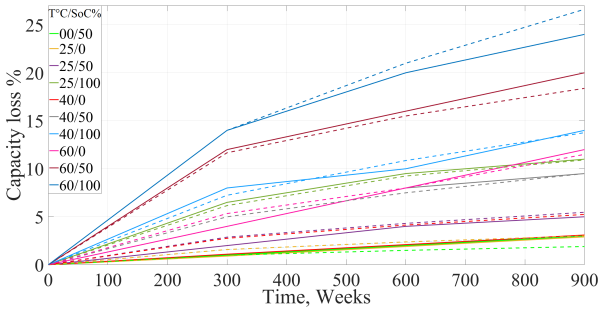


Figure 8: LFP Estimated Capacity Loss during Calendar Aging.

2011). In his work, he measured the C_{loss} during Cycling Aging in the following conditions; five different temperatures (0, 15, 25, 45, 60 °C), four levels of DOD (90%, 80%, 50%, and 10%), and four discharges rates (C/2, 2C, 6C and 10C). During discharge rate of C/2, the authors results showed that at such low rate, only temperature and FCE does have an effect on the C_{loss}, while DoD has a negligible effect on it. For the Calendar Aging C_{loss}, measured data was acquired from (Naumann et al., 2018), where the researchers studied the effect of different storage temperatures at the storage SoC = 0%,50%, and 100%.

Afterwards, the fitting process for the Cycling Aging is carried on. Similar to the LTO chemistry, the fitting process is proceed in four different scenarios; During a fixed $DoD(50\%)$, fixed temperature (60°C), and Fixed C-rate (0.5C and 10C). Figure 7 shows the estimated C_{loss} (dashed line) compared to the measured one (solid line). The average error is found to be 0.43% at half the FCE for each point, and 0.54% at the total FCE.

During the Calendar Aging process, the model has an average error of 0.45% at 450 days between the estimated C_{loss} (dashed line) and the measured one (solid line), and 0.67% at 900 days.

From the literature, it can be concluded that during Calendar Aging, the LFP chemistry degradation rate remains constant with SoC at given T, which means the C_{loss} is affected more by T and for the Cycling Aging, the C_{loss} is strongly affect by T and t on the LFP chemistry, while the DoD has almost a negligible effect especially at low C-rate (0.5C). In contrast,

Model Validation

In this section, the model is validated by comparing the estimated C_{loss} output to other known measured C_{loss}. For Cycling Aging, data needed to estimate the C_{loss} and compare it to a measured one is obtained from (Marongiu et al., 2015). Table 4, first column shows a summary of the data that used to estimate the Cycling Aging C_{loss}. The value of the battery temperature is kept at 30°C, and FCE is ranged from 1700 up to 5000.

As can been seen in Table 4, the model estimated C_{loss} error ranges from 1.2% to 1.8%, with an average error of 1.15%.

To validate the model in Calendar Aging, the same method for obtaining the required data as been used with LTO is used, and data is extracted from (Dubarry et al., 2018) . Table 5 shows a summary of the extracted data from this reference and the calculated error between the estimated C_{loss} during Calendar Aging and the measured one. The model does have an average error of 1.83% in Calendar Aging.

Table 4: Data validation during Cycling Aging measured in 30°C

Data cycling measured in 30°C	Measured Capacity Loss %	Estimated Capacity Loss %	Error %
SoC=90%, 1C	19.7	18.5	1.2
SoC=50%, 1C	17.3	18.5	1.2
SoC=20%, 1C	6.2	7.6	1.4
SoC=90%, 3C	36.6	34.9	1.7
SoC=50%, 3C	27.9	26.6	1.3
SoC=20%, 3C	56.3	54.5	1.8
SoC=90%, 6C	15.1	13.8	1.3
SoC=50%, 6C	68.1	66.7	1.4
SoC=20%, 6C	38.6	37.1	1.5

Table 5: Data validation during Calendar Aging

Temperature°C	SoC %	Capacity Loss (%/month)	Error %
0	50	0.40	0.5
10	50	0.52	0.6
20	100	1.37	1.1
25	40	1.30	1.4
30	65	0.97	2.1
40	30	1.79	1.5
45	100	3.72	2.3
50	20	3.19	4.1
60	0	1.78	2.9

Application example

After presenting, identifying and validating the battery capacity degradation model using two different types of LIBs, an attempt to predict the capacity degradation using simulated duty cycle data for a *Mining Dumper* based on (Immonen, 2003) work, is carried out. In this example, both LTO and LFP batteries are assumed to follow the same duty cycle, given in Figure 9 for Cycling Aging estimation, followed by 2 hours rest time during which the battery is cooled to the initial temperature (linear temperature decay) for the Calendar Aging estimation.

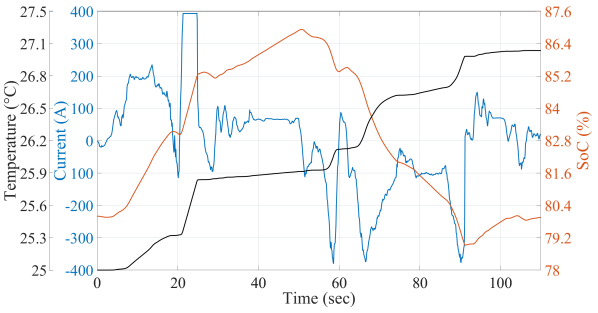


Figure 9: Simulated battery duty cycle data of a Mining Dumper.

The battery system was simulated for 1000 repeated duty and rest cycles (described above) for both LTO and LFP chemistries. The required data for predicting the capacity loss, found in Figure 9, was then fed to the proposed SEM. The results show that, for the LFP chemistry, $C_{loss}^{cyc} = 3.1\%$, and $C_{loss}^{cal} = 8.31\%$. On the other hand, for the LTO chemistry, $C_{loss}^{cyc} = 5.2\%$ and $C_{loss}^{cal} = 2.22\%$, indicating a clear difference in the aging profiles of the two battery chemistries.

CONCLUSIONS

In this work, a SEM for estimating the C_{loss} of different LIBs during Cycling Calendar Aging for different operation condition has been developed. For modeling, different LIB measurement data is required for obtaining the models' fitting parameters. Afterwards, the model is validated by comparing the estimated C_{loss} to a known one. Results show that the proposed model is able to estimate the C_{loss} during Cycling and Calendar Aging of two different LIBs (namely LTO and LFP chemistries) with a high accuracy with a brief explanation of the effect of different parameters on the two LIBs. Furthermore, a simulated battery duty cycle of a Mining Dumper has been used to estimate the C_{loss} during Cycling and Calendar Aging of a Mining Dumper.

The proposed model can be used in case of the lack of the experimental data, where it can give out an acceptable accuracy while having low computation complexity and is simpler to implement in comparison to other model-based approaches. For instance, the developed SEM has an average error of 2% between the measured C_{loss} and the estimated C_{loss}, while the adaptive-models do usually have an average error less than 1%.

In the future, the proposed SEM framework should be validated by studying different battery chemistries. Another important topic left for future work is design optimization — optimal cell chemistry selection in par-

ticular — for HDBEVs, based on the proposed battery degradation models.

REFERENCES

Andre, D., Nuhic, A., Soczka-Guth, T., & Sauer, D. (2013, mar). "comparative study of a structured neural network and an extended kalman filter for state of health determination of lithium-ion batteries in hybrid electric vehicles". *Engineering Applications of Artificial Intelligence*, *26*(3), 951–961.

Baure, G., & Dubarry, M. (2020). "battery durability and reliability under electric utility grid operations: 20-year forecast under different grid applications". *Journal of Energy Storage*, *29*, 101391.

Berecibar, M., Garmendia, M., Gandiaga, I., Crego, J., & Villarreal, I. (2016). "state of health estimation algorithm of lifepo4 battery packs based on differential voltage curves for battery management system application". *Energy*, *103*, 784–796.

Bocca, A., Sassone, A., Shin, D., Macii, A., Macii, E., & Poncino, M. (2015). A temperature-aware battery cycle life model for different battery chemistries. In *Ifip/ieee international conference on very large scale integration-system on a chip* (pp. 109–130).

Cacciato, M., Nobile, G., Scarcella, G., & Scelba, G. (2016). Real-time model-based estimation of soc and soh for energy storage systems. *IEEE Transactions on Power Electronics*, *32*(1), 794–803.

Cen, Z., & Kubiak, P. (2020). Lithium-ion battery soc/soh adaptive estimation via simplified single particle model. *International Journal of Energy Research*, *44*(15), 12444–12459.

Dubarry, M., & Devie, A. (2018). "battery durability and reliability under electric utility grid operations: Representative usage aging and calendar aging". *Journal of Energy Storage*, *18*, 185–195.

Dubarry, M., Qin, N., & Brooker, P. (2018). "calendar aging of commercial li-ion cells of different chemistries–a review". *Current Opinion in Electrochemistry*, *9*, 106–113.

Grolleau, S., Delaille, A., Gualous, H., Gyan, P., Revel, R., Bernard, J., ... Network, S. (2014). Calendar aging of commercial graphite/lifepo4 cell–predicting capacity fade under time dependent storage conditions. *Journal of Power Sources*, *255*, 450–458.

Immonen, P. (2003). *Energy efficiency of a diesel-electric mobile working machine* (Unpublished doctoral dissertation). Lappeenranta University of Technology.

Kluschke, P., Gnann, T., Plötz, P., & Wietschel, M. (2019). "market diffusion of alternative fuels and powertrains in heavy-duty vehicles: A literature review". *Energy Reports*, *5*, 1010–1024.

Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on optimization*, *9*(1), 112–147.

Liimatainen, H., van Vliet, O., & Aplyn, D. (2019). "the potential of electric trucks–an international commodity-level analysis". *Applied energy*, *236*, 804–814.

Liu, Z., Zhao, J., Wang, H., & Yang, C. (2020). "a new lithium-ion battery soh estimation method based on an indirect enhanced health indicator and support vector regression in phms". *Energies*, *13*(4), 830.

Marongiu, A., Roscher, M., & Sauer, D. U. (2015). "influence of the vehicle-to-grid strategy on the aging behavior of lithium battery electric vehicles". *Applied Energy*, *137*, 899–912.

Naha, A., Han, S., Agarwal, S., Guha, A., Khandelwal, A., Tagade, P., ... Oh, B. (2020, dec). "an incremental voltage difference based technique for online state of health estimation of li-ion batteries". *Scientific Reports*, *10*(1), 9526.

Naumann, M., Schimpe, M., Keil, P., Hesse, H. C., & Jossen, A. (2018). "analysis and modeling of calendar aging of a commercial lifepo4/graphite cell". *Journal of Energy Storage*, *17*, 153–169.

Petit, M., Prada, E., & Sauvant-Moynot, V. (2016, jun). Development of an empirical aging model for Li-ion batteries and application to assess the impact of Vehicle-to-Grid strategies on battery lifetime. *Applied Energy*, *172*, 398–407. doi: 10.1016/j.apenergy.2016.03.119

Schmalstieg, J., Käbitz, S., Ecker, M., & Sauer, D. U. (2014, jul). A holistic aging model for Li(NiMnCo)O2 based 18650 lithium-ion batteries. *Journal of Power Sources*, *257*, 325–334. doi: 10.1016/j.jpowsour.2014.02.012

Singh, P., Chen, C., Tan, C. M., & Huang, S.-C. (2019). Semi-empirical capacity fading model for soh estimation of li-ion batteries. *Applied Sciences*, *9*(15), 3012.

Standard practices for cycle counting in fatigue analysis (Standard). (2005). West Conshohocken, PA, USA: ASTM International.

Wang, J., Liu, P., Hicks-Garner, J., Sherman, E., Soukiazian, S., Verbrugge, M., ... Finamore, P. (2011). "cycle-life model for graphite-lifepo4 cells". *Journal of power sources*, *196*(8), 3942–3948.

Yang, K., Chen, Z., He, Z., Wang, Y., & Zhou, Z. (2020). "online estimation of state of health for the airborne li-ion battery using adaptive dekf-based fuzzy inference system". *Soft Computing*, *24*(24), 18661–18670.

AUTHOR BIOGRAPHIES

Mohammed Rabah PhD, is a Research Engineer at Computational Engineering and Analysis (COMEA) research group, Turku University of Applied Sciences, Finland.
E-mail:mohamed.rabah@turkuamk.fi

Eero Immonen D.Sc. (Tech.) and Adjunct Professor, leader of the Computational Engineering and Analysis (COMEA) research group at Turku University of Applied Sciences, Finland.
E-mail:eero.immonen@turkuamk.fi

Sajad Shahsavari works as Researcher at Turku University of Applied Sciences, and is a PhD student at University of Turku, Finland.
E-mail:sajad.shahsavari@turkuamk.fi

Mohammad-Hashem Haghbayan is a Post-Doctoral Researcher at Department of Future Technologies, University of Turku, Finland.
E-mail:mohammadhashem.haghbayan@utu.fi

Kirill Murashko PhD (M), is a Post-Doctoral Researcher at Department of Environmental and Biologic Science, University of Eastern Finland, Finland.
E-mail:kirill.murashko@uef.fi

Paula Immonen D.Sc. (Tech.) and Associate Professor of the School of Energy Systems, LUT University, Finland.
E-mail:paula.immonen@lut.fi

RESEARCH-AGENDA FOR PROCESS SIMULATION DASHBOARDS

Carlo Simon and Stefan Haag and Lara Zakfeld
Fachbereich Informatik
Hochschule Worms
Erenburgerstr. 19, 67549 Worms, Germany
E-Mail: {simon,haag,zakfeld}@hs-worms.de

KEYWORDS

Process Simulation, Visualization, Dashboard, Process Management, Petri nets

ABSTRACT

The *European Conference on Modeling and Simulation* is a prominent but not the only conference showing possibilities and relevance of simulation. Meanwhile, it is an important field of research worldwide and current discussions about the *industry of the future* and especially the idea of *digital twins* for the simulation of forecasts in parallel to an existing reality increase its importance.

All these efforts led to highly elaborated simulation modeling methods and tools that can be applied to different fields from air traffic management to zoo building. However, based on conference participations, literature research, and conversations with other researchers and practitioners, we observe that simulations are by far not being used as often as possible in day-to-day business. And if they are used, typically individual software solutions are developed that can hardly be transferred to other applications. So, how can we reduce the barriers for using simulation?

Any simulation comes along with a profound domain knowledge, a modeling method, a tool for the definition and simulation of models, and the visualization of the simulation results. Different roles conduct these tasks: Domain experts deliver the domain specific knowledge and – as is the case for further members of staff – must be able to interpret the simulation results. Modeling and visualization experts develop the simulations but also deliver a proper presentation for the domain experts, probably without having a deeper understanding of these results. A decision on whether a simulation is conducted at all is made by management, possibly together with the information systems department. The latter roles need information concerning the benefits both in advance as well as in retrospect.

Since we mainly work in the field of process modeling and simulation with the aid of Petri nets for production and logistics, the above made considerations encouraged further studies on the usage of simulation with a special focus on dashboard visualization of the simulation results in this field. A holistic approach includes the process of simulation development and use. The research agenda for which a grant could be won is explained within this paper and may animate other researchers to participate.

THE SIMULATION RESEARCH DILEMMA

A recent survey of 120 business decision-makers conducted in the DACH region yielded trend key topics – that is, not yet already establishing technologies – for companies in 2020. Regarding the actual utilization rate of some selected technologies in a direct process visualization context, metadata management leads with 27,3%, followed up by business activity monitoring with 17,5% and modeling digital twins for simulation of physical objects with 15,9%. (Roth and Heimann 2020)

In other words, around 70 to 85% of questioned companies do not currently utilize such technologies which leads to the question: What hurdles need practitioners to overcome in order to take more advantages from the simulation topic? We assume that – like it is the case for optimization – methods and tools are more in the focus of research than concrete applications or even an industrial usage of the methods which makes it hard for practitioners to apply them. Recent studies on optimization problems, for instance, demonstrate complex qualitative analysis and visualization results that are assessed with respect to their performance measures and indices (Koochaksaraei 2017). Although this is understandable from a research perspective, it does not help practitioners nor support the transfer into their day-to-day business.

We therefore advocate to embed simulation into a business perspective which has not been done sufficiently in the past. For example, the phases of a simulation study described in figure 1 ignore decisions on the simulation itself and does not explain the organizational roles involved in theses phases.

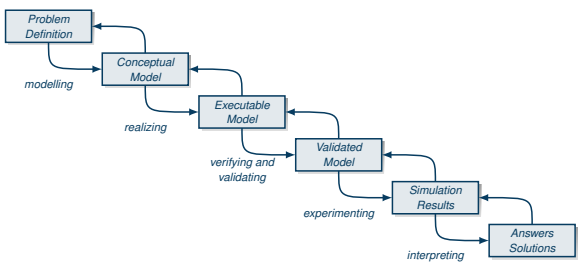


Figure 1: Phases of a Simulation Study adapted from (van der Aalst et al. 2010)

However, there is something else that matters with this phase model: It presumes that simulations are conducted for a singular purpose.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

Actually, this is true for many cases like a simulation in advance of a technical construction of a production line. But especially in the cases of forecasting and digital twins, simulation models are needed to be executed on a regular base, for example at the beginning of each planning phase. Obviously, former simulation results should then have an effect on later executions. The simulation environment could become a learning system.

Moreover, the term *simulation results* used in figure 1 leaves space for interpretation. Of course, this is a mathematical result, but the way it is presented might have consequences concerning the conclusions drawn from it. Although van der Aalst et al. (2010) see the need for interpreting the result in order to have answers to a given problem, we consider it worthwhile to think about different possibilities to present simulation results. It might even be necessary to present them from different perspectives in a dashboard like manner.

The adapted visualization pipeline of (Schumann and Müller 2000) shown in figure 2 explains the steps in which a visualization is developed. The major steps are *filtering*, *mapping*, and finally *rendering*. In comparison to figure 1, this pipeline goes one step further by also integrating the simulation results and their interpretation by the user for whom the simulation is conducted as well as a feedback loop to the visualization.

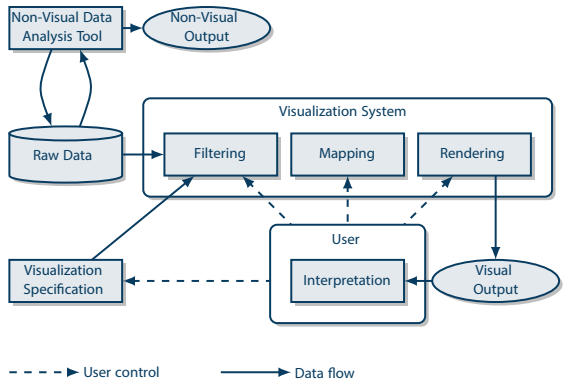


Figure 2: Extended Visualization Pipeline adapted from (Schumann and Müller 2000) based on (Robertson and De Ferrari 1994)

Visualization is meant to facilitate the assimilation of information and to compensate for the bottleneck in information processing. Through visualization, the viewer recognizes and understands connections within the data that would not have been recognized without visualization. Further, visualization can also be used to target the transmission of information.

A connection is built between insight, understanding, and explanation – also to third parties. Thus, visualization is a communicative process that needs to be constantly repeated and improved.

Although, in the last decades, many visualization systems enhanced their usability towards a better understanding of data, they are still difficult to use and are not always reliable, accurate tools for conveying information which limits their acceptance and use. (Telea 2015; Hansen and Johnson 2005)

The *Managed Simulation Process* (MSP) of figure 3 overcomes the limitations mentioned above. It integrates management objectives relevant at the beginning and for the economically successful completion of a simulation. The additional responsibility assignment matrix (also known as RACI matrix) shown in figure 4 includes the responsible roles for each activity of this process.

The MSP starts with a *decision* on whether a simulation is conducted at all. For this, a demand must be *recognized* first, followed by *choosing* a proper method and *selecting* a specific tool or consulting company. Finally, the simulation must be *approved*.

Then, the actual *simulation* is conducted which starts with an *analysis* and problem definition. Afterwards, the simulation *model* is developed (which of course may be a repetitive task) before the simulation can be *executed*. At the end, the simulation results have to be *validated* and probably need to be interpreted.

Visualization is a major outcome of a simulation. Following the simulation pipeline, the MSP considers the steps of *specifying* the intended output, *filtering* relevant data, *mapping* the data to interconnected observations, and finally *rendering* the visualization.

If simulations are conducted on a regular basis (for example for forecasts or production planning), the following steps should be iterated: First, the outcomes of the different simulation runs must be *monitored* and *reviewed* concerning their validity. Moreover, there exists a chance to *compare* the different simulation outcomes and to use former simulations to *improve* the model the simulation is based upon.

At the end, the economic outcome of the simulation must be assessed in a *control* phase. This starts with an (economic) *measure* of the improvement achieved by simulation, a *combination* of the findings with other possible approaches, a *revalue* of the entire result and an *established* routine (for or against) simulation.

