SIMULATION STUDY OF A SIGNALLING PROTOCOL EFFICIENCY IN A COMPOSITE RADIO ENVIRONMENT

Charalabos Skianis
Kimon Kontovasilis
George Kormentzas
National Center for Scientific Research "DEMOKRITOS"
Institute for Informatics & Telecommunications
GR-15310 Ag. Paraskevi, POB 60228, Greece
E-mail: {kkont, skianis, gkorm}@iit.demokritos.gr

Giordana Lisa
Motorola Electronics S.P.A - GSG Italy
Via Cardinal Massaia 83, 10147 Torino, Italy
E-mail: giordana.lisa@motorola.com

KEYWORDS
Heterogeneous Systems, Signalling, Simulation.

ABSTRACT

Key aspects of a signalling protocol governing the interactions between terminals and network entities within a composite radio environment are presented in this paper. On the basis of these aspects, simple simulation models for capturing the dynamics of the message exchanges are discussed. The simulation study allows the derivation of both qualitative and quantitative results, that provide insight on key characteristics of the environment.

INTRODUCTION

Composite radio is directly relevant to the actively researched area of Fourth Generation (4G) wireless systems (4GVision&Technologies 2004). The concept refers to the joint employment of co-operating heterogeneous wireless technologies (Mobile 2.5G or 3G telecommunications systems, such as GPRS (3GPP TS 23.060 2004) and UMTS (3GPP TS 22.101 2004); Broadband Radio Access Networks, such as IEEE 802.11 (Groups 802.11 2004) or HIPERLAN (Varshney and Vetter 2000); and wireless broadcasting technologies, such as DVB-T (Digital Video Broadcasting 2002)) towards a versatile infrastructure that can support flexible wireless access to quality-aware information services.

Taking the same approach as the CREDO (Composite Radio and Enhanced service Delivery for the Olympics) (CREDO consortium 2001) project, this paper does not regard composite radio simply as a system where terminals switch to alternative access networks through a vertical handover merely upon loss of coverage, but rather as a system where its constituent components coordinate intelligently, towards exploiting the increased potential for optimisation that becomes possible when these constituents are jointly operated. Operation at this level of intelligence presupposes the existence of appropriate management functionality; such functionality is assumed to be available at both the composite network and the wireless terminals. In the context just outlined, the following key features of the composite radio system may be identified:

- Diverse radio segments interconnected via a backbone.
- A Network Management System (NeMS) for the management of the composite radio network. It enables the joint optimisation of the alternate radio network segments, so as to deliver services efficiently, in terms of QoS.
- A Terminal Management System (TeMS) for the management of the wireless multimode terminals capable of operating in the composite radio environment and comprising functionality for interfacing to NeMS and for conducting decisions (NeMS-driven or independent) about the most appropriate radio technologies to be used for the efficient (in terms of QoS) reception of services through this terminal, under the each time applicable circumstances.
- The NeMS and TeMS exchange information towards beneficially combining the terminal's 'local view' (e.g., radio conditions in the area, services requested/received over the terminal, QoS levels associated with those services, etc) and the 'global view' of the composite network (e.g., traffic load over the various segments, QoS preservation via congestion avoidance, etc)
- Quality-aware applications and demanding service access.

It is clear that in order to fully develop a composite radio system comprised of the aforementioned key features it is imperative to adopt a signalling protocol governing the interactions between the terminal's TeMS and the network (through NeMS) towards the efficient selection (or reselection) of the most appropriate access network for the terminal to use. In this context, the paper presents a study on the efficiency of such a signalling protocol, implemented for the purposes of the IST project CREDO. The study contributes to the understanding of the protocol's dynamics and provides insight towards further improvements and modification. The concepts employed in the study are of a nature more general than the specific protocol and thus of value in other similar interaction mechanisms as well. Early work found in (Kontovasilis et al. 2003) on the model where simple analytic approximations are

Proceedings 18th European Simulation Multiconference
Graham Horton (c) SCS Europe, 2004
ISBN 3-936150-35-4 (book) / ISBN 3-936150-36-2 (CD)

employed to allow the computation of response times guided the presented work.

The rest of the paper is organised as follows: the second section outlines the technical requirements and the main structure of the TeMS-NeMS communication protocol developed in the context of the CREDO project. The third section argues on the necessity of an efficiency study and sets its context. Fourth section introduces the proposed model and exemplifies its use through the simulation of representatives case studies. Finally, the last section concludes the paper.

