[·] Tien V. Do, Patrick Wuechner, Tamas Berczes, Janos Sztrik and Hermann De Meer. A new finite-source queueing model for mobile cellular networks applying spectrum renting. *Asia Pacific Journal of Operational Research (APJOR)*, Volume 31, Number 2, pages 1440004:1–19, 2014. (c) World Scientific Publishing Co. & Operational Research Society of Singapore.

See `http://www.net.fim.uni-passau.de/papers/Do2014a` for full reference details (BibTeX, XML).

Asia-Pacific Journal of Operational Research
© World Scientific Publishing Company &
Operational Research Society of Singapore

A NEW FINITE-SOURCE QUEUEING MODEL FOR MOBILE CELLULAR NETWORKS APPLYING SPECTRUM RENTING

TIEN V. DO

Department of Networked Systems and Services,
Budapest University of Technology and Economics
Magyar tudósok körútja 2., Budapest, Hungary
do@hit.bme.hu

PATRICK WÜCHNER

Faculty of Informatics and Mathematics,
University of Passau,
Innstraße 43, 94032 Passau, Germany
patrick.wuechner@uni-passau.de

TAMÁS BÉRCZES and JÁNOS SZTRIK

Faculty of Informatics,
University of Debrecen,
Egyetem tér 1, Po.Box 12, 4010 Debrecen, Hungary
berczes.tamas@inf.unideb.hu,sztrik.janos@inf.unideb.hu

HERMANN DE MEER

Faculty of Informatics and Mathematics,
University of Passau,
Innstraße 43, 94032 Passau, Germany
hermann.demeer@uni-passau.de

This paper proposes a retrial queueing model with the finite number of sources to evaluate the performance of spectrum renting in mobile cellular networks. The model incorporates necessary ingredients such as the finite number of subscribers, their impatience and a queue for the outbound service. To consider the specific feature of spectrum renting and the current mobile cellular technology, a variable number of servers that are switched on and off in groups is introduced.

We present a novel way to take into account the renting fee, which can be used to fine-tune the operation of the spectrum renting procedure. Numerical results show that it is still profitable to initiate a spectrum renting request at high loads, even if no discount is offered by the frequency bands' owners.

This is the pre-print version accepted for APJOR (http://www.worldscientific.com/worldscinet/apjor).

Keywords: mobile cellular networks; retrial queues; finite-source; impatience; spectrum renting; discount factor

1. Introduction

The theory of retrial queues has been applied to model and evaluate the resource contention problem and the distinct user behavior in mobile cellular networks and telecommunication systems (Ajmone Marsan et al., 2001; Almási et al., 2005; Artalejo, 1998, 1999; Artalejo et al., 2007; Artalejo and Gómez-Corral, 2008; Artalejo, 2010; Artalejo and Lopez-Herrero, 2010; Economou and Lopez-Herrero, 2009; Falin and Templeton, 1997; Falin, 1990; Perel and Yechiali, 2010; Do, 2010b, 2011; Tran-Gia and Mandjes, 1997).

A typical classification of works on retrial queues and applications can be based on the assumption about the number of traffic sources. Infinite-source models are treated, e.g., by Ajmone Marsan et al. (2001), Artalejo (2010), Do (2010a), Do (2010b), Do et al. (2013) and Tran-Gia and Mandjes (1997). Finite-source retrial queues are discussed, e.g., by Artalejo and Lopez-Herrero (2012), Almási et al. (2005), Gharbi and Charabi (2012), Kulkarni and Liang (1997), Wüchner et al. (2009), Zhang and Wang (2012b), Gharbi and Dutheillet (2011), and Yang and Templeton (1987); Wang et al. (2011); Zhang and Wang (2012a). On the one hand, the assumption of the number of infinite sources covers general modelling opportunities and may lead to some efficient algorithms (Ajmone Marsan et al., 2001; Do, 2010a,b; Do and Chakka, 2010; Tran-Gia and Mandjes, 1997) for the performance evaluation of systems. On the other hand, the construction of finite-source models is justified by the inherent fact that the number of subscribers in the specific area of mobile cellular networks is finite.

Spectrum renting is a technique to relieve the temporary capacity shortages of a specific service area in wireless cellular networks, which have been intensively researched by the telecommunication industry in recent years (Buddhikot, 2007; Gandhi et al., 2008; Peha, 2009; Jabbari et al., 2010). Some queueing models for spectrum renting were proposed. Tzeng and Huang (2010) and Tzeng (2009) assumed that user channels can be rented in one unit, which is not realistic. The reason is that the separate blocks of user channels are defined in each frequency band, and each block should be controlled by a single network operator. Do et al. (2012) presented the first queueing model to take into account this technology aspect, but they did not consider the retrial phenomenon in their model. It is worth mentioning that Artalejo et al. (2005) first analysed a multi-server retrial model with a variable number of active servers, but their results cannot be applied directly for modelling spectrum renting, since servers are switched on and off one-by-one in their model.