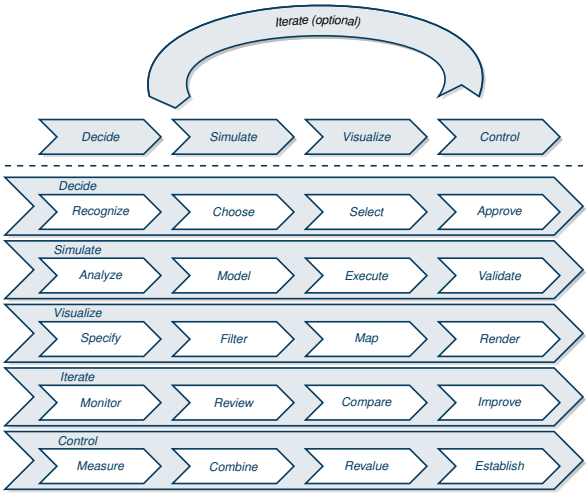


Figure 3: Phases of the Managed Simulation Process

Especially the RACI matrix uncovers the problem of establishing simulation as a regular routine in many corporations.

While management is not involved in the core tasks of simulation, it plays a major role at the beginning and the end of this process. If this is not taken into account, simulation might never be established as a method in a company. Especially a retrospective appraisal must not be forgotten, although this is typically not discussed in research papers on specific simulation methods.

We observe the following roles: *Domain experts* deliver the domain specific knowledge for the simulation and participate in the entire process. They must be the drivers for this process. *Modelers* and *visualizers* are the technical experts to perform the simulation. *Management* decides on whether time and money are invested for the simulation. Finally, the role of *procurement* might not be underestimated.

		Domain				Simulation
		Management	Procurement	Expert	Modeler	Visualizer
Decide	Recognize	I		R		
	Choose			R	C	
	Select		R	R	C	C
	Approve	R	I	I	I	I
Simulate	Analyze			R	R	C
	Model			C	R	R
	Execute			R	R	R
	Validate			R	R	R
Visualize	Specify			R	C	R
	Filter			I		R
	Map			I		R
	Render			I		R
Iterate	Monitor			R		
	Review			R		
	Compare			R		
	Improve			R	R	R
Control	Measure			R		
	Combine			R		
	Revalue	I		R	I	I
	Establish	R		I		

Figure 4: Responsibility Assignment Matrix of the Managed Simulation Process

In the following, we use this phase model to explain each step of a simulation and finally our new research approach we want to share with colleagues.

A DEMAND FOR SIMULATION

In order for a simulation to be conducted, *domain experts* must notice a demand for information that cannot be gained by other, more direct means, or alternative approaches would be much more expensive than conducting a simulation.

Choosing an appropriate method is a task for these experts together with *simulation modelers*. The term simulation modeler shall distinguish this role from other modeling experts who “only” develop conceptual models for explanation or visualization of a specific topic without claiming for simulation as a mathematical approach.

The tool selection necessitates *visualization experts* as a further role: they advise about the feasibility of different solutions with regards to the simulation goals.

The *procurement* department sets the financial borders for the investment with respect to the demands as shown by the domain experts. It might also influence the tool selection with respect to standard vendors of a company.

Lastly, the corresponding *managers* need to approve of the planned simulation project and release it.

THE SIMULATION MODEL

Domain experts, modelers, and visualizers are responsible for model and simulation execution. Cooperation among them is crucial for the success of a simulation project:

- As domain experts probably lack formal experience regarding simulation methods, they formulate the simulation goals and need to fully understand what the model does, what it does not, and what can be deduced.
- The modelers need a thorough understanding of the formal aspects of a system but may lack knowledge concerning its application or the practical impacts of the simulation results.
- The visualizers need to understand how the simulation results may be represented without having detailed knowledge on how to gather the (mathematical) simulation results.

Modeling and simulation methods can be classified by different aspects such as whether it can be conducted by an algorithm or not. Computer simulations may have several, sometimes contradictory goals that necessitate different modeling approaches. (Winsberg 2019)

If random elements have to be accounted for as for example customers entering a store in a queue, *non-deterministic* or *stochastic* models are used.

Deterministic models are used if a definite causality is given like in systems following natural principles. (Müller and Pfahl 2008)

In *chaotic* systems – which nonetheless are still deterministic – the simulation outcome varies strongly even on small deviations of the start setting (Bishop 2015). They may be of interest for scenario forecasts.

Dynamic simulations can be influenced during runtime – the simulation progress may depend on more than the factors given at the start, for example interspersed production failures or machine defects. This opens the possibility of testing adaptation capabilities. *Static* simulations allow for directly examining the implications of distinct start settings as constraints don't vary between different runs. This distinction can also be formulated by referring to the time of the last usage of new input data. (Müller and Pfahl 2008)

In static systems, *invariant* data is hardcoded in the model while *preprocessed* data is input at startup. Dynamic systems also use these data sets in addition to *runtime* data such as user input or sensor readings.

Discrete simulations run in time increments (or steps) of fixed length. A special case are *event-discrete* or *event-triggered* simulations where the steps don't have predetermined intervals. Rather, a new system state is computed on the moment of change.

In a *continuous* model, there are no steps whatsoever. Instead, the system's state gradually changes over time. This approach is more common to flow production or physics. (Müller and Pfahl 2008)

Finally, simulations may be distinguished in *quantitative* models – yielding numerical values – and *qualitative* models – with results being generally non-numerical. (Müller and Pfahl 2008)

As these classifications aren't mutually exclusive, models of interest normally combine several of the mentioned aspects, resulting in what is sometimes called hybrid models. However, this term is not uniformly defined and should be treated with care. (cf. Müller and Pfahl 2008; Brailsford et al. 2019)

THE MEANING OF NUMBERS

To convey information in a clear, concise and quickly accessible manner, different methods can be employed. Beside textual descriptions or tables, for example graphs and diagrams are often used. Obviously, the visualizers play the most important part in this step while, beyond the specification, the domain experts only have a passive role. Regarding the specification, knowledge of what is important for the real system and what is possible as output from the model are needed.

Visualizing the results

Visualizations provide high-level information about underlying data and (also simulated) processes. This data conveys insight into diverse applications such as procurement, production, computed on or distribution

To create a pertinent visualization of a given data set, the following aspects must be considered:

- The principles underlying the modeled real system.
- The presentation of the simulation models such that the domain experts can validate their correctness.
- The possibility to retrace the simulation in manageable amounts.
- Structure and quality of the input and output data.

But what exactly are the requirements that lead to a good visualization and which visualization types are suitable?

First, one should bear in mind that a short *text form* might sometimes be the better choice compared to a graphical visualization, since they do usually consume more space. Also, *tables* are still suitable in many situations, cause they both give an overview and support picking relevant data. A good visualization system is more than a mere presentation system in that it offers the possibility to uncover the contents. This is what is actually intended by figure 2.

According to the MSP, domain experts and modelers develop the data yield from the model and domain experts and visualizers do specify the visualization with the goal to transform simulation data into visualizations.

The remaining steps of this phase are at the responsibility of the visualizers with the domain experts being informed about the proceeding.

Mapping is a key element of visualization: this step converts invisible data into visible information. Mapping sets the visual attributes that encode the actual data. Rendering determines the remaining visual attributes that the visualizer can set. The final image is rendered using the processed data. (Robertson and De Ferrari 1994; Talea 2014)

Nowadays, countless variations of *diagrams* exist and with the upcoming data mining discipline their number seems to increase daily. Choosing appropriate diagrams is an art and a serious discipline which needs a lot of experience at the same time. This may be explained with the aid of some well-known diagram types:

- Point diagrams are used to depict relationships between categories.
- Line plot diagrams (cf. figure 5) are suitable to represent continuous data, but may confuse viewers if their axis-inscriptions are non-linear.
- Slope charts connect two comparison points in different categories.
- Bar charts are easy to read, especially if the x-axis-value is equal for each item, however visualizers must carefully decide on whether they draw them vertically or horizontally.
- Stacked bar charts may represent several categories at once.

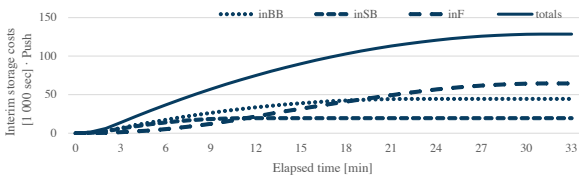


Figure 5: Example of a Line plot: Inventory Costs for Storages and Accumulated Totals (Simon et al. 2020)

Taking the following factors into account typically has a positive effect on a visualization:

- The degree of prior knowledge of involved persons.
- The visual abilities (or limitations) of the viewers and their general preferences.
- Using metaphors of the target discipline.
- Specific nature of the representation medium.

Obviously, not every visualization is target-oriented and by wrong types of visualization the simulation's audience may be lost. According to Nussbaumer Knaflic (2015), the following should be taken into account:

- Avoid pie charts, since their segments are difficult to compare for the observer, especially if they have almost the same size. Pie charts can often be replaced by bar charts which overcome these problems.
- Also, in donut diagrams, particularly if they make use of 3D effects, the different segments are difficult to compare.
- A second y-axis may cause confusion, if different areas are depicted by the x-axis. Simple labels or splitting up the information over several graphs might be better solutions to this problem.
- Lastly, avoid confusing colors, as shown in figure 5; this will be discussed later on.

There exist several other forms of diagrams which might all be useful in a specific context. An example unknown to a broader audience is the chord diagram which represents large amounts of data and their relationship to each other. (Jalali 2016)

Figure 6 shows an example of a chord diagram connecting several points of cumulative waiting time calculated by a Petri net simulation of a production process, actually the same process of figure 5.

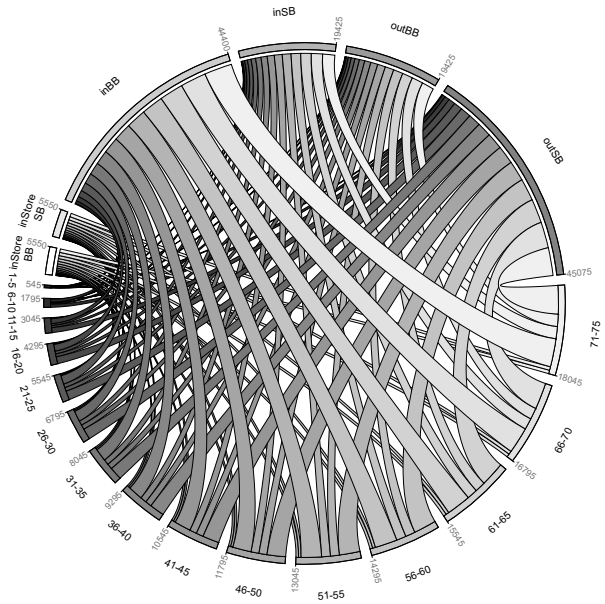


Figure 6: Example of a Chord Diagram: Waiting Times of Items in Different Settings (Haag et al. 2020)

This representation motivated to initiate the project described at the end of this paper to systematically do research on the visualization of process simulations with the aid of Petri nets.

Just the right visuals: Dashboards

The main goal of a dashboard is to create a coherent and coordinated picture that enables people to see and understand data. Dashboards convey data to ease their comprehension. Hereby, it is the aim to balance this goal with regards to efficiency and available time. On the way to effectiveness, accurate and functional thinking should be pursued (cf. Few 2013) with the option of presenting the data elegantly. The following ideas and criteria are helpful on this path (Wexler et al. 2017):

- Like in many other occasions, start with the definition of the specific goal of a dashboard to be developed and think of a user-friendly realization of this goal. Often, the possibility to adapt the dashboard for specific user needs is beneficial.
- Design a grid view which serves the purpose of clarity best, since cells of the same size make it easier to keep track of the data.
- Another design principle is described as breathing, a form of responsiveness.
- A purposeful use of color is crucial for users. Color can be a tool to determine a ranking, to define an area around a (neutral) center (e.g., the average), to separate categories by color or to set certain highlights.
- Also, different font sizes help users to keep a priority ranking. Simple numbers or statements as headlines can transport a clear message of the dashboard.
- Since dashboards are objective of discussions between different roles in a companies' organization, they must withstand different soft factors as well as subjective preferences.

In summary, a visual representation of monitoring data should be created with as little effort as possible, but this must lead to the best possible understanding for the majority of the target group consisting of management and domain experts.

REGULAR IMPROVEMENT BY SIMULATION

If simulations occur on a regular base, for instance in a digital twin of a real production environment to decide on different production strategies instantly, optimized simulation execution times become more and more important. The simulation must run fast enough to deliver results in a given reaction time. If simulations are used in planning processes instead, it must be faster and/or cheaper than other approaches, regardless of users' wishes to receive a quick simulation result.

Moreover, if simulations are executed and their results are interpreted on a regular base, a learning process should start with respect to the simulation itself. This is what the *Iteration* phase of the MSP expresses.

The iteration phase starts with observing the simulation, its results and interpretation which should be automated by a *monitoring* system. A deeper understanding of the simulation results can be gained from the current and former simulation runs which have to be *reviewed* and *compared*.

The probably most important step within this phase is *improvement*. Regular improvement cycles allow a constant adjustment of input data, simulation, and output data. A team consisting of domain experts, modelers and visualizers can use the lessons learned to improve the simulation and visualization system.

Long term, this procedure eliminates sources of variation in forecast or production data. Companies may abbreviate this phase when a sufficient level of quality is achieved.

CONTROL OF SUCCESS

Simulation is not an end in itself and the MSP therefore includes the management perspective at the beginning and the end. Ideas for a retrospective consideration of a simulation can be taken from the Control phase of Six Sigma projects. Like in Six Sigma, also simulation is intended to lead to an economic improvement.

Also, for this final phase, a visualization is crucial, however on a different level. It is not about the visualization of simulation runs but on the effect the conclusions of a simulation have on the simulated objective, in our case a business process in production or logistics. If, for instance, the simulation has the objective to reduce the throughput time of a production line, an economic result might be the new customer group that can be attracted by this.

A lot can be learned from Six Sigma at this point since also this approach considers process performance and process capability. Measures like defects per unit or the process capability index can be used.

What is crucial here is to carry out an ex-ante and ex-post analysis. An improvement (or probably a worsening) must be visible in a strategic dashboard, it must be verifiable and sustainable. The resulting improvements have to be controlled actively. (Tavasli 2007)

THE PATH TO THE GOAL

In order to deliver practitioners with suitable visualizations, a fitting modeling environment should provide possibilities for a wide range of application purposes. As theoretical foundation, we chose Petri nets for the following reasons:

1. Petri nets are *semantically clearly defined*, *well researched*, and *analyzable* with the aid of the large toolset linear algebra provides.
2. Petri nets provide a *high modeling power* and can be used in *process contexts*.
3. We have decades of *experience* with modeling and simulation using Petri nets.
4. With the Process-Simulation.Center (P-S.C), a web-based modeling and simulation environment based on Petri nets, an *advanced tool* exists that doesn't depend on commercial third-party integration.

Within a project it is the aim to do deeper research concerning the two phases *Simulate* and *Visualize* of the MSP. Since our simulation method is Petri nets, the goal is described as follows:

Create expressive visualizations of reachability sets and graphs of high-level Petri nets or excerpts therefrom both at runtime and at the end of the simulation.

To achieve this, we have established the following *research agenda*:

1. *Literature research*: Identify the state of the art of research on visualization of reachability sets and graphs of high-level Petri nets.
2. *Scenario creation*: Establish models of real-world scenarios based on logistics and production.
3. *Dashboard creation*: Identify frequently used key performance indicators and suitable visualizations. Adapt these visualizations for use in dashboards.
4. *Reference modeling*: Develop reference models for the automated and dynamic integration of visualizations and dashboards into simulation models. At this stage, applications for third party funding should be prepared.
5. *Prototyping*: Establish prototypes for the developed reference models. At this stage, applications for third party funding should be submitted.
6. *Transfer*: Infer research results to (regional) partners and enable them with the first iterations of usable prototypes and tools.
7. *Stabilization*: Expand on available simulation scenarios from other industries and departments. Stabilize the implementation.
8. *Evaluation*: Evaluate and document the results and communicate them with stakeholders and academia.

As first partners in both academia and economy have been found already, work on the first step has begun by now. We plan for the reference modeling to commence at the end of this year.

ACKNOWLEDGEMENT

This research was supported by “ProFIL - Programm zur Förderung von Forschungspersonal, Infrastruktur und forschendem Lernen der Hochschule Worms”.

REFERENCES

Bishop, R. 2015. "Chaos". In: *The Stanford Encyclopedia of Philosophy*. Stanford, CA. https://plato.stanford.edu/entries/chaos/ (last accessed 2021.04.10)

Brailsford, S. C.; T. Eldabi; M. Kunc; N. Mustafee and A. F. Osorio. 2019. "Hybrid simulation modelling in operational research: A state-of-the-art review". In: *European Journal of Operational Research* 278, 721-737.

Few, S. 2013. *Information Dashboard Design: Displaying Data for At-A-Glance Monitoring.* Analytics Press, Burlingame, CA.

Haag, S.; L. Zakfeld; C. Simon and C. Reuter. 2020. "Event Triggered Simulation of Push and Pull Processes". In: *SIMUL 2020: The Twelfth International Conference on Advances in System Simulation*, 68-73.

Hansen, C. D. and C. R. Johnson. 2011. *Visualization Handbook.* Elsevier, Amsterdam, NL.

Jalali, A. 2016. "Supporting Social Network Analysis Using Chord Diagram in Process Mining". In: *15th International Conference on Business Informatics Research*, 16-32.

Koochaksaraei, R. H.; I. R. Meneghini; V. N. Coelho and F. G. Guimarães. 2017. "A new visualization method in many-objective optimization with chord diagram and angular mapping". In: *Knowledge-Based Systems* 138, 134-154.

Müller, M. and D. Pfahl. 2008. "Simulation Methods". In: *Guide to Advanced Empirical Software Engineering*. Springer, London, UK, 119-152.

Nussbaumer Knaflic, C. 2015. *Storytelling with data: a data visualization guide for business professionals*. Wiley, Hoboken, NJ.

Robertson P. and L. De Ferrari. 1994. "Systematic Approaches to Visualization: Is a Reference Model Needed?". In: *Scientific Visualization: Advances and Challenges*. Academic Press, Cambridge, MA.

Roth, S. L. and T. Heimann. 2020. *IT-Trends 2020*. Capgemini, Berlin. https://www.capgemini.com/de-de/wp-content/uploads/sites/5/2020/02/IT-Trends-Studie-2020.pdf (last accessed 2021.04.10)

Schumann, H. and W. Müller. 2000. *Visualisierung - Grundlagen und allgemeine Methoden.* Springer, Berlin, Heidelberg, DE.

Simon, C.; S. Haag and L. Zakfeld. 2020. "Clock Pulse Modeling and Simulation of Push and Pull Processes in Logistics". In: *SIMMaApp: Special Track at SIMUL 2020: The Twelfth International Conference on Advances in System Simulation*, 31-36.

Tavasli, S. 2007. *Six Sigma Performance Measurement System: Prozesscontrolling als Instrumentarium der modernen Unternehmensführung*. Deutscher Universitätsverlag, Wiesbaden, DE.

Telea, A. 2014. *Data Visualization: Principles and Practice.* 2nd Edition. CRC Press. Boca Raton, FL.

van der Aalst, W. M. P.; K. Nakatumba-Nabende; A. Rozinat and N. Russell, N. 2010. "Business Process Simulation". In: *Handbook on Business Process Management 1*, 313-338.

Wexler, S.; J. Shaffer and A. Cotgreave. 2017. *The Big Book of Dashboards*. Wiley, Hoboken, NJ.

Winsberg, E. 2019. "Computer Simulations in Science". In: *The Stanford Encyclopedia of Philosophy*. Stanford, CA. https://plato.stanford.edu/entries/simulations-science/ (last accessed 2021.04.10)

AUTHOR BIOGRAPHIES

CARLO SIMON studied Informatics and Information Systems at the University of Koblenz-Landau. For his PhD, he applied process thinking to automation technology in the chemical industry. For his state doctorate, he considered electronic negotiations from a process perspective. Since 2007, he is a Professor for Information Systems, first at the Provadis School of Technology and Management Frankfurt and since 2015 at the Hochschule Worms. His e-mail address is: `simon@hs-worms.de`.

STEFAN HAAG holds degrees in Business Administration and Engineering as well as Economics with his main interests being related to modelling and simulation in graphs. After working at the Fraunhofer Institute for Systems and Innovation Research ISI Karlsruhe for several years, he is now a Research Fellow at the Hochschule Worms. His e-mail address is: `haag@hs-worms.de`.