SIGNALING PROTOCOL REQUIREMENTS

The main technical requirements for the communication protocol are outlined as follows:

- The interaction mechanism should allow terminals to
  - Ask from the network (through NeMS) the engagement of new/additional services.
  - Notify the network about service termination and candidate access segments within the terminal area.
  - Report to the network status parameters relating to the quality with which services are received over the terminal (e.g., status values at the radio-/ IP-, and application-level).
- The interaction mechanism should enable the network (through NeMS) to
  - Advise a terminal about the access segments that should be selected.
  - Ask from terminals to send status reports for assessing the conditions in specific radio segments.
  - Instruct terminals to switch to a different access network.
- The message protocol governing the interaction mechanism should not depend on the high-level IP means used for its implementation. In particular, the implementation should not assume that the order of messages sent from any communication end is preserved. Given the signalling nature of the protocol, low-overhead implementations requiring few packet exchanges may be preferable, but should not be imposed by the protocol's structure.
- The message protocol should be independent from the specific radio access technologies integrated into the composite network.

Following the above requirements, the messages of the signalling protocol between the terminals and the network can then be categorised as follows:

- Initial terminal registration/initialisation to the composite network.
- A core pair of messages (a Service Request/Reply pair) used from a terminal (the request) in order to report in a periodic manner the services running and the reachable network or to ask for new services, to report a change in the currently running services (e.g., a stopped service) or in the networks reachable to the terminal, and from NeMS when replying with the appropriate network to be selected.
- Messages for reporting quality status information from terminals to NeMS.
- A message from NeMS to a terminal for triggering a message exchange leading to reselection of the network segment used by the terminal.

The main advantages for the signalling protocol implemented within CREDO are its simplicity (easy to implement), its low overhead (appropriate to carry critical messages), and finally that it is specifically designed for responding to the challenges posed on a composite radio environment. An outline of the protocol in the context of the multimode terminal's architecture can be found in 0.

SIGNALING PROTOCOL EFFICIENCY STUDY

In light of the protocol's outline in the previous section, the bulk of the signalling traffic are Service Requests (sent from TeMS to NeMS), which trigger corresponding Service Replies (from NeMS back to the TeMS). If reply to a Service Request hasn't been received within time equal to half a Lifetime threshold, the request is considered to have been lost and is retransmitted by the TeMS.

Customarily each terminal should periodically send Service Request messages every T time units, even if there is no updated information to report. These periodic messages act also as "keep-alive" indicators, assuring NeMS that the terminal is "still there". Note that whenever $T \leq Lifetime/2$, there is no point in retransmitting timed-out requests, as a more recent message has been issued before deciding to repeat the old one. Retransmissions are of value when $T >> Lifetime/2$.

Since other messages occur rarely compared to the ones just mentioned, they are not so important when studying the efficiency of the signalling protocol.

Exchange of messages between TeMS and NeMS occur in a network setting like that depicted in Figure 1. The configuration includes a GPRS network, a WLAN segment and a DVB-T segment. Note that the latter is unidirectional and must use one of WLAN or GPRS as return channels. Due to the presence of Mobile IP, each message from TeMS to NeMS must traverse the uplink radio segment, reach the system's Home Agent, and be forwarded to NeMS for processing. A corresponding sequence is involved in the backwards direction.

In light of this multi-step process, it is important to ensure that the exchange of messages occurs efficiently, i.e., that excessive delays in the delivery of messages are not very probable, considering even large-scale contexts.

Towards this goal, suitable modelling is employed to capture key properties of the signalling dynamics. Focus is on a 'steady-state' setting, where terminals regularly send Service Requests to NeMS and receive back the corresponding replies from it. The quantity of primary interest is the total roundtrip time from the issuance of the Service Request to the reception of the Service Reply. This should be kept sufficiently low to ensure that the TeMS may take benefit of the advice from NeMS as soon as possible, in case this is required.