Following the common practice applied in the theory of retrial queues for performance evaluation, this paper proposes a novel finite-source retrial queue to model spectrum renting in mobile cellular networks. The model incorporates switching servers on and off in groups, considers the finite size of the subscribers' population, their impatience, and has a queue for the subscribers who requested outbound service. Additionally, we present a novel way to take into account the renting fee, which can be used to optimize the spectrum renting procedure. Numerical results are derived and show that renting an additional frequency band is still desirable in certain circumstances, even if there is no discount from the owners of unused frequency bands.

To our best knowledge, this is the first proposal for the use of the theory of retrial queues to model spectrum renting. Moreover, several model variants could be derived to investigate additional aspects that are not completely explored yet in this paper. Therefore, we hope this article can initiate a new direction for the further exploration of a novel sub-class of retrial queueing models.

The rest of this paper is organized as follows. The review of an application-specific technical background that motivates this work is given in Section 2. In Section 3, the proposed model is presented. Numerical results and their discussion are provided in Section 4. Finally, Section 5 concludes the paper.

2. Key Aspects

This section provides a short overview of some technical aspects we consider in our model. A detailed description of techniques applied for mobile cellular networks can be found in the specifications of the relevant standardisation institutes, books, and further literature (Lin and Chlamtac, 2001; Rappaport, 2002).

2.1. *Block of traffic channels*

There are main alternatives for handling multiple access in mobile cellular networks: the Frequency and the Time Division Multiple Access (FDMA, TDMA) techniques (Lin and Chlamtac, 2001) and the Code Division Multiple Access (CDMA). The latter is based on spread-spectrum technology and a code division scheme (Rappaport, 2002). In these alternatives, each frequency band hosts a block of several channels.

To provide service for customers, a specific mobile network operator should (i) purchase the exclusive right for the use of a number of frequency bands in a country, (ii) divide the whole coverage area into a number of regular shaped cells, and (iii) assign a number of frequency bands to each cell. As a result, there will be a number of channels that can be allocated to subscribers in each cell.

2.2. *Spectrum renting*

Deciding on the optimum number of frequency bands to be purchased and assigned to a cell is a non-trivial task. In particular, over-provisioning is costly and under-provisioning might quickly annoy customers due to insufficient service quality. Spectrum renting can be used to relieve temporary shortages in the capacity of mobile cellular networks and hence provides a flexibility in the management of network resources.

Several alternatives to realize spectrum-sharing models and a spectrum policy reform are suggested by Peha (2009). Based on these, Do et al. (2012) argued that network operators may cooperate with each other in a dynamic market to perform spectrum renting based on some economic principles and a bid mechanism. That is, the owners of spectrum collaborate to increase the efficiency of the spectrum usage and to relieve the temporary capacity shortage of a particular cell in a mobile cellular network. For example when the number of calls increases in a specific area, a network operator could rent an additional frequency band from another operator to reduce the blocking probability of calls.

2.3. *Outbound or callback service*

Even if spectrum renting is performed, all channels might be busy during peak hours. When the initiation of a call is unsuccessful, the calling subscriber is informed through a signaling channel (e.g., operators send a voice announcement) that by pressing a certain button he or she can request the outbound service. Later, when a free channel is available, operators initiate a call for subscribers that previously requested the outbound service in the first-come-first-served discipline.

2.4. *The impatience of subscribers*

Behavioral psychology concerning the use of service offered by mobile cellular networks includes repeated attempts and abandonments. Both phenomena reflect the impatience of subscribers when all channels are occupied.

Following the arrival of a call, if all the available channels are occupied, a call is not be admitted into a network. Later, a subscriber initiates a repeated attempt for the admission of a call.

An abandonment happens when a subscriber's call becomes rejected and the subscriber gets impatient and gives up after a certain time without getting service.

3. The System Model

The finite-source retrial queueing model is illustrated in Figure 1. It comprises the finite number of sources, the orbit of retrying customers, the queue of customers that requested the outbound service, and the service area including a renting mechanism. The details of the model, taking into account the renting mechanism, are presented below.

A particular cell is assumed to contain a finite user population of K subscribers. Each subscriber initiates a call with rate λ as long as it is neither waiting for a channel nor being served. The operator allocates n initial channels from its own frequency bands to serve incoming calls. It is assumed that the operator can also rent m additional frequency bands. Each frequency band accommodates r channels. Therefore, the maximum number of available channels for calls is $n + m \times r$ if all m frequency bands are rented for the investigated cell.