LARA ZAKFELD graduated in International Logistics Management (B.A.) after completing an apprenticeship as a management assistant in freight forwarding and logistics services. She is currently pursuing a Master's degree in Entrepreneurship and works as a Research Assistant at the Hochschule Worms. Her e-mail address is: `zakfeld@hs-worms.de`.

Modeling and Simulation for Performance Evaluation of Computer-based Systems

MODELING AND ANALYZING CLOUD AUTO-SCALING MECHANISM USING STOCHASTIC WELL-FORMED COLOURED NETS

Mohamed M. Ould Deye
Mamadou Thiongane
Mbaye Sene
Department of mathematics and computer science
Cheikh Anta Diop University
Dakar, Senegal
{mohamed.ouldeye, mamadou.thiongane, mbaye.sene}@ucad.edu.sn

KEYWORDS
Auto-scaling, Cloud computing, Stochastic Well-formed coloured Nets

ABSTRACT

Auto-scaling is one of the most important features in Cloud computing. This feature promises cloud computing customers the ability to best adapt the capacity of their systems to the load they are facing while maintaining the Quality of Service (QoS). This adaptation will be done automatically by increasing or decreasing the amount of resources being leveraged against the workload's resource demands. There are two types and several techniques of auto-scaling proposed in the literature. However, regardless the type or technique of auto-scaling used, over-provisioning or under-provisioning problem is often observed. In this paper, we model the auto-scaling mechanism with the Stochastic Well-formed coloured Nets (SWN). The simulation of the SWN model allows us to find the state of the system (the number of requests to be dispatched, the idle times of the started resources) from which the auto-scaling mechanism must be operated in order to minimize the amount of used resources without violating the service-level agreements (SLA).

INTRODUCTION

Cloud computing environments offer service such as processing, bandwidth, and storage. Customers rent these services to deploy their applications and guarantee a certain quality of service for end users. There are many important features of clouds computing but one of those features that has made these systems successful is obviously auto-scaling or elasticity. Auto-scaling is the process that automatically readjusts the cloud computing resources according to the current system load. This adaptation results in increase resources (scale out) when the workload grows, and in decrease resources (scale in) when the workload drops. The primary benefit of auto-scaling, when configured and managed properly, is that the workload gets exactly the cloud computational resources it requires (and no more or less) at any given time. Customers pay only for resources they need, when they need them. There are several techniques of auto-scaling for determining the appropriate moment to scale resource. We can cite for example the Application Profiling Technique (APT) (Hector et al. 2014; Qu et al. 2016; Sharma et al. 2011), the Static Threshold-Based Rules (STR) (Fallah et al. 2015; Han et al. 2012), the Time Series Analysis (TSA) (Kumar and Singh 2018; Roy et al. 2011), the Machine Learning (ML) (Tesauro et al. 2006), and the Queuing Theory (QT) (Villela et al. 2007).

APT is a process of finding maximum point of resource used by an application with a certain workload. It is a simple ways to find the desired resources at a different point of time. STR is the most popular technique. It defines threshold resources utilization to scale in or scale out. A simple example: if CPU > 80%, then scale out; if CPU < 20%, then scale in. It is quite difficult to set the correct thresholds and this must be done manually. In this paper we propose a method to find them. TS includes a number of methods that use a past history window of a given performance metric in order to predict its future values. Three methods are often considered: moving average, exponential smoothing and linear regression. This technique forecast the future workload and resources required. QT is one of the widely used for modeling Internet applications. It is used in the analysis phase of the auto-scaling process. It estimates the performance metrics and waiting time for the requests. Queuing theory is a field of applied probability to solve the queuing problem. ML is a technique that is used on online learning for the constructing dynamic approach for the estimation of resources. It is a self-adaptive technique as per the workload pattern available. See (Al-Dhuraibi et al. 2018; Singh et al. 2019) for more details in auto-scaling techniques.

There are also two types of auto-scaling: horizontal auto-scaling and vertical auto-scaling. Horizontal auto-scaling refers to adding more virtual machines to the auto-scaling group. Vertical auto-scaling means scaling by adding more power rather than more units, for example in the form of additional Core or CPU or RAM. Indeed, whatever the type and technique of auto-scaling used, it is difficult to determine at which state of the system (load, queue size, CPU % used, etc) the resources should be increased or decreased in order to

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

minimize the resources used and guarantee the service-level agreements (SLA).

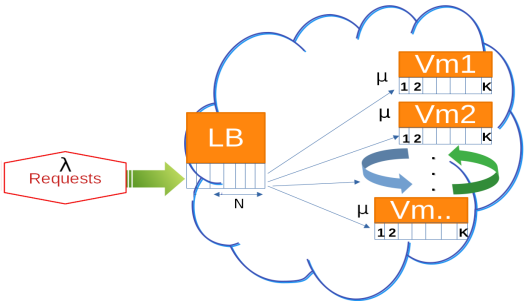
In this paper, we limit ourselves to cloud systems that allow only the STR auto-scaling method. The decision to scale out is taken when the load (the number of request) on the dispatcher is greater than or equal to N, and the decision to scale in is taken when a started resource (a virtual machine for horizontal auto-scaling, a Core or CPU for vertical auto-scaling) is idle for **α** seconds. Our goal is to determine the optimal couple (N, **α**) that minimizes the used resources without violating SLA in horizontal auto-scaling model and also in vertical auto-scaling model. We propose modeling the auto-scaling mechanism by the Stochastic Well-formed coloured Nets (SWN). We conduct simulation of the models to determine the best couple (N,**α**) for each one. Our model takes as input the arrival rate, the service time rate, the queue capacities, the SLA. The latter is defined in this work by mean response time.

The rest of the paper is organized as follows: the following section provides a description for the systems being modeled. The third section outlines the objective of the work. The fourth section presents the SWN models. The fifth presents the results of the simulation. The last section concludes the paper and gives some perspectives.

DESCRIPTION OF MODELED SYSTEMS

In order to model auto-scaling mechanism in Cloud Computing, we have considered the architecture illustrated on Figure 1 and described as follows: we have a system consisting of a set of servers (virtual machines: VM_1, VM_2, ...) and a Load Balancer (LB). These servers are identical and run the same stateless application. Each server has a service rate μ. The LB controls and manages the set of servers.

The requests arrive on LB with an arrival rate λ. The later forward them to servers for execution. Each server has a local queue of capacity K that stores requests waiting for their execution. The LB also ensures the admission control and the auto-scaling mechanisms. It has a queue with infinite capacity. In terms of admission control, when the LB receives a new request, it evaluates the load of all active servers and determines the server that should receive the request (using a certain policy). In case of saturation of all servers, the LB stores the request in its queue. When the queue size of the LB reaches a length N (a predefined parameter), the auto-scaling mechanism increases resources by adding a server in the case of horizontal auto-scaling or by adding a Core in the case of vertical auto-scaling. After this operation, the requests at the LB will be distributed with the new capacities or speeds of the system obtained from auto-scaling mechanism.



Figures 1: Architecture of the modeled system

In this system, when the horizontal auto-scaling mechanism is used, a server that remains idle for **α** (a predefined parameter) seconds is turned off, and in case of vertical auto-scaling, a core that remains idle for **α** seconds is removed.

OBJECTIVE

In this work, we sought to determine the best system states to operate the auto-scaling mechanism. The system state is defined by the number *n* of waiting requests on LB queue, the vector $\boldsymbol{v}=(x_1, x_2,\ldots,x_m)$ that contains the idle times of all started resources. x_i is the idle time for the resource *i*; $x_i=0$ if the resource *i* is executing requests; $x_i=s$ if the resource *i* is free since *s* second. If we observe a system state for which $n=N$ then we immediately add a new resource to the system to increase its capacity. If we observe a state of the system for which a value of $x_i=\boldsymbol{\alpha}$ then resource *i* is immediately switched off to decrease the capacity of the system.

The *N* and **α** that we are looking for are those that maximize the utilization rate of resources allocated to the application, minimize the amount of resources used by the application, while satisfying the level of service requested by the user through the SLA.

In this paper, the SLA is defined by mean response time. To find the best *N* and **α**, we conduct simulation of the model, monitor the mean response time and the amount of unused resources. Indeed, in the event of an overestimation, the indicator of unused capacity shows the low percentage of resource utilization. In the event of an underestimation, the response time indicators will show us high response time. These indicators, gathered together, can be used to fine the best *N* and **α** for an auto-scaling approach adapted to the context of the application concerned.

SWN MODELING FOR AUTO-SCALING

In this paper, we model the auto-scaling mechanism by SWN (Chiola et al. 1993). The SWN is an extension of colored petri nets. The main interest of this model is the possibility to have a reduction due to the symmetries of the Markov chain derived from the stochastic colored Petri nets (Chiola et al. 1993; Haddad and Moreaux 2009). In a SWN model the tokens have a color of a given set

which allows to model the characteristics of different entities of the same type (for example different types of servers or different types of requests); Places and transitions have a color domain. A color domain of a place identifies the tokens it can contain; that of a transition defines the type of values used to make it enabled. A color domain is a finished Cartesian product of elementary classes. Each elementary color class is a finite, non-empty set of terminal colors whose definition does not depend on any other color.

The other interesting aspect of SWN for our work here is the implicit synchronization between tokens of the same class. This allows us to easily define synchronizations based on the membership of a token or the membership of a part of a token. For example, when we are going to dispatch requests in the queues of different servers, we are going to do it associating for each request the name of the server that is going to execute it and there the implicit SWN synchronization guarantees that this request will only be executed by a CPU or Core of the designated server.

In this work, we are interested in horizontal auto-scaling and vertical auto-scaling. We propose a SWN model for each of them.

The SWN Model for Horizontal Auto-scaling

The Horizontal Auto-scaling SWN model that we propose is shown in Figure 2. For this model, we considered two types of servers: permanent servers and dynamic servers. Permanent servers refer to virtual machines that are started from the beginning and kept running all time. They represent the initial capacity of the system. The dynamic servers, on the other hand, are the virtual machines that will be created and later stopped, if necessary, by the auto-scaling mechanism.

We have two color classes to model the servers in our SWN model: the "PSRV" class represents permanent servers and the "DSRV" class represents dynamic servers. The "PSRV" class is used to initialize the marking of the "RunningVM" place, the "DSRV" class initialize marking of the "MaxCapa" place which represents the maximum capacity that auto-scaling can allocate to our system. The "RunningVM" place gives the number of servers running at a given time.

The requests arrive on the Load Balancer (the "LB" place) through the "IncomingRequests" transition. The "IncomingRequests" transition models the incoming flow of requests with arrival rate **λ**. The immediate transition "Affect" expresses the random dispatching of these requests to virtual machines. This dispatching is conditioned by the existence of free capacity on the servers.

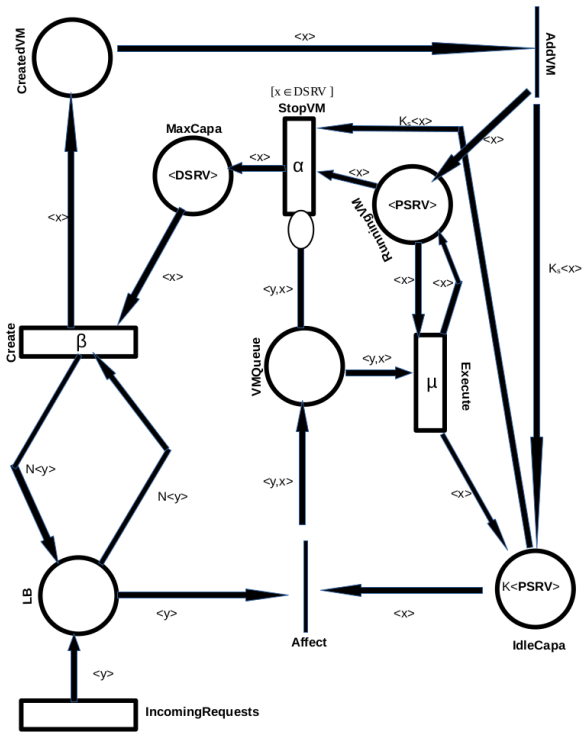


Figure 2: The SWN model for horizontal auto-scaling

The "IdleCapa" place models the available capacity in the system at a given time. The initial marking M_0 of this place is the length of all queues of all permanent servers.

$$M_0(\text{IdleCapa})=K<\text{PSRV}>$$

where <PSRV> gives the number of permanent servers. The "VMQueue" place represents the queues of all running servers. Each token in this place is <y, x> tuple composed of a request "y" and the name of the virtual machine "x" that will execute this request. The "Execute" transition models execution of requests and is defined using the **μ** service rate. This transition is configured in "infinite server" mode to model the parallel execution of virtual machines.

In this model, the auto-scaling mechanism is modeled by the "Create" and "StopVM" transitions. The "Create" transition models the creation of a new VM by the auto-scaling mechanism. The "StopVM" transition models the release of a VM that has no more requests to execute. For each of the two transitions, we have an exponential distribution to model the time required to complete the task in question. For the "Create" transition, the rate is $1/\beta$ where β is average time needed to create a new VM. For the "StopVM" transition, the rate is $1/\alpha$ where α is the average time to wait before stopping an idle VM.

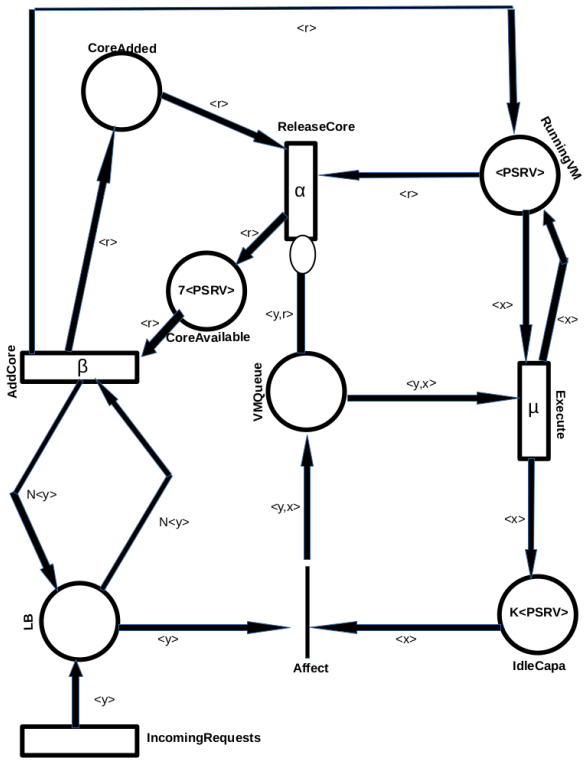
The "Create" transition requires a number N of tokens to create a new VM. This parameter "N" corresponds to the number of requests exceeding the capacity of the system and from which the scale out is operated. The

worst case with this parameter N is N=1 which means that a new VM is added as soon as there is a single request exceeding the system capacity. Indeed, delaying the addition of new resources pushes the system to make better use of the resources already allocated, hence the idea of looking for the largest N that respects the QoS.

The defined guard on "StopVM" transition indicates that only dynamic servers can be shut down. Stopping a VM causes the disappearance of "K" tokens from the "IdleCapa" place. This disappearance expresses the removal of its queue from the overall system capacity. The "AddVM" transition adds "K" tokens into the "IdleCapa" place when a new VM is created with the auto-scaling mechanism. The addition of "K" tokens into the "IdleCapa" place represents the creation of a new queue for the newly created VM.

The SWN Model for Vertical Auto-scaling

Figure 3 gives vertical auto-scaling SWN model. Here, we only have permanent servers. The computing capacity of these servers can be increased by adding new cores. The "CoreAvailable" place contains the number of cores available on the physical machines hosting the virtual servers. A core can be added to one of the running servers, if waiting requests on LB is equal to N.



Figures 3: The SWN model for vertical auto-scaling

The time required to add a Core is an exponential with rate **1/β** represented by "AddCore" transition. The "CoreAdded" place helps to recognize the Cores added by the auto-scaling mechanism. The "ReleaseCore" transition rate defines the time **α** after which an unused Core is released. The other elements of the model play the same roles as in the horizontal model (see Figures 3).

SIMULATION AND ANALYSIS OF RESULTS

Performance Measures

In this work, we use the following performances measures:

1. Average Response Time of the system (ART);
2. Average number of Resources Used (ARU).
3. Resource Utilization Rate (RUR)

These performance are calculated as follows:

$$\text{ART}=\frac{E(\text{LB})+\text{E}(\text{VMQueue})}{X(\text{IncomingRequests})}$$

$$\text{ARU}=\text{E}(\text{RunningVM})$$

where "E" is a function of GreatSPN(Amparore et al. 2016) that gives the average number of tokens of the place whose name is passed as an argument; and "X" is the function of GreatSPN that gives the average throughput of the transition whose name is passed as an argument.

$$\text{RUR}=100-\text{IC}$$

where

$$\text{IC}=\begin{cases}\text{AFP}-\text{ARS} & \text{IfAFP}\geq\text{ARS}\\ 0 & \text{if AFP}<\text{ARS}\end{cases}$$

where ARS is the average number of requests stored in the load balancer's queue.

$$\text{ARS}=\frac{E(\text{LB})*100}{E(\text{RunningVM})*K}$$

and AFP is the average number of free places in server queues:

$$\text{AFP}=\frac{E(\text{IdleCapa})*100}{E(\text{RunningVM})*K}$$

Simulated Examples

To analyze these models, we used GreatSPN's simulator (Amparore et al. 2016). The servers on which the application is running have each one a service rate μ = *4 requests per time unit.* In this work, we take the second as the basic unit of time for our model. We have defined our basic configuration with two permanent servers. The maximum number of extra resources that auto-scaling mechanism can create for our system is limited to 10 VMs in horizontal case, and 14 Cores in vertical case. The queue capacity K of a VM is equal to

50. For the sake of simplicity and without loss of generality, we will assume that all flavours are single-core in the case of Horizontal Auto-scaling. The creation time of a new resource is an exponential distribution with rate 1/β=0.5. This means that it will take 2s on average to create a new VM or add a Core in the vertical case. The incoming flow rate **λ**=7. The QoS constraint to be guaranteed is the "average response time" (ART), and it must remain ≤ 7second. Table 1 summarizes the different parameters used in our simulation.

Table 1: Simulation parameters

Parameter	**Values**
Mean response time	≤ 7s
Service rate (μ)	4 requests per second
VM's Queue capacity (Ks)	50
LB's queue capacity (K_L)	∞
Permanent servers	2 VM
Number of extra resources	10 VMs or 14 Cores
Arrival rate λ	7
Rate $1/\beta$	0.5
Rate $1/\alpha$	{0.1, 0.5}
N	{1, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120}

In order to determine the best couple (N,α), we simulate the models and calculated the performance measures for several N values and for several α values, see Table 1.

To quickly find the best couple (N,α), we can begin research with large N values and small values of α. For example we can start with N=K and depending on the average response time observed and its acceptable limit value, we will know if it is necessary to increment or decrement the N. However, if we want to show the general behavior of Auto-scaling mechanism, we can begin with the worst value N=1. Here, we chose to present simulation result with the last case. The increment step to reach the optimal value is 10.

Table 2 shows the average response time (ART), the resource utilization rate (RUR), the average number of VMs (ARU) used for horizontal auto-scaling as a function of N and α. We observe that small values of N give small ART but lead to low resource utilization rates. The increase of N increases the RUR; that is good things but increase also the ART. The increase of the latter is blocked by constraint of the SLA. So we see clearly with the constraint on the SLA (ART must be inferior or equal to 7 second), that the maximum resource utilization rate is 77.52% with N=70. We simulate the system with several value of α varying from 2 to 10 per step 2, but in Table 2, we report result for only α =2 and α =10. Comparing performance for α=2 and α=10, we observe that α=2 give better rate of utilization resource.

Table 2: System performances as function of N and α with an horizontal auto-scaling mechanism

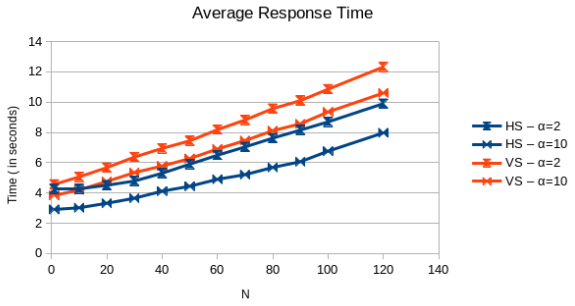
H-Scaling	α=2			α=10		
N	ART	RUR	ARU	ART	RUR	ARU
1	4.25	45.85	2.6	2.91	29.28	2.78
10	4.28	46.46	2.59	3.02	30.52	2.77
20	4.51	48.52	2.61	3.32	33.88	2.76
30	4.78	51.68	2.6	3.65	37.42	2.74
40	5.3	57.58	2.58	4.11	42.5	2.7
50	5.9	64.39	2.56	4.44	45.96	2.71
60	6.49	71.47	2.54	4.9	51.2	2.69
70	7.04	77.51	2.54	5.21	54.4	2.68
80	7.62	84.57	2.52	5.68	59.71	2.67
90	8.15	91.11	2.51	6.06	63.62	2.66
100	8.69	96.64	2.52	6.76	71.93	2.64
120	9.9	100	2.49	7.97	85.72	2.61

Table 3 shows the average response time (ART), the resource utilization rate (RUR), the average number of Cores (ARU) used for vertical auto-scaling as a function of N and α. We observe the same result as in auto-scaling horizontal.