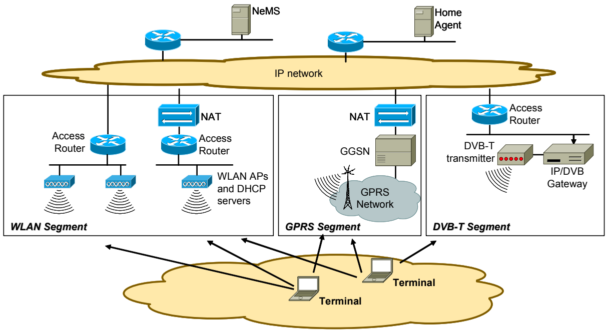


Figure 1: Composite radio network architecture

MODEL SPECIFICATION

This section introduces, in a detailed manner, the model developed for the study of the efficient exchange of signalling messages between network and terminal entities followed by the simulation study.

The Model

Figure 2 identifies the various components contributing to the roundtrip time between the emission of a TeMS-initiated Service Request message and the reception of the Service Reply message returned by NeMS in response. These components include network traversal (one-way delay) times and queueing (i.e., waiting plus service) times.

Random variables of potentially distinct probability distribution function (PDF) are used for the traversal times of different radio networks. Without further modelling complexity, it is also possible to assign different PDFs for uplink and downlink traversal of the same radio segment (this issue being void in the case of DVB-T, as it is unidirectional). The main simplifying assumption adopted by the model is that the times for different traversals along a given radio segment are assumed independent, for messages associated either to the same or to different terminals. This is an assumption typically employed in end-to-end networking studies and has been proven quite reasonable to adopt.

The traversal times over the IP backbone are associated with a PDF different from the previous ones. An independence assumption like the one just mentioned is assumed for this case too. It should be pointed out that although IP traversal does not always correspond to one-way delay between the same endpoints (e.g., there is traversal from the radio network to the Home Agent and from the Home Agent to the machine hosting NeMS) nevertheless, the model follows the assumption that all such times are distributed according to the same PDF. This assumption, supported by measurements, appears a reasonable one to make for high-speed backbones that employ optimised routing between their major nodes. These characteristics are quite representative of the setting being modelled.

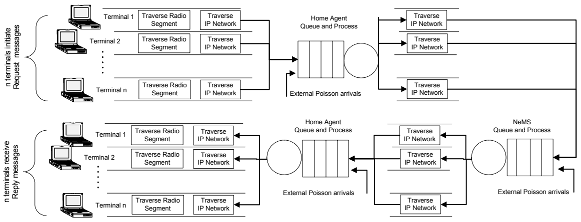


Figure 2: Flow of messages in Service Request/Reply chains

For the functional form of the traversal time PDFs, a rather conservative choice (i.e., one that is likely to yield estimates worse than the actual values) can be provided by the distribution corresponding to the sum of a constant time plus one term with a truncated exponential PDF. The second term signifies an additional occasional delay, due to congestion. Measured statistics for the mean, variance and percentile of exceeding a preset threshold could serve matching purposes. For a simpler choice, a plain exponential distribution may also be used. As already noted, different parameters apply to each of the three radio segments and to the IP backbone.

Besides network traversals, the time components in Figure 2 include two queueing times, at the Home Agent and at the NeMS, where messages pool for processing. Both queues are assumed to operate according to the FIFO policy. The service time at the Home Agent queue is the time required for processing/forwarding the packet carrying the message. The service time at the NeMS queue is the amount required for the internal bookkeeping (including possibly the execution of short-term optimisation algorithms) and for creating the Service Reply message.

The traffic load to these queues primarily consists of the periodically generated Service Request messages (for both queues) and the corresponding Service Replies (only for the Home Agent queue). The model assumes that each active terminal in the composite network generates a service request every T time units. Consequently, the average rate of message arrivals due to these periodic messages is n/T at the NeMS queue and 2n/T at the Home Agent (because for each request there is a corresponding reply). It is conservatively assumed that the packets carrying the requests and replies do not get lost along the way and thus the queues are presented with the full 'nominal' load.

Besides the explicitly considered messages (and possibly retransmitted Service Requests, discussed shortly), external Poisson load is introduced by the model to each of the two queues. For the NeMS queue, the Poisson load models packets carrying other messages (e.g., corresponding to non regular Service Requests for invocation of a new service, or messages of a type not considered in this study). This load is typically low, compared to the main one. For the Home Agent queue, the Poisson load, besides implicit messages also represents actual data packets (i.e., carrying content from servers to terminals) and (possibly) packets coming from other parts of the Internet, external to the composite network infrastructure.