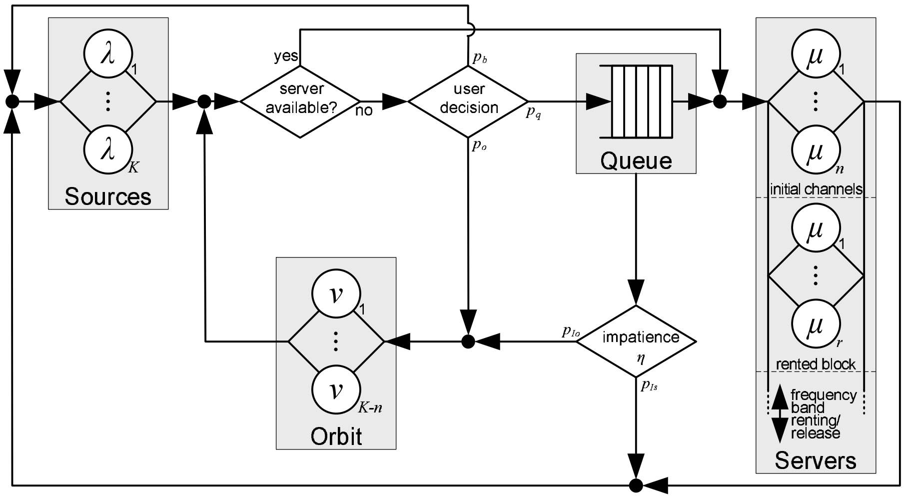

Fig. 1. A retrial queue with components.

3.1. *Subscribers' behavior*

The subscribers' behavior, including the outbound service and the impatience of subscribers, is considered in our model as follows. Upon the arrival of either a fresh call from the sources or a retrying call from the orbit, it checks whether there is an idle channel.

- If there is an idle channel, the call occupies this channel immediately for a call duration that is exponentially distributed with parameter μ. After the call duration, the channel is released and the subscriber returns to the sources.
- If all channels are busy, the subscriber chooses one of three alternatives:
  - The subscriber requests the outbound service with probability p_q. In our model, those subscribers join the virtual "outbound queue". Whenever an idle channel gets available (i.e., a busy channel gets idle or more channels are rented), the system automatically and immediately initiates the call for subscribers "waiting" in the outbound queue following a first-come-first-served scheduling discipline.
  - The subscriber joins the orbit with probability p_o. Subscribers waiting in the orbit retry to obtain a free channel with retrial rate ν.
  - The subscriber gives up and becomes idle, i.e., returns to the pool of sources, with probability $p_b = 1 - p_o - p_q$.

Note that the subscriber in the outbound queue might get impatient (after an exponentially distributed time) with rate η. He/she leaves the outbound queue either to the orbit (retrying later) with probability p_{Io} or "goes back" to the pool of the sources (i.e., giving up the call) with probability $p_{Is} = 1 - p_{Io}$.

3.2. *Spectrum renting*

To ensure the smooth operation related to renting and releasing frequency bands, the operation rule based on a hysteresis control with two thresholds is introduced by Do et al. (2012) as follows.

- *Initiation of renting a frequency band.* Whenever the number of free channels reduces to the lower threshold t_1 (which is due to the occupation of one channel by a call from either the finite-source pool or the orbit when the number of free channels is $t_1 + 1$), the network operator initiates the request to rent a new frequency band. Let p_r denote the probability that the request for a new frequency band is successful.

  If the request for a new frequency band fails (with probability $p_f = 1 - p_r$), the network operator retries with rate ν_r as long as the number of free channels is equal to t_1 or less.
- *Release of a frequency band.* Whenever the number of free channels increases to $t_2 + r$, then r channels are returned back to their owner after an exponentially distributed release time with mean $1/\mu_r$. Note that there is a positive release time and the immediate serving of subscribers that requested the outbound queue always has priority over releasing frequency bands.

The two spectrum renting parameters t_1 and t_2 are used to tune the performance of the system. The later a specific operator returns back rented channels, the higher the chance that more calls will be realized. However, there is a trade-off between the blocking probability and the fees a network needs to pay for a rented spectrum. To make a trade-off, we will define the average profit rate in Section 4.

3.3. *The underlying Markov chain*

To analyze the proposed queueing model (see Figure 1), we introduce the following notations.