Table 3: System performances as function of N and α with an vertical auto-scaling mechanism

V-Scaling	α=2			α=10		
N	ART	RUR	ARU	ART	RUR	ARU
1	4.56	71.71	2.56	3.83	68.24	2.56
10	5.06	76.82	2.52	4.2	70.81	2.52
20	5.67	83.62	2.52	4.75	75.97	2.52
30	6.37	91.39	2.5	5.34	81.8	2.5
40	6.94	97.56	2.46	5.78	86.04	2.46
50	7.44	100	2.47	6.27	91.02	2.47
60	8.18	100	2.45	6.9	97.5	2.45
70	8.82	100	2.47	7.46	100	2.47
80	9.57	100	2.43	8.12	100	2.43
90	10.1	100	2.43	8.57	100	2.43
100	10.85	100	2.45	9.35	100	2.45
120	12.32	100	2.43	10.6	100	2.43

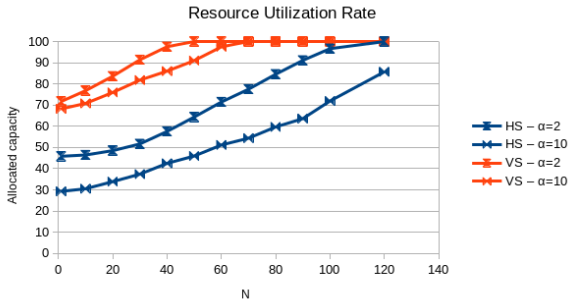
Figure 4 shows the evolution of the mean response time for the two types of auto-scaling and for two *α* values.



Figures 4: Average Response Time as function of N

These results show that delaying the addition of resources leads to an increase in the mean response time for all types of auto-scaling. They also show that delaying the release of inactive resources improves the mean response time.

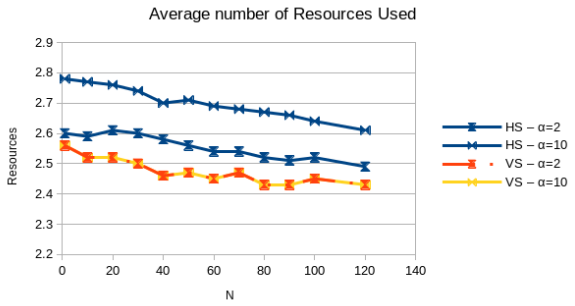
It can also be seen that the average response time recorded by the Horizontal Auto-scaling in this specific case is smaller than the average recorded with the Vertical Auto-scaling mechanism. However, in terms of resource utilization rates, the results in Figure 5 show a higher utilization rate with Vertical Auto-scaling compared to Horizontal Auto-scaling.



Figures 5: Resource Utilization Rate as function of N

We observe in Figure 5 that delaying scale-out increases the rate of resource utilization. It can also be seen that delaying the release of resources naturally leads to a significant waste of resources, particularly for the horizontal mechanism.

The same reading can be made from Figure 6. Here, the vertical auto-scaling recorded a better average of Cores used. In addition we can also see that this score remained constant with the two **α** values.



Figures 6: Average number of Resources Used as function of N

According to these results and in order not to violate the fixed QoS constraint, the N to be retained can be chosen among the following values in Table 4. The best couples were highlighted.

Table 4: The best couples

	Candidate couples	Recorded utilization rate	Mean number of Cores used
Horizontal scaling	**N=70,α=2**	77.52%	2.54
	N=100,α=10	71.93%	2.64
Vertical scaling	N=40, α =2	97.56%	2.46
	N=60, α =10	97.50%	2.45

CONCLUSION AND PERSPECTIVES

In this paper, we proposed a modeling of the auto-scaling mechanism with SWN models. We studied two types of systems. The first system uses the horizontal auto-scaling mechanism, and the second system uses the vertical auto-scaling mechanism. For each system, the arrivals requests follow a Poisson process, and the service times follow an exponential distribution.

The simulation of models allowed us to fine the best load N of LB from which we have to scale out, and the best idle time α of a resource from which we have to scale in. The best couple (N, α) allows us to minimize the number of resources to be allocated to the system on the one hand and on the other hand to ensure a high rate of use of allocated resources. In a future work, we want to test our using real data. It would also be interesting to study self-scaling mechanisms with time-varying arrival rates.

REFERENCES

Al-Dhuraibi Y., Paraiso F., Djarallah N., Merle P.. Elasticity in Cloud Computing: State of the Art and Research Challenges. *IEEE Transactions on Services Computing, IEEE*, 2018, 11 (2), pp.430-447. ⟨https://dx.doi.org/10.1109/TSC.2017.2711009⟩.

Amparore E.G., Balbo G., Beccuti M., Donatelli S., Franceschinis G. (2016) 30 Years of GreatSPN. In: Fiondella L., Puliafito A. (eds) Principles of Performance and Reliability Modeling and Evaluation. *Springer Series in Reliability Engineering. Springer*, Cham. https://doi.org/10.1007/978-3-319-30599-8_9

Chiola G., Dutheillet C., Franceschinis G. and Haddad S., "Stochastic well-formed colored nets and symmetric modeling applications," in *IEEE Transactions on Computers*, vol. 42, no. 11, pp. 1343-1360, Nov. 1993, https://doi.org/10.1109/12.247838

Fallah, M., & Arani, M.G. (2015). ASTAW: Auto-Scaling Threshold-based Approach for Web Application in Cloud Computing Environment. *International Journal of u- and e- Service, Science and Technology*, 8, 221-230.

Haddad S., Moreaux P., Stochastic Well-formed Petri Nets. *M. Diaz. Petri Nets: Fundamental Models, Verification and Applications, Wiley-ISTE*, pp.303-320, 2009. ⟨http://hal.univ-smb.fr/hal-00441928⟩.

Han R., Guo L., Ghanem M., and Guo Y., "Lightweight Resource Scaling for Cloud Applications," in *2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, May 2012, pp. 644–651, (https://doi.org/10.1109/CCGrid.2012.52).

Hector F., Guillaume P., and Thilo K.. 2014. Autoscaling Web Applications in Heterogeneous Cloud Infrastructures. In *Proceedings of the 2014 IEEE International Conference on Cloud Engineering (IC2E '14). IEEE Computer Society*, USA, 195–204. DOI:https://doi.org/10.1109/IC2E.2014.25.

Kumar J., and Singh A. K. 2018. Workload prediction in cloud using artificial neural network and adaptive differential evolution. *Future Gener. Comput. Syst. 81, C* (April 2018), 41–52. DOI:https://doi.org/10.1016/j.future.2017.10.047

Qu C., Calheiros R. N., Buyya R., 2016. A reliable and cost-efficient auto-scaling system for web applications using heterogeneous spot instances. *J. Netw. Comput. Appl. 65, C* (April 2016), 167–180. DOI:https://doi.org/10.1016/j.jnca.2016.03.001

Roy, N., Dubey, A., and Gokhale, A. 2011. Efficient Autoscaling in the Cloud Using Predictive Models for Workload Forecasting. In Proceedings of the *2011 IEEE 4th International Conference on Cloud Computing (CLOUD '11). IEEE Computer Society*, USA, 500–507. DOI:https://doi.org/10.1109/CLOUD.2011.42

Sharma U., Shenoy P., Sahu S. and Shaikh A., "A Cost-Aware Elasticity Provisioning System for the Cloud," *2011 31st International Conference on Distributed Computing Systems*, Minneapolis, MN, USA, 2011, pp. 559-570, https://doi.org/10.1109/ICDCS.2011.59.

Singh P, Gupta P, Jyoti K and Nayyar A (2019). Research on Auto-Scaling of Web Applications in Cloud: Survey, Trends and Future Directions. *Scalable Computing: Practice and Experience*, 20(2), 399–432.

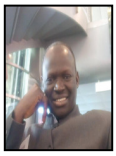
Tesauro, G., Jong, N. K., Das, R., and Bennani, M. N. 2006. A Hybrid Reinforcement Learning Approach to Autonomic Resource Allocation. In Proceedings of the *2006 IEEE International Conference on Autonomic Computing (ICAC '06). IEEE Computer Society*, USA, 65–73. DOI:https://doi.org/10.1109/ICAC.2006.1662383

Villela, D., Pradhan, P., and Rubenstein, D. 2007. Provisioning servers in the application tier for e-commerce systems. *ACM Trans. Internet Technol.* 7, 1 (February 2007), 7–es. DOI:https://doi.org/10.1145/1189740.1189747

AUTHOR BIOGRAPHIES

Mohamed M. O. DEYE is an Assistant Professor at the Department of mathematics and computer science at Cheikh Anta Diop University of Dakar, Senegal. His areas of interest are Cloud Computing, Web Services and performance evaluation of distributed systems. His email address is mohamed.oulddeye@ucad.edu.sn

Mamadou THIONGANE is an Assistant Professor at Cheikh Anta Diop University of Dakar, Senegal. . His main research interests are waiting time prediction in call centers, modeling service time in service system. He are also interest by Cloud Computing system. His email address is mamadou.thiongane@ucad.edu.sn



MBAYE SENE is a full professor in the department of Mathematics and Informatics of the Faculty of Sciences and Technology in the university of Cheikh Anta Diop of Dakar. He received the PhD degree in Computer Science from the University of Paris Dauphine, France, in 2002 and a master degree of Management of organizations in the same university in 2003. He was during 2 years an associate professor in Paris-Dauphine before he integrates the most important university in Senegal, UCAD. His main research interests include distributed database systems, design of inter-operate open systems, wireless de sensor networks and performance evaluation of stochastic complex systems. He has authored or co-authored more than 20 international conference papers and journals. Professor Mbaye SENE mange also since 2008 big projects in the ministry of employment and vocational training of Senegal; he has worked with France Agency of Development (AFD), GIP International (France).

Telling faults from cyber-attacks in a multi-modal logistic system with complex network analysis

Dario Guidotti, Giuseppe Cicala, Tommaso Gili, Armando Tacchella

KEYWORDS

Cyber-Security and Critical Infrastructure Protection, Complex Networks, Discrete Event Simulation.

ABSTRACT

We investigate the properties of systems of systems in a cybersecurity context by using complex network methodologies. We are interested in *resilience* and *attribution*. The first relates to the system's behavior in case of faults/attacks, namely to its capacity to recover full or partial functionality after a fault/attack. The second corresponds to the capability to tell faults from attacks, namely to trace the cause of an observed malfunction back to its originating cause(s). We present experiments to witness the effectiveness of our methodology considering a discrete event simulation of a multimodal logistic network featuring 40 nodes distributed across Italy and daily traffic roughly corresponding to the number of containers shipped through in Italian ports yearly averaged daily.

INTRODUCTION

Complex networks are significantly present in many science disciplines and have recently received much attention [Bar]. Many studies have been devoted to measuring networks' robustness against attacks or random degradation failures causing deletion of nodes or connections. Such measures are used to increase the security of complex systems and possibly to improve their robustness [AJB00], [KG14]. Edges of a network usually play the role of transmitting information or load and maintaining network connectivity. The load model of cascading failures can be used to investigate small-world network performance subject to deliberate attacks on node and edge. Results show that edge attacks produce more significant cascading failures than node attacks. On the other hand, in real-world networks, the nodes vulnerable to attack are often well protected, while edges are a relatively easy target for attackers [NGZL15].

Systems of systems, e.g., water treatment plants, electric grids (power plants and associated distribution networks), industrial plants, transportation networks, and smart homes, are the ideal field of application for complex network theory to obtain useful insights about the behavior of the systems under scrutiny. In such systems, wireless communication among components and external network access for supervisory control and data acquisition (SCADA) make them an ideal target for cyber-attacks. It has been demonstrated that malicious users can gain control of such systems and/or disrupt their functionality severely [FR11]. This is also true for systems that are part of critical national infrastructure (CI). As such, intentional or accidental incidents that alter their normal behavior can have dramatic effects on the safety of citizens [WFD10].

Among other security-related issues, resilience is recognized as one of the keys to understanding how much damage can be brought to a system and its surrounding environment in case of a successful cyber-attack [DRKS08]. The concept of resilience — defined as "*the quality of being able to return quickly to a previous good condition after problems*" — emerges as an additional target, complementary to protection from external threats, but not subordinate to it. More recently, the term *cyber-resilience* has been coined to identify specifically "*the ability to continuously deliver the intended outcome despite adverse cyber events*" [BHSZ15], and this is the interpretation we consider in this paper, where we are interested in applying complex networks analysis to obtain a measure of resilience.

Our research goal is to discriminate whether a random fault or a cyber-attack causes the performance degradation a system incurs into and the amount of such degradation. We call this *attribution* and we hypothesize that it relates strongly to the ability to trace the cause of an observed malfunction back to its cause(s). We need to stress that we are not interested in the specific originating event but rather differentiating system-related events from cyber-related events. We find that a clear answer to this question may be the solid basis for any attribution process targeted to spot attackers.

We implemented complex network metrics on a realistic multi-modal logistic system. We embedded a hypothetical network into the Italian railway system, providing coverage of the entire national territory using 40 terminals (nodes) so that each one serves an area of approximately 150Km in radius. Simulation parameters – e.g., number of trains with their schedules and routes, number of containers with their origins and final destinations – are chosen according to stochastic models. Events generated by the simulator are stored (i) in a database, and basic KPIs can be computed out of these data. The user can inject both faults and attacks in the simulation, so that their effects can be observed in the results. Simulations tested our methodology's effectiveness considering daily traffic scenarios approximately corresponding to the number of containers shipped through Italian ports yearly.

Results clearly show that complex network analysis enables the assessment of cyber-resilience and gives us indications to understand whether a system's performance degra-

Dario Guidotti, Giuseppe Cicala 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: dario.guidotti@edu.unige.it, giuseppe.cicala@unige.it, armando.tacchella@unige.it. Tommaso Gili is with IMT Lucca ... E-mail: tommaso.gili@imtlucca.it. The authors wish to thank ... The corresponding author is Armando Tacchella.

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

dation is due to a fault or some malicious activity.

BACKGROUND

In our methodology, we consider different topological measures from undirected graphs.

Definition 1 (Graph) A Graph is a pair $G = (V, E)$ where V is a set whose elements are called vertices and E is a set of paired *vertices*, whose elements are called *edges*. We briefly present our measures of interest in the following. The first measure we consider is Laplacian Energy, i.e., the sum of the absolute values of the eigenvalues of the Laplacian matrix of the graph. This quantity is often studied in the context of spectral graph theory and chemistry studies.

Definition 2 (Laplacian Energy) [GZ06] Let G be a graph with n vertices and no loops or parallel edges. Let L be the Laplacian matrix of G and μ_i, $i = 1, ..., n$ the eigenvalues of L. Then the Laplacian energy of the graph is defined as:

$$E(G) = \sum_{i=1}^{n} \mu_i^2 \tag{1}$$

Another measure of interest for graphs, in general, is the centrality of vertices. In this work, we have chosen to consider Laplacian Centrality [QFW+12] and Betweenness Centrality [Fre77]. We use these measures to understand the relevance of the nodes and edges of the graph.

Definition 3 (Betweenness Centrality) Let G be a graph with n vertices and no loops or parallel edges. The Betweenness centrality of the vertex i is defined by the number of shortest paths that pass through i. Precisely, let L_{hj} be the total number of shortest paths from a vertex h to another vertex j and $L_{hj}(i)$ be the number of shortest paths that pass through the vertex i. The Betweenness centrality of vertex i can be defined as

$$\frac{2}{(n-1)(n-2)} \sum_{h \neq i} \sum_{j \neq i, j \neq h} \frac{L_{hj}(i)}{L_{hj}} \tag{2}$$

Definition 4 (Laplacian Centrality) Let G be a graph with n vertices and no loops or parallel edges. Let $E_L(G)$ be the laplacian energy of G and $E_L(G_i)$ the laplacian energy of G after the vertex i has been removed. The laplacian centrality of vertex i is defined as

$$\frac{E_L(G) - E_L(G_i)}{E_L(G)} \tag{3}$$

Definition 5 (Community Structure and Communities) A graph is said to have a community structure if the graph's nodes can be easily grouped into sets of nodes such that each set of nodes is densely connected internally: each set of nodes is a community. One of the reasons for the importance of communities is that they often present very different properties than the average properties of the corresponding network, therefore concentrating only on the average property usually misses important and interesting features of the network. In this work, we consider the communities generated using the Louvain algorithm [BGLL08].

Definition 6 (Giant Component) It is the largest connected component of a given graph that contains a finite fraction of the vertices. We partitioned the graph into several connected components by removing the least important edges according to a percolation approach [LLL+21].

Fig. 1: Graphical representation of ONTOMIL network.

Edges removal stops when the difference of the two largest connected components' size is less or equal to a specific value.

ONTOMIL SIMULATOR

The simulator models an *Intermodal Logistics System* (ILS) which support receiving, storing and shipping goods packaged in *Intermodal Transport Units* (ITUs, also known as "containers"). A detailed description of the context is provided in [CCT13]. Here we restrict our attention to ILSs wherein rail transportation is supported by a network of terminals equipped with systems for fast ITU handling. The overall network is "covered" by relatively frequent short-distance trains with a fixed composition and a predefined daily schedule. ITUs enter the network at some terminal and travel to their destination according to a predefined route, usually boarding more than one train along the way. While this solution enables efficient utilization of resources, information technology is vital to operate it effectively.

Operation of the simulated ILS involves several customers forwarding their goods through the system and an handling agent, i.e., the business responsible for managing the entire network. Given its role, the handling agent is also the main stake-holder, and the one who is thought to collect *key performance indicators* (KPIs) to be computed on data about the system. Transportation across the network is organized by having customers emit *requests for work* which contain ITUs to be sent from a given terminal to other destinations on the network. The handling agent associates to each request for work a number of *transport orders*, one for each ITU listed in the request for work. The transport order contains all the data related to the shipping, like ITU route through the network and expected time of delivery. Once the ITU corresponding to a given transport order is collected at a terminal, it is boarded on the first outgoing train whose

destination is compatible with its route. Since trains travel across relatively short distances, it is possible to dispatch ITUs more than once during a 24 hours time-span.

The main activity of the simulated ILS is to satisfy the supplied demand of transportation in a timely way in spite of events potentially disrupting the service like, either due to natural causes, e.g., network and rolling stock failures, or due to malicious activity, e.g., cyber-attacks targeting single terminals or the whole network. We describe how such events are injected in the simulator later on as part of our methodological approach, but here we observe that monitoring ITUs from the departure terminal to their final destination is a key enabler for every kind of analysis on the network. In particular, the data obtained through monitoring enables the computation of KPIs which summarize the overall status of the system and its ability to handle a given workload over time. In particular, ONTOMIL provides the following "standard" indicators:

1. Late transport orders on a daily basis, i.e., the number of transport orders issued on a given date whose ITUs did not reach the final destination on the same date.
2. Cumulative number of ITUs handled in terminals.
3. Average number of ITUs unloaded per hour in the network terminals.
4. Late trains, i.e., trains suffering one or more delays with respect to their schedule on a specific route.
5. Recent sink/source terminals, i.e., terminals wherein the only operations were loadings or unloadings over the past hour.
6. Number of customers whose request for work contains transport orders backlogged for more than two days – calculation done on a daily basis.
7. Number of loading and unloading operations for each terminal on an hourly basis.
8. High-activity customers on a daily basis, i.e., those customers shipping more ITUs than a given daily threshold.
9. Average train utilization on an hourly basis, i.e., number of ITUs vs. number of trains travelling across the network.
10. Route utilization, i.e., cumulative number of ITUs which traveled along a given route.

The above KPIs can be grouped in different categories. For instance, KPI 2 and 9 are considered *critical success factors*, in the sense that they highlight potential flaws in network organization which should be corrected to maintain efficiency, e.g., wrong train scheduling and routing. Most of the remaining KPIs are so called *dashboard indicators*, in the sense that they provide useful information to quantify the overall health status of the network, and can support tactical decision making. Change in some of the dashboard indicators impacts on the ability of the whole system to generate revenues for the handling agent.