The choice of Poisson process is a reasonable assumption when modelling a large aggregation of independent packet arrivals. In fact, when modelling a composite network with a large number of terminals, all traffic arriving at the two queues can be reasonably assumed of Poisson type. This further simplification is not employed in the simulation study. However, comparisons of mean value results between the simulation and an approximate analytic technique (c.f., (Kontovasilis et al. 2003)) support the validity of the simpler model in a context associated with many terminals.

In following the functionality of the signalling protocol, the model uses a Lifetime constant for determining 'expired' Service Requests (retransmitted only if the model's input parameters satisfy T>Lifetime/2).

The model treats the events relevant to disjoint service request/replies (associated to either the same or different terminals) through independent "process threads" (meaning that events corresponding to drawing a time sample for e.g., traversing the radio segment, occur in parallel and independently between different threads) as depicted in Figure 2. These "threads" coordinate (become dependent) at the two common queues, which handle all packets arriving at them.

It is clear that the roundtrip delay involved in a request/reply pair is equal to the sum of all the random variables associated with the time components of Figure 2, distinguishing when applicable between uplink and downlink radio traversal (e.g., a required distinction for a terminal in DVB-T, as the return path is through another radio segment). However, when the parameters of the model suggest that retransmissions be modelled, the computation of the roundtrip is a bit more complicated when timeouts occur. Indeed, assume that a request timeouts k times; this means that k+1 requests were ultimately generated and only the last one received a reply before time equal to Lifetime/2 passed from its issuance. In that case, the roundtrip delay should be computed by subtracting the time when the first request was generated from the time when the (k+1)-th reply was received at the terminal.

Given this model description the following list provides suitable Input/Output parameters for use when the model is implemented on a simulator:

- Input Parameters
    - Number of terminals active (in each of the three radio segments; for each of those in DVB-T, the uplink segment must also be specified).
    - Lifetime.
    - Time interval T between the periodic issuance of Service Request messages from each terminal (not including those generated due to timeouts, should the notion apply).
    - Delay PDFs for the IP traversal and the traversal of radio segment.
    - Service time PDFs for the Home Agent and the NeMS queues.
    - Poisson load at the NeMS and the Home Agent queues.
- Output parameters:
    - Statitistics for the roundtrip time for each segment hosting active terminals.
    - Probability of timeout (per segment).
    - Average arrival rate at each of the two queues (non-redundant only if retransmissions are enabled).

A Typical Case Study.

Table 1 summarises typical input parameters for an application of the methodology just described. Some parameters (like the f factor and the mean service times) have been intentionally kept greater than the values one would normally expect them to have, in order to come up with conservative predictions.

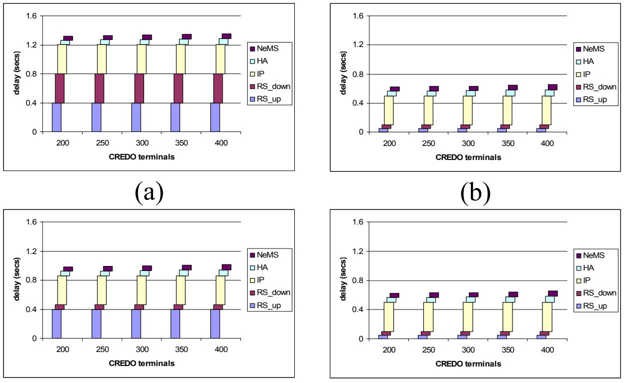
Table 1: Typical input for the study of system's reaction to terminals'requests

Network Parameters			
Network Size		240 terminals	
Terminal distribution among radio segments			
GPRS	WLAN	DVB/GPRS	DVB/WLAN
110	110	15	5
Traversal Times (ms)			
GPRS	WLAN	DVB-T	IP
400	50	60	100
Home Agent Parameters			
average service time		25 ms (i.e., 40 messages/sec)	
Sq. coefficient of variation for service time		1	
Arrival_rate_HA		9.6 messages/sec	
Utilisation, ρ		0.24	
NeMS Parameters			
average service time		50 ms (i.e., 20 messages/sec)	
Sq. coefficient of variation for service time		1	
Arrival_rate_NeMS		3.47 messages/sec	
Utilisation, ρ		0.174	
Additional Parameters			
Lifetime		180 sec	
f-factor for additional load		30%	
r_external		100% of n/T	

Nevertheless, traversal times are typical, as is the value of the Lifetime. The parameter T is half the Lifetime. The composite radio hosts a large number of terminals, so that the Poisson assumption is realistic.