- $k(t)$ is the number of active sources at time t,
- $s(t)$ denotes the number of occupied channels at time t,
- $b(t)$ $(0 \leq b(t) \leq m)$ is the number of rented frequency bands at time t,
- $q(t)$ denotes the number of calls in the outbound queue at time t,
- $o(t)$ is the number of calls in the orbit at time t.
- $u(t)$ $(u(t) \in \{0, 1\})$ indicates whether there is an unsatisfied block rental retrial at time t.

Note that $k(t) = K - s(t) - q(t) - o(t)$.

The main model parameters are summarised in Table 1 (the complete list of parameter values for the investigation can be found in Table 2). In addition, the system's steady-state probabilities are defined as

Table 1. Overview of main model parameters.

Parameter	Maximum	Value at time t	Mean $(t \to \infty)$
Number of active sources	K (population size)	$k(t)$	$\overline{K}$
Calls in queue	$Q = K - n$ (queue size)	$q(t)$	$\overline{Q}$
Orbiting calls	$O = K - n$ (orbit size)	$o(t)$	$\overline{O}$
Busy channels	$n + m \times r$	$s(t)$	$\overline{S}$
Number of rented bands	m	$b(t)$	$\overline{B}$
Activity of block rental retrial	1	$u(t)$	$\overline{U}$
Call generation rate		λ	
Total call generation rate	λK	$\lambda k(t)$	$\widetilde{\lambda} = \lambda \overline{K}$
Impatience rate		η	
Service rate		μ	
Retrial rate		ν	

$$p_{s,b,q,o,u} = \lim_{t \to \infty} P(s(t) = s, b(t) = b, q(t) = q, o(t) = o, u(t) = u)\,.$$

To obtain the steady state probabilities and performance measures within the Markovian framework, the mathematical tractability of the proposed model should be preserved. Therefore, we follow the classical approach frequently applied in the theory of retrial queues for the performance evaluation of wireless cellular networks (Ajmone Marsan et al., 2001; Almási et al., 2005; Artalejo, 1998, 1999; Artalejo et al., 2007; Artalejo and Gómez-Corral, 2008; Artalejo, 2010; Artalejo and Lopez-Herrero, 2010; Do, 2010b, 2011; Economou and Lopez-Herrero, 2009; Falin and Templeton, 1997; Falin, 1990; Perel and Yechiali, 2010; Tran-Gia and Mandjes, 1997). That is, all inter-event times (i.e., request generation time, impatience time, service time, retrial time, and times related to the spectrum renting) are assumed to be exponentially distributed and regulated by their rate parameter.

Based on the assumptions, the system is modelled by the five-dimensional Continuous Time Markov Chain (CTMC) $\{s(t), b(t), q(t), o(t), u(t)\}$, which is driven by the following types of events:

1) the arrival of calls,
2) the impatience and retrials of calls,
3) the departure of calls,
4) the request and retrial process for renting spectrum,
5) the release of rented spectrum.

Note that in the present case, the unique stationary distribution always exists, because the underlying CTMC is irreducible and the state space of the CTMC is finite. A graphical illustration of the five-dimensional CTMC appears to be infeasible and is omitted in this paper. Suitable analysis tools (like MOSEL-2, see Section 4 for details) allow the automatic generation of the CTMC and its infinitesimal generator matrix. These tools use standard techniques (Takacs, 1962; Bolch et al., 2006) to solve the global-balance equations for the system's steady-state probabilities.

As soon as the steady-state probabilities are known, further interesting performance measures to characterize spectrum renting can be defined as follows.

- The *blocking probability*[a] is

$$p_{\text{block}} = \sum_{(s,b,q,o,u)\in X\,|\,s=n+br} P(s,b,q,o,u).$$

- The *mean number of rented frequency bands* is

$$\overline{B} = \sum_{u=0}^{1}\sum_{b=1}^{m}\sum_{s=0}^{n+b\times r}\sum_{q=0}^{\min(K-n,K-s)}\sum_{o=0}^{\min(K-n-q,K-s-q)} bP(s,b,q,o,u).$$

- The *mean number of busy channels* is

$$\overline{S} = \sum_{u=0}^{1}\sum_{b=0}^{m}\sum_{s=1}^{n+b\times r}\sum_{q=0}^{\min(K-n,K-s)}\sum_{o=0}^{\min(K-n-q,K-s-q)} sP(s,b,q,o,u).$$

- The *mean queue length* is

$$\overline{Q} = \sum_{u=0}^{1}\sum_{b=0}^{m}\sum_{s=0}^{n+b\times r}\sum_{q=1}^{\min(K-n,K-s)}\sum_{o=0}^{\min(K-n-q,K-s-q)} qP(s,b,q,o,u).$$