The actual simulated model consists of a hypothetical logistic network covering the entire Italian territory using 40 intermodal terminals connected between them through railroads. A representation of such network is given in Figure 1. Albeit ONTOMIL simulates an ideal system, the railroad connection correspond to the actual freight lines along which goods are forwarded by train. As shown in Figure 2, ONTOMIL enables the customization of different parameters, including the number of simulation days, the minimum

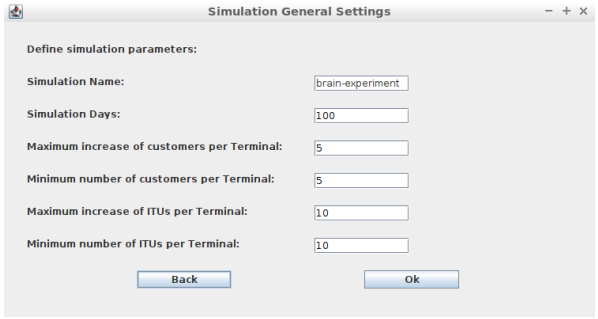


Fig. 2: Graphical Interface for the selection of the parameters of the ONTOMIL simulation.

number of customers insisting on any terminal on each single day and others that affect the number of ITUs in the system. The actual values are chosen randomly for each terminal before the beginning of each day of simulation respecting the bounds defined by the user. In our simulations we considered a number of customers and ITUs that results in a number of ITUs circulating in the network which is comparable to the number of ITUs handled daily by Italian ports and forwarded on railway trains.

METHODOLOGY

Analyzing the ILS

To analyze the ILS we abstract it as two different kinds of graph and we analyze how they change in different temporal intervals. The terminals correspond to the vertexes of the graphs whereas the connections between the terminals correspond to the edges of the graphs. The first graph we considered is the Flux Graph (FG) whose weights are the number of ITUs present on the corresponding connection during the chosen interval of time. The second graph is the Difference Flux Graph (DFG) whose weights are the difference in absolute value between the number of ITUs present on the corresponding connection during the chosen interval of time and the number of ITUs present on it during the previous interval of time. In this work we have chosen a single day as the interval of time of interest given the characteristic of the simulation. The idea behind the FG is to provide a snapshot of the performance of the ILS during a particular day, whereas the idea behind the DFG is to provide a snapshot of the evolution of the ILS between a particular day and the next.

Fault and attack injection

In order to test the proposed methodology to analyze the ILS, we added on top ofthe ONTOMIL simulator the capability of injecting faults and attacks. We think of the former as naturally occurring events, e.g., delay along a line due to a locomotive malfunction, and the latter as the result of a malicious activity, e.g., an hacker infiltrating the shipping network and altering the transport orders. The capability to inject faults and attacks, alone or combined, is crucial to demonstrate the effectiveness of the methodology proposed. As shown in Figure 3, ONTOMIL currently supports the injection of two kinds of anomalies:

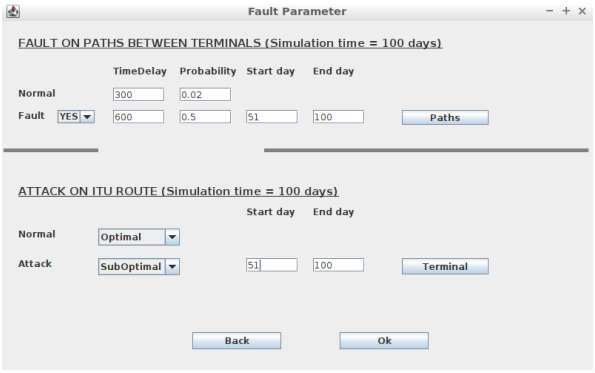


Fig. 3: Graphical Interface for the control of faults and attacks in the ONTOMIL simulation.

- Fault on paths between terminals, where the user can tune the "physiological" delay that affects trains traveling across all routes with a given probability ("Normal" time delay and related probability), and introduce a "pathological" delay which is larger in magnitude and probability of occurrence and it is meant to affect specific paths — including the possibility to target all the paths.
- Attach on ITU route, where the user can decide to change the normal routing of ITUs through the network, based on the minimum number of terminals hops, to some anomalous routing that could be just suboptimal or plain wrong, e.g., an ITU that should go from terminal A to terminal B is sent to a terminal C where no connection to B exists; the attack can affect all the terminals or just specific one, again based on user's choice.

It is important to observe that these anomalies do not cover all potential incidents due to natural causes or malicious attacks that can happen in a network like the one simulated by ONTOMIL. However, train delays represent a frequent event and alteration of transport orders is the easiest way that an hacker has to alter the normal behavior of the network and cause disruption in service. Also, they represent two fundamentally different ways to cause such a disruption, one that relates to the physical nature of the process, the other that relates to the control of the same. In principle, further anomalies can be injected into the ONTOMIL simulator, but these will be the subject of future work.

EXPERIMENTAL EVALUATION

The results of our experiments can be seen in Tables I and II. For each measure of interest, we have computed the p-value with the Wilcoxon Signed-Rank test and the Cohen's d coefficient, comparing the samples collected during normal operation of the system and the ones collected during the fault/attack. The goal is to understand whether the distribution of a specific measure differs, in a statistically significant way, when considering normal operation and fault/attack injection. Given the results, it is clear that some of the measures of interest present distributions which have this characteristic, because they are significantly different during the normal operations and the fault/attack and/or they are different based on the event injected (fault or attack). In the tables, we have highlighted in green all the p-values smaller than 10^{-4} and all the Cohen's d coefficients greater than 1.

As we would expect, the simulations without any attack or fault do not present any statistically significant difference between the various measures. However, when the network is under attack, the Louvain Energy measures computed over the Flux Graphs present a significant difference. Moreover, when the attack is applied to the high importance terminals also the Giant Energy measure present a comparable difference. The same phenomenon occurs for the high importance fault simulations regarding the Louvain Energy measure computed over the Difference Flux Graph. Regarding the simulation in which both the attack and faults occurs on the high relevance terminals and routes, all the measures except the Giant Energy computed over the DFGs present a significant difference.

Given the observations above we can define a set of rules (represented as a decision tree in Figure 4) based on the statistical significance of the difference of the measures computed on the data sampled during different time-windows of the simulations. In particular, it appears clear that an attack can be identified using the Louvain Energy or Giant Energy measures computed on the FGs, a fault can be identified using the Louvain Energy measure computed on the DFGs and the presence of both fault and attack can be identified by the significance of both Louvain Energy computed on FGs and DFGs at the same time. In general, variations of the Louvain Energy computed on the FGs pinpoint attacks both on low and high importance Terminals, whereas variations of the Louvain Energy computed on the DFGs pinpoint faults only on the high importance routes.

To understand the motivations behind the general inability to identify faults on low importance routes we must refer to Images 5, 6 and 7 in which we show the trend of two KPIs of interest during a simulation in which the system was under attack, one in which it was experiencing a fault and one in which it was experiencing both. As it can be seen, the performances of the system clearly deviates from normal conditions either under attack and under both attack and fault, and this happens both when their intensity is low and when it is high. On the other hand, when only the low intensity fault is applied, the trend of the KPIs is almost identical to the baseline one. This shows clearly that we are unable to identify low intensity faults because their effects on the system are negligible.

CONCLUSIONS

We have shown that, considering a hypothetical but realistic case study, complex network analysis is capable of quantifying the decrease of resilience in a system under fault/attack and to tell the difference between the two. As a future work, we plan to consolidate our methodology by further integrating by formalizing the theoretical connections between the observed measures and the dynamics of the underlying system. On the engineering side, we wish to extend our analysis to cover other systems of systems, and further validate our methodology by extending it to evaluate resilience of other critical-infrastructure facilities, with a focus on energy production plants and distribution networks.

Experiment	Simulation	Louvain Energy		Giant Energy	
		FG	*DFG*	*FG*	*DFG*
EXP 1	Standard	0.341	0.701	0.233	0.142
	Attack (H)	$1.302 * 10^{-9}$	0.039	$3.091 * 10^{-8}$	0.716
	Attack (L)	$1.339 * 10^{-8}$	0.001	$1.150 * 10^{-4}$	0.059
	Fault (H)	0.009	$7.774 * 10^{-6}$	0.717	0.411
	Fault (L)	0.437	0.134	0.060	0.124
	Both (H)	$1.383 * 10^{-9}$	$2.789 * 10^{-9}$	$1.585 * 10^{-8}$	0.126
	Both (L)	$1.983 * 10^{-8}$	$6.569 * 10^{-4}$	0.260	0.043
EXP 2	Standard	0.653	0.020	0.919	0.289
	Attack (H)	$1.302 * 10^{-9}$	0.527	$1.417 * 10^{-8}$	0.838
	Attack (L)	$4.790 * 10^{-8}$	0.026	0.012	0.043
	Fault (H)	0.454	$4.458 * 10^{-7}$	0.081	0.694
	Fault (L)	0.143	0.988	0.114	0.452
	Both (H)	$2.661 * 10^{-9}$	$1.309 * 10^{-8}$	$4.790 * 10^{-8}$	0.005
	Both (L)	$2.478 * 10^{-8}$	0.069	0.105	0.924
EXP 3	Standard	0.151	0.489	0.813	0.716
	Attack (H)	$7.159 * 10^{-9}$	0.002	$4.012 * 10^{-9}$	0.389
	Attack (L)	$1.549 * 10^{-7}$	0.005	$1.150 * 10^{-4}$	0.754
	Fault (H)	0.015	$1.544 * 10^{-7}$	0.382	0.988
	Fault (L)	0.881	0.608	0.739	0.650
	Both (H)	$7.556 * 10^{-10}$	$9.773 * 10^{-9}$	$4.067 * 10^{-8}$	0.208
	Both (L)	$1.468 * 10^{-9}$	0.001	0.017	0.020

TABLE I: Table of p-values for our experiments. The **Experiment** column indicates the experiments of interest. The **Simulation** column identifies the specific simulation inside a specific experiment, L indicates that the low importance terminals/routes were chosen for the attack/fault whereas H indicates that the high importance ones were chosen. **Louvain Energy** and **Giant Energy** indicate the measure considered and they represent the Laplacian Energies of the Louvain Communities graph and the Giant Component graph respectively. *FG* and *DFG* represent the Flux Graph and the Difference Flux Graph respectively.

Experiment	Simulation	Louvain Energy		Giant Energy	
		FG	*DFG*	*FG*	*DFG*
EXP1	Standard	0.135	0.003	0.288	0.310
	Attack (H)	2.341	0.290	1.623	0.167
	Attack (L)	1.682	0.576	0.802	0.409
	Fault (H)	0.497	1.005	0.102	0.171
	Fault (L)	0.208	0.191	0.376	0.344
	Both (H)	2.033	1.683	1.567	0.272
	Both (L)	1.460	0.702	0.338	0.261
EXP2	Standard	0.116	0.430	0.004	0.245
	Attack (H)	2.233	0.104	1.749	0.017
	Attack (L)	1.617	0.466	0.548	0.467
	Fault (H)	0.242	1.285	0.321	0.200
	Fault (L)	0.296	0.060	0.369	0.157
	Both (H)	1.827	1.613	1.609	0.485
	Both (L)	1.591	0.416	0.431	0.027
EXP3	Standard	0.260	0.031	0.040	0.017
	Attack (H)	1.698	0.585	1.969	0.200
	Attack (L)	1.295	0.598	0.940	0.118
	Fault (H)	0.499	1.535	0.203	0.044
	Fault (L)	0.052	0.188	0.060	0.194
	Both (H)	2.012	1.681	1.446	0.339
	Both (L)	2.043	0.672	0.481	0.488

TABLE II: Table of the Cohen's d coefficients for our experiments. The **Experiment** column indicates the experiment of interest. The **Simulation** column identifies the specific simulation inside a specific experiment, L indicates that the low importance terminals/routes were chosen for the attack/fault whereas H indicates that the high importance ones were chosen. **Louvain Energy** and **Giant Energy** indicates the measure considered and they represent the Laplacian Energies of the Louvain Communities graph and the Giant Component graph respectively. *FG* and *DFG* represent the Flux Graph and the Difference Flux Graph respectively.

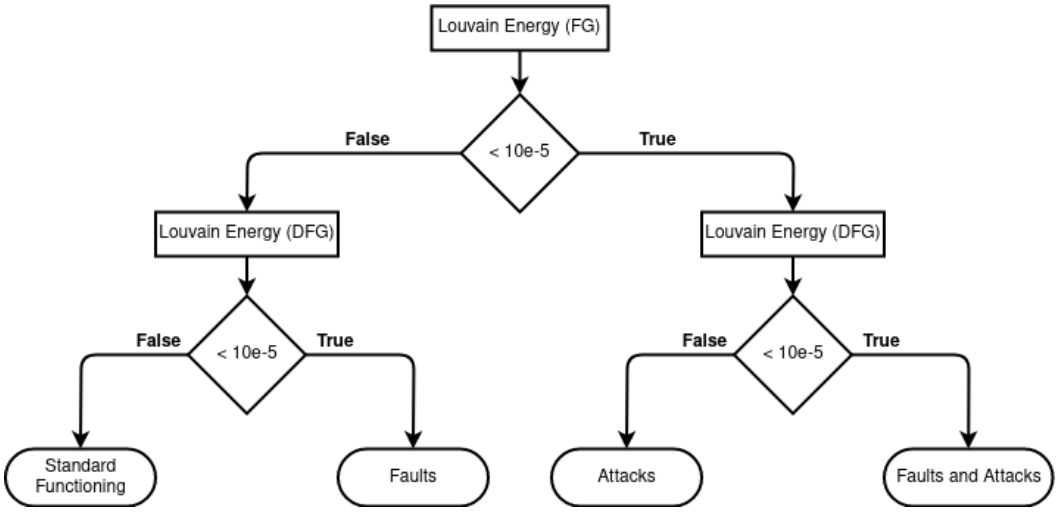
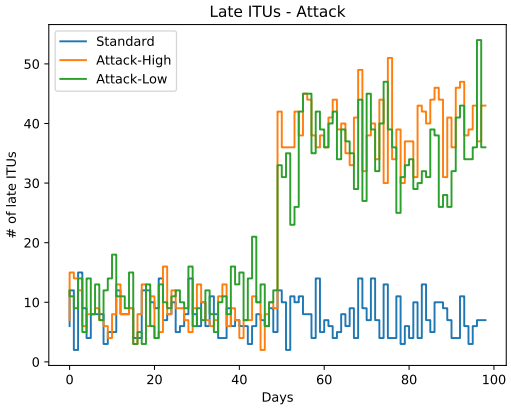


Fig. 4: Graphical representation of the rules extracted by our analysis on the ILS.



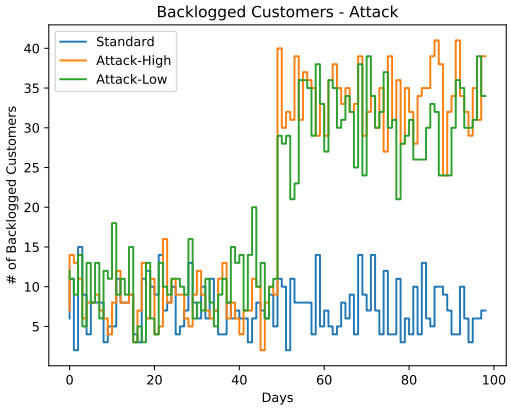
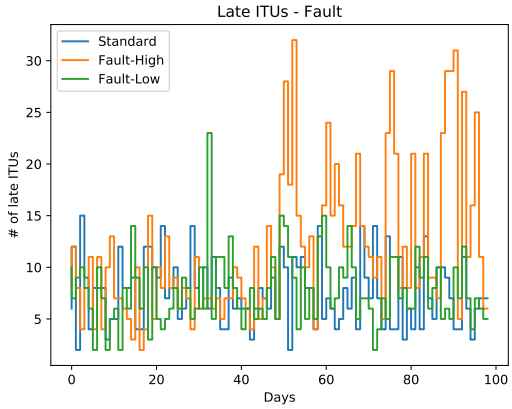


Fig. 5: Trends of the KPIs Late ITUs and Number of Backlogged Customers during a simulation subject to an attack.



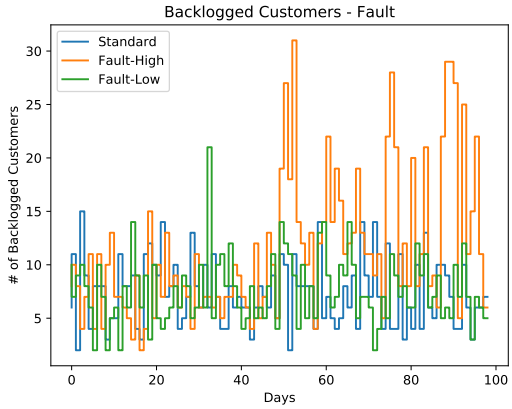


Fig. 6: Trends of the KPIs Late ITUs and Number of Backlogged Customers during a simulation subject to a fault.

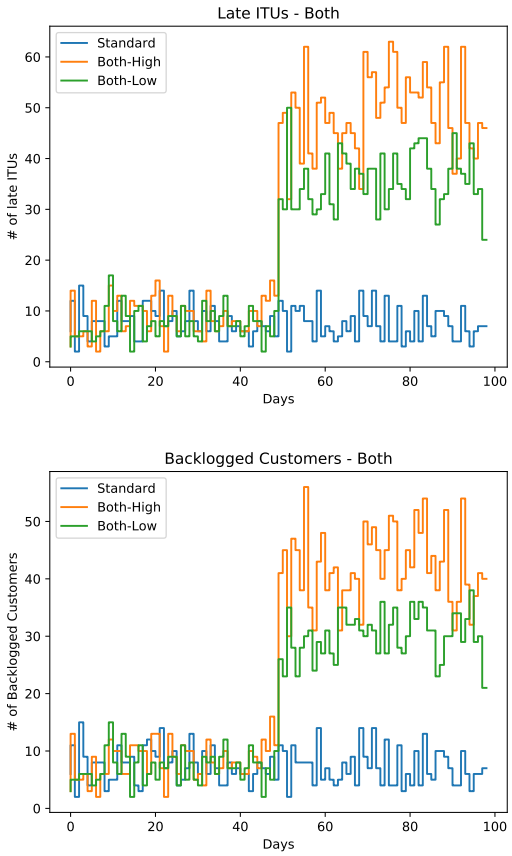


Fig. 7: Trends of the KPIs Late ITUs and Number of Backlogged Customers during a simulation subject both to a fault and an attack.

REFERENCES

[AJB00] Reka Albert, Hawoong Jeong, and Albert-Laszlo Barabasi. Error and attack tolerance of complex networks. *Nature*, 406:378382, 2000.

[Bar] Albert-Laszlo Barabasi. *Network Science*. Cambridge University Press.

[BGLL08] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[BHSZ15] Fredrik Björck, Martin Henkel, Janis Stirna, and Jelena Zdravkovic. Cyber resilience-fundamentals for a definition. In *WorldCIST (1)*, pages 311–316, 2015.

[CCT13] Matteo Casu, Giuseppe Cicala, and Armando Tacchella. Ontology-based data access: An application to intermodal logistics. *Inf. Syst. Frontiers*, 15(5):849–871, 2013.

[DRKS08] Salvatore DAntonio, Luigi Romano, Abdelmajid Khelil, and Neeraj Suri. Increasing security and protection through infrastructure resilience: the inspire project. In *International Workshop on Critical Information Infrastructures Security*, pages 109–118. Springer, 2008.

[FR11] James P Farwell and Rafal Rohozinski. Stuxnet and the future of cyber war. *Survival*, 53(1):23–40, 2011.

[Fre77] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.

[GZ06] Ivan Gutman and Bo Zhou. Laplacian energy of a graph. *Linear Algebra and its applications*, 414(1):29–37, 2006.

[KG14] Marek Korytar and Darja Gabriska. Integrated security levels and analysis of their implications to the maintenance. *Journal of Applied Mathematics, Statistics and Informatics*, 10:3342, 2014.

[LLL+21] Ming Li, Run-Ran Liu, Linyuan Lu, Mao-Bin Hu, Shuqi Xu, and Yi-Cheng Zhang. Percolation on complex networks: Theory and application. *Physics Reports*, in press, 2021.

[NGZL15] Tingyuan Nie, Zheng Guo, Kun Zhao, and Zhe-Minhg Lu. New attack strategies for complex networks. *Physica A: Statistical Mechanics and its Applications*, 424:248–253, 2015.

[QFW+12] Xingqin Qi, Eddie Fuller, Qin Wu, Yezhou Wu, and Cun-Quan Zhang. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*, 194:240–253, 2012.

[WFD10] Chunlei Wang, Lan Fang, and Yiqi Dai. A simulation environment for scada security analysis and assessment. In *Measuring Technology and Mechatronics Automation (ICMTMA), 2010 International Conference on*, volume 1, pages 342–347. IEEE, 2010.