The model has been simulated on the OPNET simulation platform. In all cases simulation time was large enough to guarantee reaching steady state conditions.

Figure 3 displays the mean roundtrip delay for each radio segment as a function of the number of terminals, with values ranging from 200 up to 400 terminals. With magnification of the size of the system, appropriate dimensioning is presumably applied to the radio segments, so that the traversal times do not vary. In the graph, different traversal and queueing time components are indicated by different colour (tones of grey in black-and-white printouts). It can be observed that most of the delay is due to traversing networks, while queueing is negligible, even for higher number of terminals, resulting in higher traffic at the queues. The most dominant part of the delay is due to IP traversal and (where applicable) to the traversal of the 'slow' GPRS segment.



(a) (b)

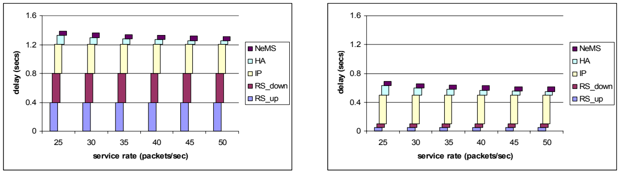
(c) (d)

Figure 3: Mean roundtrip delay (and parts of it) for exchange of service request/reply messages vs number of terminals, for terminals in GPRS (a), WLAN (b), DVB/GPRS (c) and DVB/WLAN (d)

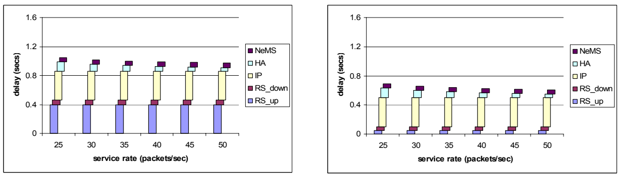
Additionally Figure 4 and Figure 5 display the roundtrip delay for each radio segment as a function of the Home Agent (rates from 25 up to 50 packets/sec) and the NeMS (rates from 10 up to 35 packets/sec) service times respectively. In each case a trend is exhibited in respective delays within the queues with the service times modified. However, the overall effect on the total round-trip time remains small in all cases due to the significantly larger delays imposed by the IP and the radio segment traversals, which are independent of the service times within the queues.

In all cases, in a qualitative sense the higher delays are found at the 'slow' GPRS segment, the faster segment is the WLAN, while DVB-T is affected by the choice of the uplink route with WLAN being the faster choice.

It is also notable that even in the most stringent case, the mean roundtrip delay remains below 2sec, i.e., about 50 times smaller than the timeout threshold. This illustrates that timeouts are highly improbable and would have remained so for considerably smaller values of T. Chebyshev's inequality (see e.g., (Feller 1968)) gives a conservative upper bound for this probability, through Pr{T>50·mean_delay}≤SCV_delay/502= SCV_delay·4·10-4, which is less than 1% even when the squared coefficient of variation for the delay is as high as 25.

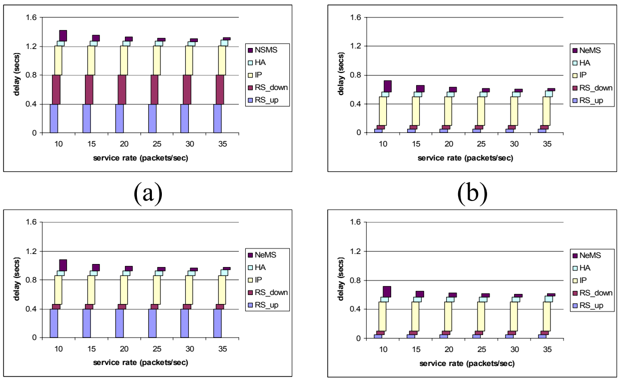


(a) (b)



(c) (d)

Figure 4: Mean roundtrip delay (and parts of it) for exchange of service request/reply messages vs Home Agent service rate, for terminals in GPRS (a), WLAN (b), DVB/GPRS (c) and DVB/WLAN (d)



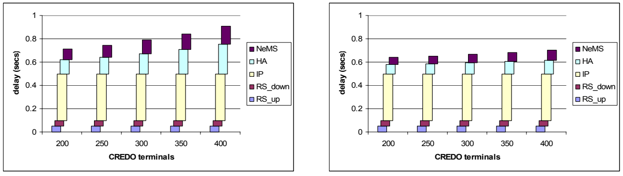
(a) (b)