METADATA FOR ROOT CAUSE ANALYSIS

Alexander A. Grusho, Nick A. Grusho, Michael I. Zabezhailo,
Elena E. Timonina and Vladimir V. Senchilo
Federal Research Center "Computer Science and Control"
of the Russian Academy of Sciences
Vavilova 44-2, 119333, Moscow, Russia
Email: grusho@yandex.ru, info@itake.ru, zabezhailo@yandex.ru,
eltimon@yandex.ru, volodias@mail.ru

KEYWORDS

Anomalies in distributed information systems, Approximate method of searching for anomalies, Metadata.

ABSTRACT

The paper is devoted to the task of finding the root cause of anomaly in a distributed information and computing system. An approximate approach is considered to detect implicit anomalies with accuracy to the object (of a component of the technical device, a node of a network infrastructure, an application or of an information resource). The approximate solution is based on the use of integral parameters that allow you to identify an anomaly, but do not allow you to indicate its cause. To work with such methods for determining the root causes of anomalies, auxiliary data is required, which is called metadata in the work.

The work describes a metadata construction algorithm and shows ways of using metadata to build an object in which the root cause of the anomaly is located. An approximate solution to the problem of finding the root cause of an anomaly with a help of quickly computable values of integral parameters is necessary to reduce the time of interruption of work processes due to implicit anomalies. It is assumed that small subsystems and nodes are easier to replace than to delve into the study of the cause.

INTRODUCTION

The increase in the size of distributed information and computing systems (DICS) exacerbates the problem of timely detection of failures, errors and shutdowns due to conflicts with information security (IS) requirements (hereinafter, anomalies). Implicit anomalies do not demonstrate their damage immediately, and Root Cause Analysis (RCA) is required to analyze and restore them. The problem of a remote search of implicit anomalies root causes is often a complicating factor. Modern operating systems can collect a large amount of data for RCA. However, remote analysis requires to transfer this data to the system administrator or IS officer (Grusho et al., 2020a), but most of this data will not be in demand, and their transfer over communication channels is not cheap and is not always possible. Big data requires more computing resources. RCA acceleration is often achieved by successful localizing of an anomalous DICS subsystem, which allows you to replace this subsystem and reduce the time to restore workflows.

Such a search requires certain auxiliary data, the collection and the organization of which depends on the speed and the effectiveness of the finding of a minimum possible object containing the root cause of the anomaly sought.

In the paper these supporting data are called metadata (MD). This name was chosen due to the fact that this term has already been used to solve such problems (Grusho et al., 2020b). The basis of MD is formed by a special case of Knowledge Graphs in which the arc from the object O_1 going to the object O_2, means that the anomaly in O_1 can initiate anomaly in O_2.

Arbitrary technical devices, network infrastructure, software and information resources will be called elements of the DICS. Arbitrary sets of elements are called DICS objects. Anomalies can occur in case of errors in data and data transformations, in case of technical breakdowns, as well as in case of unacceptable interactions of DICS objects. Anomalies can manifest themselves in the improper development of computing processes, obtaining incorrect calculation results. The most obvious manifestation of the anomaly is the unexpected shutdown of the computing process, the failure of the technical devices of the DICS. However, in all cases, implicit anomalies are possible, the location of which and the causes of their occurrence require a deep analysis (Grusho et al., 2018).

The model of implicit anomalies and search of root causes using integral parameters (IP) is constructed in (Grusho et al., 2020c). Such parameters are determined on the basis of generating algorithms (GA) that create the values of these parameters, and effective algorithms that display the values of these parameters. The RCA procedure proposed in (Grusho et al., 2020c) creates a DICS object containing a damaged GA fragment, the IP of which showed the presence of an anomaly.

Close problems arise when finding vulnerabilities in software using intelligent fuzzing (Jurn et al., 2019). Fuzzing is a technique for generating input values suitable for generating an error through target software analysis (Bekrar et al., 2012). Smart fuzzing has the advantage of knowing where errors can occur through software analysis. The tester can create test cases for that branch to extend the coverage of the code and generate valid conflicts. However, analyzing the target soft-

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

ware requires expert knowledge and takes a long time to generate a template suitable for software input.

Thus, it is recognized that the most difficult problems, equivalent to RCA, require a lot of human labor. In this paper the reduction of labor is achieved by using the anomaly's IP which is used in search for the root cause.

PROPERTIES OF GENERATING ALGORITHMS

Further there are often used two terms: an algorithm and a process. The complete definition of the term "an algorithm" can be found in (Uspensky and Semenov, 1987)[7], and process concepts in a computer system can be found in (Hoare, 1985)[8]. In this paper, the links of these concepts are used. The process implements some algorithm or fragment of the algorithm. The algorithm in the computer system is implemented using one or more processes.

It is said that the algorithm passes through the object O if processes as elements of O are involved in its implementation.

The generating algorithm (GA) of the parameter I is the algorithm for generating and calculating values of parameter I (Grusho et al., 2020c), further will be denoted through GA(I). If I is an integral parameter, the values of I can be calculated using an algorithm independent of GA, but giving the same value as GA. In the case of an anomaly, both algorithms show the anomaly. At the same time, only GA is involved in the formation of an anomaly that both algorithms are able to identify. In (Grusho et al., 2020c) it is shown that there can be several GA for the parameter I. In this case it is necessary to determine the influence of several GA on the values of the parameter I. However, in the assumption of the presence of an anomaly, we do not take into account the influence of several GA on reducing the capability of the detection of the anomaly, and in terms of increasing for the detection capability of the anomaly with the help of I, it is not necessary to take into account the influence of other GA. Therefore, instead of considering the effect on the value I of several algorithms, we can talk about GA that have failures of fragments of algorithms or do not have them. It has been proved (Grusho et al., 2020c) that the uniqueness of the root cause of the anomaly is a sufficient condition for the manifestation of the anomaly in the integral parameter I, when the anomaly is generated by the failed fragment of the GA(I).

Consider a few questions about the content of DICS facilities. Since it is not always possible to enumerate all DICS elements included in the object O, inductive definitions should be used. Let the objects $O_1, ..., O_k$ be defined. Then the theoretical-plural union O of objects $O_1, ..., O_k$ is also an object. At the same time, these objects as sets of elements may not intersect, but may interact.

We define the interaction of objects O_1 and O_2 as the presence of processes ξ_1 in the object O_1 and ξ_2 in the object O_2 such that ξ_1 can at a given time of its implementation transmit information about its state to the process ξ_2 over a certain channel. In this case, the process ξ_2 is able to receive this information and use it in its algorithm (Hoare, 1985).

However, the anomaly in GA(I) is not always the root cause of the anomaly. Acceptable localization of the anomaly with the help of integral parameters is possible when the source of the anomaly is some unobserved process mediated by interacting with GA(I), through which the anomaly is detected. It is said that in the ξ_1 process, the anomaly error is spread to GA(I) if there are a number of interactions ξ_1 with ξ_2,..., ξ_k with ξ_{k+1}, where the last process belongs to GA(I). Hence, an abnormal fragment GA(I)(t) of GA(I) arises, where t is the value in a certain enumeration of fragments of GA(I). If in the process of ξ_1 is the root cause of the anomaly and it is the only one in the computer system, then (Grusho et al., 2020c) it can be proved that the anomalous fragment of GA allows you to see the anomaly in I. The process of such an anomalous transmission can be represented in the form of the following graph, which we will call the attachment.

$$GA(I)(t-1) \to GA(I)(t) \to$$

$$\uparrow$$

$$\xi_k$$

If you allow an attachment operation, you must define a branch operation. GA(I)(t) having an anomaly, when interacting with some ξ process, may initiate an anomaly in ξ.

$$GA(I)(t-1) \to GA(I)(t) \to GA(I)(t+1)$$

$$\downarrow$$

$$\xi$$

In this case as ξ can be GA fragment of another integral parameter I^*, so GA(I^*) will show an anomaly (with the only root cause of the anomaly, this will happen necessarily).

Let $O_1 \to O_2$ be objects in which ξ_1 can transmit to ξ_2 anomaly. Let O_{11}, O_{12} be partition of the object O_1 into two nonintersect objects, and O_{21}, O_{22} be a partition of the object O_2 into two nonintersect objects. The condition $O_1 \to O_2$ means that there are at least a pair of objects from different partitions for which it is carried out, for example, $O_{12} \to O_{21}$. Indeed, the process is an indivisible entity. Then if ξ_1 is in O_{12} and ξ_2 is in O_{21} then $O_{12} \to O_{21}$. Back, if $O_{12} \to O_{21}$, then in the object of merge we get $O_1 \to O_2$. If objects O_1 and O_2 intersect and ξ_1 is the common process, then when dividing each of them into two objects, there is a pair where ξ_1 will be in, and for this pair it remains possible to transmit an anomaly. From here we get the following statement.

Statement 1. Objects O_1 and O_2 satisfy the condition $O_1 \to O_2$ if and only if there are subsets of O_1^* in O_1 and O_2^* in O_2 such that $O_1^* \to O_2^*$.

The relation $\to$ is reflexive, transitive and antisymmetric, that is, it is a partial order.

For RCA using the integral parameter I for GA(I), the following important property is required. Let the objects $O_1, ..., O_k$ in MD can transfer anomaly to the object O. Then $O_1, ..., O_k$ satisfy the requirement of completeness for O if for parameter I provided that GA(I) passes through O, there is m, $m = 1, ...k$, such that GA(I) passes through O_m.

Requirements of completeness for O and $O_1, ..., O_k$ are difficult to verify. The following statement simplifies the situation somewhat.

Statement 2. The objects $O_1, ..., O_k$ satisfy the condition of completeness with respect to the object O for GA(I) if and only if the union $O_1, ..., O_k$ is the complete object with respect to the object O for GA(I).

Proof. The necessity follows from the following simple reasoning. If GA(I) passes through O_m, then it passes through the union $O_1, ..., O_k$.

Sufficiency. If GA(I) passes through the union $O_1, ..., O_k$, then there exists a fragment of this algorithm passing through this union. Each fragment of GA(I) is implemented by one or more processes $\xi_1, ..., \xi_l$, which are indivisible. The totality of these processes interacts in this order with the union $O_1, ..., O_k$. If there is an anomaly in fragment GA(I), then it must be at least in some ξ_n. If the indivisible process ξ_n passes through the union $O_1, ..., O_k$, then as an algorithm it passes through some set of this union. So the fragment GA(I) of ξ_n passes through at least one of the objects $O_1, ..., O_k$. The statement is proved.

Consequence. This statement is true for any parameter I and its GA(I).

Considering large disjoint components of the DICS, for which completeness is easy to check, we will transfer completeness to their partitions, which will automatically follow from Statement 2. Further, it will be assumed that for the considered sets of MD the completeness conditions are met.

If GA(I) passes through an object O in which interaction with an anomalous process ξ occurs, then it is easy to propose a simple probabilistic model for the propagation of the anomaly. Let p be the probability that GA(I) will receive an anomaly from ξ. If there are n independent interactions of GA(I) with abnormal processes in O, then the probability of an anomaly in GA(I) is $1-(1-p)^n$. It is interesting to look at this formula in various extreme cases of the values of these parameters.

If $p = 1$, then the propagation of the anomaly in interactions towards O occurs with probability 1. If error propagation occurs with probability 1, then the widespread propagation of the anomaly poses the greatest difficulties for RCA. An example of such an error is given in (Grusho et al., 2018, 2017).

If p is close to or equal to 0, then the probability of propagation of the anomaly is small, but it is easier to localize the cause of the anomaly (if such an anomaly can be seen). An example of such an anomaly is the prefailed state of the hard drive when it is still working, but an anomaly in its device will inevitably cause it to fail. It is possible to see such an anomaly by indirect signs, in particular, by a significant increase in the time of access to the disk (Grusho et al., 2020c).

Further, it is believed that the DICS propagates errors with $p = 1$. In order to effectively use the MD to search for root causes, it is necessary to construct a hierarchical decomposition of the MD and use it to accelerate algorithms for searching for objects containing an anomaly.

CONSTRUCTION AND USAGE OF MD

In the introduction, it was noted that MD are auxiliary information for a system administrator or an IS officer, which is built on the basis of knowledge graphs (Brandón et al., 2020; Nickel et al., 2016) of the form $X \to^c Y$, where X and Y are objects of the DICS, c is the action that is transmitted from X to Y. In our case, the hierarchy of the MD is built sequently. Let the object X may pass an anomaly to the object Y. This binary relationship was entered earlier and is denoted by $X \to Y$. As shown above, this relationship generates a partial order on a subset of objects. The largest element of this order is the DICS. Assuming the completeness of the partitions of objects and the finiteness of the tree representing the sequences of such partitions, we will build the MD of three parts.

1. The set of objects with the help of successive partitions into meaningful nodes and technologies and their hierarchy are formed from structural model of DICS (Denisov and Kolesnikov, 1982). They are used to search for objects that cover the anomaly using IP and object relationships derived from the structural model.

2. The set of integral parameters IP is defined in the constructed set of objects (where this can be done).

3. From IP, the completeness properties of object partitions and their relationships $O \to O"$, we build the MD as trees whose roots coincide with objects containing IP, and the arrows are directed to the roots.

Consider the issue of constructing integral parameters that lie in the objects constructed in item 1. Let O be an object (node, device, data conversion information technology, information resource, communication fragment, program, etc.) – this is a set of elements of the DICS. Each DICS element is described by a plurality of characteristics (Ashby, 1956). Each characteristic is described by one or more parameters and areas of their normal values (Ashby, 1962). That is, any system can be described by a plurality of its parameters (Ashby, 1962; Grusho et al., 2016). An anomaly in a parameter value is the appearance of a value that goes beyond normal values. Since the function of belonging to a set of normal values of a parameter can be calculated without knowledge of the GA of this parameter, then formally any parameter can be integral. However, in practice, not all system parameters can be seen and learned their values in the real system, then only a few set of integral parameters should be found that cover the most important subsystems of the DICS if possible. At the same time, the importance lies not only in the allocation of

risk objects, but also in the set of interactions of the selected integral parameters based on observations of the system. Therefore, it is necessary to build IP with an orientation on the following principles.

If IP could not be allocated in the desired object, then it is necessary to divide this object into subobjects before IP appears in each branch.

If it is not possible to achieve completeness in the selection of subobjects, then you need to supplement more subobjects before splitting.

If two or more IP are selected in the object, the object must be divided into subobjects so that horizontal links are converted to vertical or simply one of IP remains in each subobject.

From the set of IP, from objects which contain these IP, and from their relations $O_1 \rightarrow O_2$, we build influence trees. If IP is contained in the object O, then the influence tree has the root O, and the arcs of all objects are oriented to O. If an arc emerges from the object of the tree to a lower object of the tree, then this object can affect objects that are located deeper in this tree, then this object can be met again in the tree.

Consider the usage of MD for RCA.

1. IP allows you to identify an anomaly in objects of the constructed object hierarchy.

2. The search of the following object with the root cause is based on the search of the IP with an anomaly value on the branch of the corresponding MD tree.

3. If IP is detected on a branch of the tree, the tree generated by the object with this IP is considered and iteration is repeated. The lowest (from the root of the tree) object with an anomaly determines the coverage of the anomaly in the created diagram of GA. The anomaly may not be covered only when the root cause is in an object through which GA of the last identified IP with the anomaly value does not pass. Therefore, the last object with the anomaly must be extended to merge all the objects from which the arrows enter the object with the anomaly (including closure).

The effectiveness of RCA by the considered method is evaluated further.

1. When the condition of completeness is fulfilled, all potential carriers of the anomaly are covered by the last built object with the anomaly.

2. The coverage of the anomaly is determined by the object with the last detected IP with the abnormal value, provided that the root cause is unique.

3. The usage of last identified IP after objects located in the tree that do not have IP is based on the transitivity of the $\rightarrow$ relations.

CONCLUSION

The use of integral parameters in RCA allowed to get a new look at the problem of causality analysis. The idea of RCA based on IP is presented for the first time in the work (Grusho et al., 2020c). In this work, the regions covering the anomaly are more clearly defined and the organization of the creation and use of MD by the system administrator and the IS officer has been developed. Note that in remote cause analysis, search by IP with an abnormal values reduces the amount of information transmitted.

In addition, the interactions between IP and objects of DICS were investigated, which in fact can give an answer to the question about the coverage of the anomaly. A hierarchical organization of MD has been determined, which allows formalizing the algorithm for finding anomaly coverage using IP. It is shown under what conditions a good coverage of the anomaly is achieved using objects containing IP and when additional information needs to be collected and used to reduce the region of coverage of the anomaly.

Unfortunately, all experiments on the practical application of the proposed approach were carried out manually by organizing the actions of the system administrator according to the constructed algorithm. This does not mean that it is impossible to automate the collection, storage and application of the IP method in RCA. However, the authors are confident that the automated system should be built on the basis of algorithms that maximize usage of the experience of system administrators and IS officers. It should be noted that during the research, the initial vision of the problem has changed significantly.

Further research will be related to the construction of algorithms for big data in real large DICS.

Acknowledgements

This work was partially supported by the Russian Foundation for Basic Research (grant No. 18-29-03081).

REFERENCES

Grusho, N. A., A. A. Grusho, M. I. Zabezhailo, and E. E. Timonina. 2020. "Methods of finding the causes of information technology failures by means of metadata". *Informatics and applications* 14, No. 2, 33–39.

Grusho, N. A., A. A. Grusho, and E. E. Timonina. 2020. "Localizing failures with metadata". *Automatic Control and Computer Sciences* 54, No. 8, 988–992.

Grusho, A. A., M. I. Zabezhailo, A. A. Zatsarinny, A. V. Nikolaev, V. O. Piskovski, V. V. Senchilo, I. V. Sudarikov, and E. E. Timonina. 2018. "About the Analysis of Erratic Statuses in the Distributed Computing Systems". *Systems and Means of Informatics* 28, No. 1, 99–109.

Grusho, A.A., M.I. Zabezhailo, A.A. Zatsarinny, A.V. Nikolaev, V.O. Piskovski, V.V. Senchilo, and E.E. Timonina. 2017. "Erroneous states classifications in dictributed computing systems and sources of their occurences". *Systems and Means of Informatics* 27, No 2, 29–40.

Grusho, A. A., N. A. Grusho, M. I. Zabezhailo, and E. E. Timonina. 2020. "Root Cause Anomaly Localization". *Information Security Problems. Computer Systems* 4 (in press).

Jurn, J., T. Kim, and H. Kim. 2019. "A Survey of Automated Root Cause Analysis of Software Vulnerability". In: *Innovative Mobile and Internet Services in Ubiquitous Computing*, L. Barolli, F. Xhafa, N. Javaid, and T. Enokido (Eds), Advances in Intelligent Systems and Computing 773. Springer, Cham, 756–761.

Bekrar, S., C. Bekrar, R. Groz, and L. Mounier. 2012. "A taint based approach for smart fuzzing". In: *2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, Montreal, QC, 818–825.

Uspensky, V.A., and A.L. Semenov. 1987. *Theory of algorithms: the main discoveries and applications*, Moscow: Science, 288 p. (in Russian).

Hoare, C. A. R. 1985. *Comucating Sequential Processes*, Englewood Cliffs (N.J.): Prentice-Hall, 256 p.

Brandón, Álvaro, Marc Solé, Alberto Huélamo, David Solans, María S. Pérez, Victor Muntés-Mulero. 2020. "Graph-based root cause analysis for service-oriented and microservice architectures". *Journal of Systems and Software* 159, 1–17.

Nickel, M., K. Murphy, V. Tresp and E. Gabrilovich. 2016. "A Review of Relational Machine Learning for Knowledge Graphs". *Proceedings of the IEEE* 104, No. 1, 11–33.

Denisov, A.A., and D.N. Kolesnikov. 1982. *Theory of large control systems: Textbook for universities*, Leningrad: Energoizdat, 288 p. (in Russian).

Ashby, W. Ross. 1956. *An Introduction to Cybernetics*, London: Chapman and Hall, 295 p.

Ashby, W. Ross. 1962. *Design for a Brain. The Origin of Adaptive Behavior*, 2nd Ed. Revised.Moscow: Foreign literature, (Russian translation).

Grusho, A., N. Grusho, and E. Timonina. 2016. "Detection of anomalies in non-numerical data". In: *2016 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops* (ICUMT), Lisbon, 273–276.

AUTHOR BIOGRAPHIES

ALEXANDER A. GRUSHO,Professor (1993), Doctor of Science in physics and mathematics (1990). He is principal scientist at Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences 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`.

NICK A. GRUSHO has graduated from the Moscow Technical University. He is Candidate of Science (PhD) in physics and mathematics. At present he works as senior scientist at 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, simulation theory and practice, computer sciences.

His email is `info@itake.ru`.

MICHAEL I. ZABEZHAILO has graduated from the Institute of Physics and Technology and gained the Candidate degree (PhD) in theoretical computer science (1983). He is Doctor of Science in physics and mathematics (2016). Now he works as Head of laboratory in Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences.

Research interests: mathematical foundations of artificial intelligence, reasoning modeling, information security, theoretical computer sciences.

His email is: `zabezhailo@yandex.ru`.

VLADIMIR V. SENCHILO has graduated from the Moscow Institute of Physics and Technology. He is scientist at Institute of Informatics Problems, Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences.

Research interests: computer sciences, machine learning, optimization theory, data mining, financial risk.

His e-mail address is: `volodias@mail.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`.

MINIMIZING MEAN RESPONSE TIME IN BATCH-ARRIVAL NON-OBSERVABLE SYSTEMS WITH SINGLE-SERVER FIFO QUEUES OPERATING IN PARALLEL

Mikhail Konovalov and Rostislav Razumchik
Institute of Informatics Problems of the FRC CSC RAS
Vavilova, 44-2, 119333, Moscow, Russia
Email: mkonovalov@ipiran.ru, rrazumchik@ipiran.ru

KEYWORDS

dispatching, unobservable, mean response time, batch arrivals, load balancing, unreliable servers

ABSTRACT

Consideration is given to dispatching systems, where jobs, arriving in batches, cannot be stored and thus must be immediately routed to single-server FIFO queues operating in parallel. The dispatcher can memorize its routing decisions, but at any time instant does not have any system's state information. The only information available is the batch/job size and inter-arrival time distributions, and the servers' service rates. Under these conditions, one is interested in the routing policies which minimize the job's long-run mean response time. The single-parameter routing policy is being proposed which, according to the numerical experiments, outperforms best routing rules known by now for non-observable dispatching systems: probabilistic and deterministic. Both batch- and job-wise assignments are studied. Extension to systems with unreliable servers is discussed in short.

INTRODUCTION

Efficient resource allocation is the typical problem faced by system designers in various fields of study (transportation, distributed/parallel computing, customs inspection etc.). The particular problem studied in this paper has its roots in the volunteer computing. Consider a system in which jobs (of a single class) arrive according to some stochastic process in batches and have to be assigned to one of the single-server queues immediately upon the arrival. Scheduling in each queue is FIFO. The dispatcher, which performs this operation (see Fig. 1), can memorize its routing decisions, but has no online information about the system except for: the cumulative distribution function (CDF) of inter-arrival times, batch size and job size CDFs and servers' service rates. Thus for the dispatcher the system is non-observable (see, for example, Anselmi and Gaujal (2011); Lingenbrink and Krishnamurthy (2017)). The task is to find the routing policy, which minimizes the job's long-run mean sojourn time in the system[1]. Since the dispatcher may decide to split batches, when making decisions, both cases — batch- and job-wise assignments — need to be considered.

[1] Or, equivalently, the system's mean response time.

The described system is closely related to the basic dispatching problem studied extensively in the literature (Harchol-Balter, Crovella and Murta (1999); Hyytiä and Aalto (2013); Feng, Misra and Rubenstein (2005)). But, due to the absence of any online information, from the long list of routing rules, only the following two naive (in terminology of (Mengistu and Che, 2019, Section 2.3)) policies are feasible: probabilistic and deterministic. According to the numerical experiments (see Konovalov and Razumchik (2020)), both these rules can be outperformed by the algorithm, which implements the so-called arrival-aware policy (see Konovalov and Razumchik (2018)) for single-arrival non-observable systems. In this paper numerical evidence is given that all the three policies (probabilistic, deterministic and arrival-aware) can be applied by the dispatcher in a batch-arrival system. We rank the policies with respect to the minimal mean response time (and its standard deviation) and discuss extensions of the system as well as of the arrival-aware policy.

The paper is organized as follows. In the next section the detailed description of system is given. The third section contains the overview of the routing policies available to the dispatcher in the considered setting. The numerical examples, which follow, demonstrate the performance of the policies. In the Section 5 the study is extended to systems with unreliable servers. The main conclusions are briefly summarized in Section 6.

SYSTEM DESCRIPTION AND THE PROBLEM STATEMENT

The system considered in the paper is illustrated in Fig. 1.

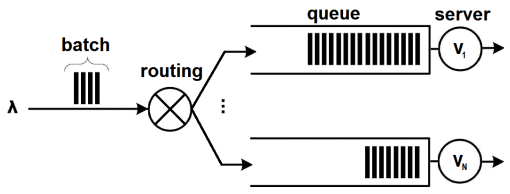


Figure 1: Jobs of a single class arrive in batches and are immediately routed to one of the queues, where they wait in FIFO order for service.

It consists of $N \geq 2$ single server infinite capacity queues, operating in parallel. The queues are numbered from

Communications of the ECMS, Volume 35, Issue 1,
Proceedings, ©ECMS Khalid Al-Begain, Mauro Iacono,
Lelio Campanile, Andrzej Bargiela (Editors)
ISBN: 978-3-937436-72-2 / 78-3-937436-73-9(CD) ISSN 2522-2414

1 to N and the server's service rate of queue m is $v^{(m)} > 0$, $1 \le m \le N$. Jobs (of a single class) arrive to the system in batches and the inter-arrival times are independent and identically distributed (i.i.d.) random variables with the known CDF $A(x)$ and the mean λ^{-1}. Consecutive batches consist[2] of random numbers $B_1, B_2, \dots$ of jobs having the known distribution[3] $\{g_i, i = 1, 2, \dots\}$, with the mean $\mathsf{E}[B]$ and the variance $\mathsf{Var}[B]$. The sizes $X_1, X_2, \dots$ of individual jobs are i.i.d. random variables with the known CDF $S(x)$, mean $\mathsf{E}[X]$ and variance $\mathsf{Var}[X]$. Jobs are served one-by-one and the service discipline employed in each queue is FIFO. Service pre-emption and jockeying between queues is not allowed.

The dispatcher operates either in the batch-wise or in the job-wise mode. The former means that it treats a batch as if it were a single (macro-) job and makes a single routing decision (i.e. all the jobs in the batch are assigned to the same queue). In the latter case the dispatcher splits batches and chooses a queue for each job individually. Switching between the modes is not allowed. Upon receiving a batch the dispatcher must immediately[4] make a routing decision.

Fix an arbitrary integer $n \ge 1$. Let $0 \le t_1 < \dots < t_n$ and $b_1, b_2, \dots, b_n$ denote the arrival instants and the sizes of the first n batches, correspondingly. Let $y_1, y_2, \dots, y_k$ be the sequence of routing decisions[5] made so far at the instants $t_1, t_2, \dots, t_{n-1}$. Each y_j takes a value from the set $\{1, 2, \dots, N\}$. In order to make routing decisions at the instant t_n, the dispatcher may use only the following information: the values $t_1, t_2, \dots, t_n$ and $b_1, b_2, \dots, b_n$; the values $y_1, y_2, \dots, y_k$; the distributions $A(x)$, $S(x)$ and $\{g_i, i = 1, 2, \dots\}$; the service rates $v^{(1)}, v^{(2)}, \dots, v^{(N)}$ in each queue. No online information (like the arriving job size, current queues' sizes etc.) is available. The dispatcher's task is to minimize job's long-run mean sojourn time in the system.

OVERVIEW OF THE AVAILABLE DISPATCHING POLICIES

The list of dispatching policies, which are available to the dispatcher under the assumptions made (i.e. in the absence of any system's state information), is very short: probabilistic (see, for example, Bell and Stidham (1983)), deterministic (see Hordijk and van der Laan (2004)) and arrival-aware (see Konovalov and Razumchik (2020)).

According to a probabilistic routing policy (further referred to as RND) a job is routed to the queue n with the probability p_n independently of the previous decisions. Recall that in the batch-wise mode the dispatcher assigns the same queue for each job within a batch. Let the arrival flow of batches be Poisson with the rate λ. Then the n^{th} queue is the $M^{[X]}/GI/1$-FIFO queue with the arrival rate λp_n and, whenever it is stable, the mean response time $\mathsf{E}[T_n]^{\mathrm{FIFO}}$ of an arbitrary job is equal to (see (Cooper, 1981, p. 241))

$$\mathsf{E}[T_n]^{\mathrm{FIFO}} = \frac{(\mathsf{E}[B]+1)\mathsf{E}[X^{(n)}]}{2} + \frac{\mathsf{Var}[B]\mathsf{E}[X^{(n)}]}{2\mathsf{E}[B]} + \\ + \frac{\lambda p_n \mathsf{E}[B]^2 \mathsf{E}[X^{(n)}]^2 \left(1 + \frac{\mathsf{E}[B]\mathsf{Var}[X^{(n)}] + \mathsf{E}[X^{(n)}]^2 \mathsf{Var}[B]}{\mathsf{E}[B]^2 \mathsf{E}[X^{(n)}]^2}\right)}{2(1 - \lambda p_n \mathsf{E}[B]\mathsf{E}[X^{(n)}])}, \quad (1)$$

where the notation $X^{(n)} = X/v^{(n)}$ is used. The optimal probabilistic routing policy, further referred to as RND-opt, is the probability distribution $(p_1, p_2, \dots, p_N)$ that minimizes the mean response time given by

$$\sum_{n=1}^{N} p_n \mathsf{E}[T_n]^{\mathrm{FIFO}} \quad (2)$$

under the constraint $0 \le \lambda p_n \mathsf{E}[X^{(n)}]\mathsf{E}[B] < 1$ for each n. This problem can be approached either analytically[6] or numerically and, in the latter case, it can be solved at a satisfactory level. If the batch-arrival process is not Poisson, but a general renewal process with the mean λ^{-1} and SCV (squared coefficient of variation) equal to C_A^2, then the sequence of arrival instants to the queue n constitutes the renewal process with the mean $(\lambda p_n)^{-1}$ and the SCV equal to $1 + (C_A^2 - 1)p_n$. But (1) is not valid any more and we are unaware of any feasible way to compute numerically both $\mathsf{E}[T_n]^{\mathrm{FIFO}}$ and the optimal N-tuple $(p_1, p_2, \dots, p_N)$. For small values of N simulation can be used to find the approximate solution of the minimization problem. In general, load balancing (i.e. $p_n = v^{(n)} / \sum_{i=1}^{N} v^{(i)}$) seems to be the only reasonable trade-off, but in most cases it is far from being optimal. When the dispatcher operates in the job-wise mode, a queue is assigned for each job individually according to the given probability distribution $(p_1, p_2, \dots, p_N)$. But even if the batches arrive according to a Poisson flow, the optimal N-tuple is not the solution of (2). For a batch of size k the probability that k_1 jobs are routed to the queue 1, ..., k_N jobs — to the queue N, is equal to

$$\frac{k!}{k_1! k_2! \dots k_N!} (p_1)^{k_1} \dots (p_N)^{k_N}, \quad (3)$$

where $\sum_{i=1}^{N} k_i = k$, $0 \le k_i \le N$. Since the consecutive batches are treated independently, for the systems with Poisson arrivals and exponential job size distributions (3) can be used to write out the balance equations[7], which can be solved numerically (after truncation of the state space). Once the joint stationary distribution of

[2] No precedence constraints are imposed on jobs within a batch.

[3] It is worth noticing that in batch-arrival queues the limiting distributions of some quantities may not exist without additional restrictions on the batch-size distribution $\{g_i\}$ (see, for example, (van Ommeren, 1990, p. 679)).

[4] I.e. it does not have a room for storing the jobs.

[5] The total number k of decisions before the instant t_n depends on the operation mode of the dispatcher.

[6] See, for example, traffic/resource allocation problems in Ibaraki and Katoh (1988).

[7] For example, for $N = 2$ the evolution of the system's content can be described by the QBD process. Assuming the capacity of either queue to be finite, the generator is then of $M/G/1$-type.

queue-sizes' is found, standard argument can be used to find (approximately) the system's mean response time. This procedure theoretically can be used to search for the (close to) optimal N-tuple $(p_1, p_2, \ldots, p_N)$. But in practice (especially in heavy traffic) the computational complexity becomes prohibitive and, as well as in the case of a general renewal arrival process, one has to resort to simulation.

According to a deterministic policy jobs are dispatched according to the infinite sequence $a_1, \ldots, a_n, \ldots$, where a_i is the queue number, whereto the i^{th} job is routed[8]. Finding optimal deterministic sequence for N single server queues, operating in parallel with single arrivals, is a difficult problem for which no general procedures exist; obviously the same holds for batch-arrival systems. Yet when jobs arrive to the dispatcher not in the batches but one at a time, very good results can be achieved by special deterministic sequences – billiard sequences[9], which can be constructed using greedy algorithms. One of such algorithms further referred to as SG (Special Greedy) is due to (Hordijk and van der Laan, 2004, p.184). According to the SG policy the i^{th} job is routed to the queue a_i:

$$a_i = \text{argmin}_{1\le n\le N}\left(\frac{x_n + \kappa^n(i-1)}{d_n}\right), \qquad (4)$$

where $\kappa^n(i)$ is equal to the the number of jobs (among the first i jobs) sent to the queue n so far, d_n is the fraction of jobs, which has to be routed to the queue n, and x_n are properly chosen non-negative rational numbers[10]. Finding the optimal densities $(d_1, d_2, \ldots, d_N)$ is an open problem and, in general, can be approached only using simulation. As a rule of thumb one can use in (4) instead of $(d_1, d_2, \ldots, d_N)$ the N-tuple $(p_1, p_2, \ldots, p_N)$ computed for the probabilistic routing policy. Even though such choice usually results not in the optimal solution, deterministic routing is more efficient[11] than the probabilistic routing. We are unaware of any deterministic policy developed particularly for batch-arrival systems. And since all the quantities in (4) remain properly defined for both batch-wise and job-wise routing, we choose (4) as the basic deterministic dispatching policy.

The third class of routing policies further referred to as AA (Arrival-Aware) is based on the following intuitive idea (see Konovalov and Razumchik (2018)): longer inter-arrival times increase the probability of those system's states, which correspond to lower workloads in queues and vice versa. It is not straightforward to implement this idea in a batch-arrival system[12]. But in Konovalov and Razumchik (2020) for single-arrival unobservable systems it was suggested to put the idea into practice by means of the algorithm[13], which is reproduced below. Fix the positive real $c > 0$. Let us associate with the i-th job arriving at the dispatcher, N numbers, say $u_i^{(1)}, \ldots, u_i^{(N)}$, which are defined recursively as follows:

$$\tilde{u}_i^{(n)} = \max\left(0, u_{i-1}^{(n)} - (t_i - t_{i-1})\right),\ 1 \le n \le N,\ i \ge 1,$$

$$u_i^{(n)} = \begin{cases} \tilde{u}_i^{(\tilde{y}_i)} + \frac{c}{v^{(\tilde{y}_i)}}, & \text{if } n = \tilde{y}_i, \\ \tilde{u}_i^{(n)}, & \text{otherwise,} \end{cases}$$

where $u_0^{(1)} = b_1, \ldots, u_0^{(N)} = b_N$ and[14]

$$\tilde{y}_i = \text{argmin}_{1\le n\le N}\left(\tilde{u}_i^{(n)} + \frac{c}{v^{(n)}}\right).$$

The AA policy prescribes to route the i^{th} job to the queue $y_i = \tilde{y}_i$. The pseudocode for the policy is given below (see Algorithm 1). Unlike the RND and SG policies, the AA policy depends only on a single parameter, which, in general, must be estimated using simulation. All the quantities in the Algorithm 1 remain[15] properly defined for batch-arrival systems as well. Thus we choose it as the basic AA policy.

Algorithm 1 Pseudocode of the AA policy

function NextDecision(N,$v^{(1)}$,…,$v^{(N)}$,$u_{i-1}^{(1)}$,…,$u_{i-1}^{(N)}$,t_i, t_{i-1},c)
  for $n = 1 \to N$ **do**
    $u_i^{(n)} = \max\left(0, u_{i-1}^{(n)} - (t_i - t_{i-1})\right)$
  end for
  $y_i = \text{argmin}_{1\le n\le N}\left(u_i^{(n)} + c/v^{(n)}\right)$
  $u_i^{(y_i)} = u_i^{(y_i)} + c/v^{(y_i)}$
  return $y_i, u_i^{(1)}, \ldots, u_i^{(N)}$
end function

[a] The function NextDecision(N,$v^{(1)}, \ldots, v^{(N)}, u_{i-1}^{(1)}, \ldots, u_{i-1}^{(N)}, t_i, t_{i-1}$) returns for the i^{th} job the routing decision y_i based on the i^{th} job arrival instant t_i and the arrival instant t_{i-1} of the $(i-1)^{th}$ job, servers' speeds $v^{(1)}, \ldots, v^{(N)}$, auxiliary values $u_{i-1}^{(1)}, \ldots, u_{i-1}^{(N)}$ and c.
[b] The values $u_0^{(1)}, \ldots, u_0^{(N)}$ are the initial remaining workloads (including server). For the initially empty system $u_0^{(1)} = \cdots = u_0^{(N)} = 0$.
[c] The positive real value c is the parameter of the algorithm, which must be set manually.

Numerical examples in the next section give some impression on how these dispatching policies are ranked,

[8] One of the well-known examples of such a policy is the Round-Robin policy. It rotates the jobs between the queues in the cyclic order and thus is applicable only in the systems with homogeneous servers (i.e. having equal service rates). For heterogeneous systems some generalizations do exist (see Arian and Levy (1992)).

[9] Although other deterministic policies besides (4) do exist (see, for example, GRR, CGRR, and mBS policies in Arian and Levy (1992); Sano and Miyoshi (2000)), according to our experience with scheduling problems in non-observable queues (see, for example, Konovalov and Razumchik (2018, 2020)), (4) shows the best performance.

[10] For example, one can put $x_n = 1$ if the queue n has the fastest server and $x_n = 0$ otherwise.

[11] Some numerical and analytic evidences are given, for example, in Anselmi (2017).

[12] We conjecture that the algorithm of Konovalov and Razumchik (2018) can be adopted to the considered system at least in the case when the dispatcher operates in the batch-wise mode, i.e. assigns the whole batch to the same server.

[13] See also the footnote 9 on page 400 in Konovalov and Razumchik (2020).

[14] When argmin() is being evaluated, ties are broken in the favour of the fastest server and randomly between the fastest servers.

[15] The presented version of the AA policy (Algorithm 1) is tuned for the batch-wise assignment. In case of job-wise assignment the Algorithm 1 has to be applied to each job in the batch.

when one seeks to minimize the mean (and even the standard deviation of the) sojourn time of an arbitrary job as well as of a whole batch.

NUMERICAL EXPERIMENT

In the first example we consider an elementary system with the two servers processing jobs with the total service rate equal to 1. Let $v^{(1)} = 0,326$ and $v^{(2)} = 0,674$. Batches arrive according to the Poisson process with the rate λ. The batch size distribution is geometric with the mean $\mathsf{E}[B] = 3$ ($\mathsf{Var}[B] = 6$) and the job size distribution is exponential with the mean $\mathsf{E}[X] = 1$. The offered load to the whole system is thus $\rho = \lambda \mathsf{E}[B]\mathsf{E}[X]/\sum_{i=1}^{2} v^{(i)} = 3\lambda$. Since all the three considered policies (RND, SG and AA) can be applied with batch-wise and job-wise assignment, this yields 3 batch- and 3 job-specific policies. In the considered two-server fully exponential system the approximate values of the optimal parameters (p_1, p_2) for the RND policies are found numerically following the guidelines in the previous section. The (close to) optimal parameters (d_1, d_2) and c of the SG and the AA policies, respectively, are estimated through simulation. The numerical results (values of the mean and the standard deviation of the response time depending on the offered load ρ) for the batch-wise assignment are given in the Table 1 and for the job-wise assignment — in the Table 2.

The first observation from the data in the Tables 1 and 2 is that the SG policy and the AA policy, which were to be used for single-arrival systems, show meaningful results for the batch-arrival system as well. And this observation remained valid in all our numerical experiments. Next, both the SG and the AA policies give lower mean (and standard deviation of the) response time than the RND policy (with the optimal parameters' values). As the system's load increases it becomes less and less appealing to route the jobs/batches according to the RND-opt (even though it has the strongest theoretical support among the three). The performance of the SG and the AA policies is (roughly speaking) the same. Yet it can be noticed that for low and moderate system's load the AA policy is slightly (but persistently) better. The optimal values (p_1, p_2) and (d_1, d_2) are not equal (except for the case of high system's load): using (p_1, p_2) instead of (d_1, d_2) in the SG policy (4) leads to worse results.

When the dispatcher splits the batches (i.e. performs job-wise assignment), it may be important to route jobs in such a way so as to minimize the mean sojourn time of the whole batch[16]. The extent at which the RND, SG and AA policies cope with this task can be assessed from the Tables 2 and 3. By comparing the data it can be said that the results follow the intuition: the mean response time of an individual job is lower than of a whole batch. In can be also seen that the ranking of the policies remains unchanged.