(c) (d)

Figure 5: Mean roundtrip delay (and parts of it) for exchange of service request/reply messages vs NeMS service rate, for terminals in GPRS (a), WLAN (b), DVB/GPRS (c) and DVB/WLAN (d)

As an indicative additional test, Figure 6 displays the mean roundtrip delay for DVB/WLAN, this time with unreasonably high variability (SCV=10) in the service times (at both the Home Agent and NeMS). The impact of queueing in the total delay has increased, although it has remained much smaller than the IP traversal part (see subfigure (a), when T has the previous value). The right subfigure (b) maintains the high value in the SCV parameter, doubling though the value of T (effectively halving the load at the queues) with the end-result of bringing the queueing times back to the previous very small levels. This illustrates that the ability of the signalling protocol to negotiate the Lifetime parameter provides a versatile mechanism for controlling signalling congestion.

Note that in the case where the squared coefficient of variation is higher than one the Generalised Distribution (see e.g., (Kouvatsos 1994)) was adopted which is completely specified by just the mean and the variance.



(a) (b)

Figure 6: Mean roundtrip delay (and parts of it) for exchange of service request/reply messages vs number of terminals in DVB/WLAN and SCV=10, with T=1/90sec (a) and T=1/180sec (b)

CONCLUSIONS

This paper outlined key aspects of composite radio environments, focusing on a signalling protocol that governs the interactions between the network management system and intelligent multi-mode terminals. A framework for studying the impact of the signalling messages exchanged between the terminals and the Network and Service Management system, in a context of large scale was reported. Models for capturing the message exchanges were discussed, along with simulation studies for representative cases. In all cases the implemented protocol was shown to work efficiently. The model and the simulation allow the display of response times broken into separate time components for network traversal and queueing. Results from case studies showed that IP traversals contribute significantly to the overall time, while queueing times are less important. Future work would involve exhaustive comparison studies between simple analytic approximations and simulation additionally expanded to other types of exchanged signalling messages. Further degrees of complexity are envisaged towards a finer set of approximation and simulation tools for the study of the behaviour of the signalling protocol.

ACKNOWLEDGEMENT

This work has been performed in the framework of the project IST-2001-33093 CREDO, partly funded by the European Union under the IST programme.
The work outlined herein has been benefited from contributions by all partners in the CREDO consortium, which consists of: Motorola Labs (MOTLABS)—France, National Center for Scientific Research (NCSR) "Demokritos"—Greece, Motorola Technology Center of Italy (MTCI)—Italy, Thales Broadcast & Multimedia (TBM)—France, National Technical University of Athens (NTUA)—Greece, Vodafone Hellas—Greece, IBM Hellas—Greece.

REFERENCES

3GPP TS 22.101, 2004. 3GPP TS 22.101: *UMTS Principles*, http://www.3gpp.org.

3GPP TS 23.060, 2004. 3GPP TS 23.060: *General Packet Radio Service (GPRS)*,http://www.3gpp.org.

4GVision&Technologies, 2004. Electronically available information in http://www.mobileinfo.com/3G/4GVision&Technologies.htm

M. Catalina-Gallego and P. Stathopoulos, 2003. "Terminal Station Management System for Optimised Service Delivery in Composite Radio Environments", *IST Mobile & Wireless Communications Summit 2003*, 15-18 June 2003, Aveiro, Portugal.

CREDO consortium, 2001. *Technical Annex* (September).

Digital Video Broadcasting (DVB), 2002. http://www.dvb.org (Jan).

W. Feller, 1968. "An Introduction to Probability Theory and its Applications, Volume 1", John Wiley and Sons.

Groups 802.11, 2004. Electronically available information in http://grouper.ieee.org/groups/802/11/

K. Kontovasilis, C. Skianis, G. Kormentzas, 2003. 'Estimating Signalling Efficiency in a Composite Radio Environment', Proc. Of 1st International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks, D. Kouvatsos (ed.), Ilkley pp. 65/1-65/12.

D.D. Kouvatsos, 1994. "Entropy maximisation and queueing network models" in *Annals Oper. Res.* vol. 48, 63-126.

U. Varshney and R. Vetter, 2000. "Emerging Mobile and Broadband Wireless Networks", *Commun. of the ACM,* Vol. 43, No 6 (June).