If one ranks the policies, depending on the number of parameters requiring estimation, then in the first example all the policies are identical. The second example, considered further, is intended to bring in the difference in this issue. Consider the system with 128 servers processing jobs with the total service rate equal to 1. Servers are aggregated into 8 groups of equal size i.e. each of the 16 servers in the group number i, $1 \le i \le 8$, has the service rate $i/576$. Batches arrive according to the hyper-exponential process with SCV= 4, two phases and balanced means i.e. the phase probabilities are

$$\alpha_1 = \frac{1}{2}\left(1 + \sqrt{\frac{\mathrm{SCV}-1}{\mathrm{SCV}+1}}\right) \approx 0,8873, \quad \alpha_2 \approx 0,1127.$$

Thus the arrival rate is $\lambda = (0,8873/\lambda_1 + 0,1127/\lambda_2)^{-1}$. The values of λ_1 and λ_2, fixed to fit the chosen values of the system's load, are reported in the Tables 4 and 5. The job size distribution is chosen to be bimodal: a job has size either $0, 5$ or 9 with probabilities of $16/17$ and $1/17$ respectively. Thus the mean job size is $\mathsf{E}[X] = 1$ ($\mathsf{Var}[X] = 4$). The assumed batch size distribution is specified in the following table:

i	1	2	4	8
g_i	$0,375$	$0,125$	$0,125$	$0,375$

The mean batch size is $\mathsf{E}[B] = 4,125$ ($\mathsf{Var}[B] \approx 9,86$). The offered load to the whole system is thus $\rho = \lambda \mathsf{E}[B]\mathsf{E}[X]/\sum_{i=1}^{128} v^{(i)} = 4,125\lambda$. In order to determine the parameters of the RND and SG policies, we notice that in each group the service rates are identical; thus it is reasonable[17] to use Round-Robin routing within a group. Consequently for each of the two policies, RND and SG, one has to estimate 7 parameters. The values of $(p_2, \ldots, p_8)$ for the RND policy can be set so as to balance the load (this choice is further denoted by RND-lb) or they can be optimized through simulation (RND-opt, see Table 6). By analogy there are two options for the SG policy: SG-lb and SG-opt[18]. The AA policy requires a single parameter c, which is estimated through simulation. The values of the mean and the standard deviation of the response time depending on the offered load ρ for the batch-wise assignment are given in the Table 4 and for the job-wise assignment — in the Table 5.

Compared to the first example here all the random quantities (inter-arrival time, batch-size and job-size) are more variable. Yet qualitatively the results are similar. The SG-opt and AA policies always outperform (any of the two) the RND policies. With respect to both the mean response time and its standard deviation the AA policy is

[16] Under such a requirement, all jobs belonging to the same batch wait until the last member of the batch is processed, and thus the considered system becomes somewhat similar to a fork-join or a split-merge system.

[17] And such a rule is superior to random routing.

[18] We note that in the presented examples the parameters $(d_2, \ldots, d_8)$ of the SG-opt policy were, in fact, set equal to the parameters of the RND-opt, given in the Table 6. Thus the results under the SG policy can be potentially improved but not affecting the conclusions made.

Table 1: Job's mean (and the standard deviation of the) response time in the two-server system. Batch-wise assignment. Service rates: $\nu^{(1)} = 0,326$ and $\nu^{(2)} = 0,674$. Poisson batch-arrivals with the rate λ. Batch size is geometrically distributed with the mean $\mathsf{E}[B] = 3$. Job size distribution is exponential with the mean $\mathsf{E}[X] = 1$. The offered load is $\rho = 3\lambda$.

	$\lambda = 0,05$ $\rho = 0,15$	$\lambda = 0,1$ $\rho = 0,30$	$\lambda = 0,18$ $\rho = 0,54$	$\lambda = 0,2$ $\rho = 0,60$	$\lambda = 0,23$ $\rho = 0,69$	$\lambda = 0,25$ $\rho = 0,75$	$\lambda = 0,32$ $\rho = 0,96$
RND-opt	5,73 (5,74) $p_1 = 0$	7,68 (7,87) $p_1 = 0,129$	12,29 (12,69) $p_1 = 0,254$	14,22 (14,63) $p_1 = 0,269$	18,48 (18,97) $p_1 = 0,288$	22,99 (23,66) $p_1 = 0,297$	145,12 (145) $p_1 = 0,322$
SG-opt	5,725 (5,74) $d_1 = 0$	7,2 (7,35) $d_1 = 0,212$	10,42 (10,44) $d_1 = 0,27$	**11,81 (12,08)** $d_1 = 0,3$	**14,96 (15,07)** $d_1 = 0,3$	**18,4 (18,7)** $d_1 = 0,31$	**112 (110)** $d_1 = 0,322$
AA	**5,709 (5,74)** $c = 1$	**6,99 (7)** $c = 3,9$	**10,35 (10,58)** $c = 4,26$	**11,81 (12,2)** $c = 4,25$	15,08 (15,59) $c = 3,85$	18,5 (18,97) $c = 3,6$	112 (118) $c = 3,5$

Table 2: Job's mean (and the standard deviation of the) response time in the two-server system. Job-wise assignment. The input parameters are the same as in the Table 2.

	$\lambda = 0,05$ $\rho = 0,15$	$\lambda = 0,1$ $\rho = 0,30$	$\lambda = 0,18$ $\rho = 0,54$	$\lambda = 0,2$ $\rho = 0,60$	$\lambda = 0,23$ $\rho = 0,69$	$\lambda = 0,25$ $\rho = 0,75$	$\lambda = 0,32$ $\rho = 0,96$
RND-opt	4,64 (4,69) $p_1 = 0,25$	5,64 (5,66) $p_1 = 0,27$	8,59 (8,66) $p_1 = 0,3$	9,89 (10) $p_1 = 0,31$	12,8 (13,04) $p_1 = 0,315$	15,9 (16,12) $p_1 = 0,32$	100 (100) $p_1 = 0,322$
SG-opt	4,27 (4,24) $d_1 = 0,25$	5,11 (5,1) $d_1 = 0,29$	7,69 (7,68) $d_1 = 0,31$	8,83 (8,89) $d_1 = 0,32$	**11,4 (11,4)** $d_1 = 0,315$	14,1 (14,14) $d_1 = 0,31$	**88,6 (88,31)** $d_1 = 0,322$
AA	**4,17 (4,17)** $c = 0,75$	**5,05 (5,1)** $c = 1,2$	**7,66 (7,75)** $c = 1,25$	**8,79 (8,83)** $c = 1$	**11,4 (11,4)** $c = 1,1$	**14 (14,14)** $c = 1,15$	86,7 (86,6) $c = 1$

Table 3: Mean (and the standard deviation of the) response time of a batch in the two-server system. Job-wise assignment. See Table 1 for input parameters, Table 2 for values of policies' parameters.

	$\lambda = 0,05$ $\rho = 0,15$	$\lambda = 0,1$ $\rho = 0,30$	$\lambda = 0,18$ $\rho = 0,54$	$\lambda = 0,2$ $\rho = 0,60$	$\lambda = 0,23$ $\rho = 0,69$	$\lambda = 0,25$ $\rho = 0,75$	$\lambda = 0,32$ $\rho = 0,96$
RND	5,39 (5,1)	6,58 (6,24)	10,1 (9,53)	11,7 (11,13)	15,2 (14,49)	19 (18,16)	120 (110)
SG	5 (4,58)	6,05 (5,66)	9,1 (8,54)	10,5 (9,9)	**13,5 (12,65)**	**16,5 (15,49)**	**100 (94)**
AA	**4,84 (4,69)**	**5,98 (5,75)**	**9,06 (7,75)**	**10,3 (9,85)**	13,6 (13)	16,7 (15,8)	100 (96)

Table 4: Job's mean (and the standard deviation) response time in the 128-server system. Batch-wise assignment. Batches arrive according to the two-phase hyper-exponential process with the rate $\lambda = (0,8873/\lambda_1 + 0,1127/\lambda_2)^{-1}$. Mean size is $\mathsf{E}[B] = 4,125$, mean job size is $\mathsf{E}[X] = 1$. The offered load is $\rho = 4,125\lambda$. The RND-opt (and SG-opt) policy parameters values are given in the Table 6.

	$\lambda_1 = 0,0355$ $\lambda_2 = 0,0045$ $\rho = 0,0825$	$\lambda_1 = 0,0709$ $\lambda_2 = 0,0090$ $\rho = 0,165$	$\lambda_1 = 0,1479$ $\lambda_2 = 0,0188$ $\rho = 0,34375$	$\lambda_1 = 0,2218$ $\lambda_2 = 0,0282$ $\rho = 0,515625$	$\lambda_1 = 0,3227$ $\lambda_2 = 0,0409$ $\rho = 0,75$
RND-lb	481 (800)	488 (806)	579 (927)	792 (3873)	1620 (2450)
RND-opt	401 (583)	389 (557)	483 (600)	700 (849)	1510 (1673)
SG-lb	483 (800)	491 (806)	572 (905)	780 (1183)	1550 (2280)
SG-opt	400 (574)	390 (548)	478 (600)	**686 (825)**	**1510 (2049)**
AA	**298 (361)** $c = 14$	**338 (400)** $c = 13$	**473 (575)** $c = 10$	**703 (819)** $c = 7$	1560 (2280) $c = 6$

(persistently) the best choice when $\rho < 0,5$; but the SG-opt policy does better under the heavy traffic. An important observation from the data in the Tables 4 and 5 is that the performance of the RND and SG policies is sensitive to the choice of the parameter's values: load balancing leads to visibly larger mean response times. But in large heterogeneous systems searching for good values of $(p_1, \ldots, p_N)$ (and especially $(d_1, \ldots, d_N)$) may not be feasible. Thus, in general, the performance achieved under the RND and SG policies with load balancing is the only one, which can be guaranteed. From this point of view the ranking of the policies becomes independent of the system's load: the AA policy is uniformly the best choice among the three policies. Moreover it does not depend on the system's size: for a system with parallel single-server queues it requires estimation[19] of the single parameter.

[19] Yet no "rule of thumb" can be suggested for its estimation and simulation always has to be engaged.

Table 5: Job's mean (and the standard deviation) response time in the 128-server system. Job-wise assignment. See Table 4 for the input parameters, Table 6 for the policies parameters values.

	$\lambda_1 = 0,0355$ $\lambda_2 = 0,0045$ $\rho = 0,0825$	$\lambda_1 = 0,0709$ $\lambda_2 = 0,0090$ $\rho = 0,165$	$\lambda_1 = 0,1479$ $\lambda_2 = 0,0188$ $\rho = 0,34375$	$\lambda_1 = 0,2218$ $\lambda_2 = 0,0282$ $\rho = 0,515625$	$\lambda_1 = 0,3227$ $\lambda_2 = 0,0409$ $\rho = 0,75$
RND-lb	123 (272)	147 (302)	229 (424)	374 (640)	982 (1612)
RND-opt	111 (232)	131 (251)	209 (346)	354 (520)	912 (1183)
SG-lb	134 (332)	157 (361)	241 (500)	383 (721)	899 (1483)
SG-opt	111 (232)	130 (249)	**207** (**346**)	351 (510)	**830** (**1049**)
AA	**97**, **7** (**187**) $c = 6$	**126** (**228**) $c = 4$	**208** (**331**) $c = 2, 3$	**348** (**510**) $c = 1, 7$	857 (1225) $c = 1, 3$

Table 6: Parameters of the RND-opt (and SG-opt) policies for batch-wise and job-wise assignments for the values of the offered load ρ considered in the Tables 5 and 6: batch-wise | job-wise.

	$\rho = 0,0825$	$\rho = 0,165$	$\rho = 0,34375$	$\rho = 0,515625$	$\rho = 0,75$
p_2	0,001|0,001	0,001|0,001	0,003|0,029	0,040|0,018	0,043|0,042
p_3	0,068|0,067	0,025|0,013	0,052|0,062	0,063|0,054	0,077|0,071
p_4	0,119|0,127	0,101|0,111	0,102|0,106	0,116|0,120	0,116|0,122
p_5	0,159|0,167	0,157|0,166	0,157|0,154	0,152|0,162	0,145|0,152
p_6	0,186|0,183	0,204|0,207	0,201|0,189	0,180|0,187	0,176|0,173
p_7	0,213|0,214	0,233|0,234	0,222|0,213	0,212|0,216	0,206|0,204
p_8	0,246|0,239	0,270|0,265	0,261|0,247	0,235|0,243	0,236|0,233

UNRELIABLE SERVERS

Assume that servers in the queues are unreliable in the following sense. Each server is subject to breakdowns independently of whether it is busy of not and how long it has been busy. A breakdown makes the queue inoperative for a while. Thus each queue alternates between operative (up) and inoperative (down) periods. Once the "down" period is over, job's processing is resumed at the point where it was interrupted. It is assumed that the "up" and "down" periods constitute an alternating renewal process with the known CDFs $U(x)$ and $D(x)$ having means $\mathsf{E}[U]$ and $\mathsf{E}[D]$.

The list of available routing policies in the presence of breakdowns remains the same: RND, SG and AA. For Poisson arrivals and under the batch-wise assignment the arrival process to each queue with RND routing remains Poisson. And under some assumptions on $S(x)$, $U(x)$ and $D(x)$ it may be feasible to find the N-tuple $(p_1, p_2, \ldots, p_N)$ minimizing the mean response time of an arbitrary job. In general, the best parameters' values can be found only through simulation. Even though the structures[20] of the SG policy (4) and of the AA policy (Algorithm 1) do not take into account the presence of breakdowns, numerical experiments show that, when applied as is, both outperform the RND-opt policy. Nevertheless, one can go further and make various amendments in the AA policy, which make it aware of servers' unreliability. One of them is to independently[21] sample (for each queue) "up" and "down" periods (from $U(x)$ and $D(x)$) and to update the values of $\tilde{u}_i^{(n)}$ only if the queue n is in the "up" period. In order to demonstrate the performance of this modified AA policy we again consider the first example, and additionally assume that servers are unavailable for processing 10% of time. Let the time between the breakdowns (for each server) be exponentially distributed with the mean $\mathsf{E}[U] = 30$ and the repair time be exponentially distributed with the mean $\mathsf{E}[D] = 30/9$. Thus the fraction of time each server is in the "up" state is equal to $\mathsf{E}[U]/(\mathsf{E}[U] + \mathsf{E}[D]) = 0,9$. The whole system is stable if and only if $\rho = \lambda \mathsf{E}[B]\mathsf{E}[X]/(0,9 \sum_{i=1}^{2} v^{(i)}) \approx 3,33\lambda$. The values of the mean and the standard deviation of the mean response time depending on the offered load ρ for the batch-wise assignment are given in the Table 7. It can be seen that the presence of the breakdowns does not affect the ranking of the policies. If the markovian assumptions are dropped, according to our numerical experiments, the conclusion remains unchanged.

SUMMARY

The arrival-aware policy (Algorithm 1) always leads to better performance than the RND-opt policy. If the system's load is not high, then the AA policy is also better than the SG policy (with the (close to) optimal densities). Whenever the RND and SG policies adopt load balancing, the AA policy outperforms both of them across all values of the system's load. The performance improvement comes for free: the AA policy (as well as the SG

[20] But implicitly the presence of breakdowns is taken into account through the values of $(d_1, d_2, \ldots, d_N)$ and c.

[21] We note that if the online information about the "up" and "down" periods is available to the dispatcher, then the sampling is not needed and $\tilde{u}_i^{(n)}$ are updated only during the "up" periods of the queue n. This makes the AA policy outperform the SG policy (and RND-opt) for all $0 < \rho < 1$. The AA policy allows one also to take into account (in Bayesian framework) the uncertainty in the information about the "up" and "down" periods (for example, their means).

Table 7: Job's mean (and the standard deviation of the) response time in the system with two unreliable servers, which are available 90% of time. Batch-wise assignment. The input parameters are the same as in the Table 1. The offered load is $\rho \approx 3,33\lambda$.

	$\lambda = 0,05$ $\rho \approx 0,166$	$\lambda = 0,1$ $\rho \approx 0,333$	$\lambda = 0,18$ $\rho \approx 0,6$	$\lambda = 0,2$ $\rho \approx 0,666$	$\lambda = 0,23$ $\rho \approx 0,766$	$\lambda = 0,25$ $\rho \approx 0,833$	$\lambda = 0,32$ $\rho \approx 1,066$
RND-opt	7, 012 (7, 2) $p_1 = 0,001$	9, 603 (10) $p_1 = 0,157$	16, 71 (17) $p_1 = 0,260$	20, 20 (21) $p_1 = 0,276$	29, 15 (30) $p_1 = 0,289$	40, 56 (41) $p_1 = 0,305$	— —
SG-opt	7, 011 (7, 2) $d_1 = 0,001$	8, 810 (9, 2) $d_1 = 0,25$	14, 01 (14, 6) $d_1 = 0,31$	**16, 60 (17)** $d_1 = 0,315$	**23, 38 (24)** $d_1 = 0,32$	**32, 62 (34)** $d_1 = 0,325$	— —
AA	**6, 762 (7)** $c = 5,7$	**8, 569 (8, 9)** $c = 5,5$	**14, 00 (14, 5)** $c = 4,7$	16, 66 (17) $c = 4,4$	23, 53 (24) $c = 4,0$	32, 63 (34) $c = 3,8$	— —

policy) can be implemented in the dispatcher at very limited costs. In general the gain depends on the properties of the job/batch size distribution, number of queues, service rates: numerical experiments shows that it increases with the decrease of variability of the involved random quantities. Unfortunately the AA policy lacks, so far, the theoretical support. On the other hand, it allows various modifications (for example, accounting for servers' unreliability) and, for a system with $N \geq 2$ queues, requires estimation of the single parameter (compared to the RND and SG policy, which require $N - 1$ parameters). This all taken together makes the AA policy an appealing routing rule for batch-arrival unobservable dispatching systems.

REFERENCES

Anselmi, J. 2017. Asymptotically optimal open-loop load balancing. Queueing Systems. Vol. 87. No. 3-4. Pp. 245–267.

Anselmi, J. and B. Gaujal. 2011. The price of forgetting in parallel and non-observable queues. Perform. Eval. Vol. 68. No. 12. Pp. 1291–1311.

Arian, Y., Levy, Y. 1992. Algorithms for generalized round robin routing. Oper. Res. Lett. Vol. 12. Pp. 313–319.

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.

Cooper, R.B. 1981. Introduction to queueing theory. Elsevier North Holland, New York.

Feng, H., Misra, V., Rubenstein, D. 2005. optimal state-free, size-aware dispatching for heterogeneous $M/G/$-type systems. Performance Evaluation. Vol. 62. No. 1-4. Pp. 475–492.

Harchol-Balter, M., Crovella, M.E., Murta, C.D. 1999. On choosing a task assignment policy for a distributed server system. Journal of Parallel and Distributed Computing. Vol. 59. Pp. 204–228.

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.

Hyytiä, E., Aalto, S. 2011. To Split or not to Split: Selecting the Right Server with Batch Arrivals. Operations Research Letters. Vol. 41. No. 4. Pp. 325–330.

Ibaraki, T.I., Katoh, N. 1988. Resource allocation problems. Cambridge: MIT Press.

Konovalov, M.G., Razumchik, R.V. 2020. A simple dispatching policy for minimizing mean response time in non-observable queues with SRPT policy operating in parallel. Communications of the ECMS. Vol. 34. No. 1. Pp. 398–402.

Konovalov, M.G., Razumchik, R.V. 2018 Improving routing decisions in parallel non-observable queues. Computing. Vol. 100. No. 10. Pp. 1059–1079.

Lingenbrink, D., Iyer K. 2017. Optimal Signaling Mechanisms in Unobservable Queues with Strategic Customers. In Proceedings of the 2017 ACM Conference on Economics and Computation, New York, NY, USA. Pp. 347–347.

Mengistu, T.M., Che, D. 2019. Survey and taxonomy of volunteer computing. ACM Comput. Surv. Vol. 52. No. 3. Art. ID 59.

Sano, S., Miyoshi, N. 2000. Applications of m-balanced sequences to some network scheduling problems. Discrete Event system: Analysis and Control. Proceedings of the 5th Workshop on Discrete Event Systems. Pp. 317–325.

van Ommeren, J. C. W. 1990. Simple approximations for the batch-arrival $M^X/G/1$ queue. Operations Research. Vol. 38, No. 4. Pp. 678–685.

AUTHOR BIOGRAPHIES

MIKHAIL KONOVALOV is a Doctor of Sciences in Technics and holds position of the principal scientist at the Institute of Informatics Problems of the Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS). 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 the leading research fellow at the Institute of Informatics Problems of the Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences (FRC CSC RAS). 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`

AUTHOR INDEX