- The *mean orbit length* is

$$\overline{O} = \sum_{u=0}^{1}\sum_{b=0}^{m}\sum_{s=0}^{n+b\times r}\sum_{q=0}^{\min(K-n,K-s)}\sum_{o=0}^{\min(K-n-q,K-s-q)} oP(s,b,q,o,u).$$

- The *mean number of active sources* is

$$\overline{K} = K - \overline{S} - \overline{Q} - \overline{O}.$$

- The *mean system throughput* is

$$\overline{\lambda} = \overline{K}\lambda.$$

- To compute the mean time spent in orbit, we follow the approach by Wüchner et al. (2009)). Let λ_O denote the throughput of the orbit. Using the Little Law, the mean time spent in the orbit (per visit) by every request arriving at the orbit is $\overline{T_O^\nu} = \frac{\overline{O}}{\lambda_O}$. The mean number of visits to the orbit per requests is $e_O = \frac{\lambda_O}{\overline{\lambda}}$. Hence, the overall *mean time spent in the orbit* is

$$\overline{T_O} = e_O\overline{T_O^\nu} = \frac{\lambda_O}{\overline{\lambda}}\frac{\overline{O}}{\lambda_O} = \frac{\overline{O}}{\overline{\lambda}}.$$

---

[a] In this paper, the term *blocking probability* refers to the probability (as seen by an outside observer) that all initial and rented servers are busy.

- The *mean time spent in queue* is

$$\overline{T_Q} = \frac{\overline{Q}}{\overline{\lambda}}.$$

4. Numerical Results

In this paper, the MOSEL-2 tool[b] is applied to compute the steady-state probabilities and the performance measures of the system. In particular, for generating the underlying CTMC and calculating the steady-state probabilities, MOSEL-2 calls SPNP (Ciardo et al., 1989; Hirel et al., 2000), which uses the successive over-relaxation method to solve the system of global-balance equations numerically.

Note that further evaluation tools such as MOSES (Greiner and Bolch, 1995) or TimeNET (German et al., 1995) can be invoked by MOSEL-2. Other evaluation methods, like discrete-event simulation, allows further generalizations for the presented model. For example, the assumption of exponentially distributed state transitions could be relaxed. However, these are out of the scope of this paper.

Table 2. Numerical values of default model parameters.

Parameter	Symbol	Value	Unit
Normalized traffic intensity	$\rho_0 = \frac{K\lambda}{n\mu}$	[0.6, 4.6]	-
Number of active sources	K	100	-
Retrial rate	ν	1	$1/s$
Prob. that subscriber gives up	p_b	0.1	-
Prob. that subscriber joins the queue	p_q	0.5	-
Prob. that subscriber joins the orbit	p_0	$1 - p_b - p_q$	-
Impatience rate	η	1/300	$1/s$
Prob. that impatient user goes back to orbit	p_{Io}	0.8	-
Prob. that impatient user gives up	p_{Is}	$1 - p_{Io}$	-
Number of channels per block	r	8	-
Maximum number of rented blocks	m	5	-
Number of channels without renting	n	$2 \times r$	-
Service rate	μ	1/53.22	$1/s$
Block renting threshold	t_1	$0 < t_1 < t_2 < r$	-
Block renting rate	λ_r	1/5	$1/s$
Block release threshold	t_2	$0 < t_1 < t_2 < r$	-
Block release rate	μ_r	1	$1/s$
Prob. for successfull renting	p_r	0.8	-
Prob. for unsuccessfull renting	p_f	$1 - p_r$	-
Block rental retrial rate	ν_r	1/7	$1/s$

In this section, we study the scenario where $K = 100$ subscribers reside in the investigated cell of a network operator. If not stated otherwise, the values of further

---

[b] MOSEL-2 is maintained at the Chair of Computer Networks and Communications, University of Passau, Germany. The project homepage is available at `http://mosel2.net.fim.uni-passau.de/` (last accessed: 8 April 2013).

parameters are identical to the default model parameter set summarized in Table 2. The normalized traffic intensity $\rho_0 = (K\lambda)/(n\mu)$ is changed in the interval $[0.6, 4.6]$ which translates to the blocking probability of approximately 1% to 2% for the default model parameter set.

To illustrate the capability of spectrum renting, we first compare a case where spectrum renting is not applied with a case where the spectrum renting procedure is carried out. In Figure 2, performance measures versus the normalized traffic intensity are plotted when spectrum renting is not applied. The probability that all n channels are busy reaches 1 when $\rho_0 = 2.1$ for $n = 16$. As expected, the performance is degraded by the increase of traffic load. Now, we are going to investigate how spectrum renting can help the network operator in such circumstances.

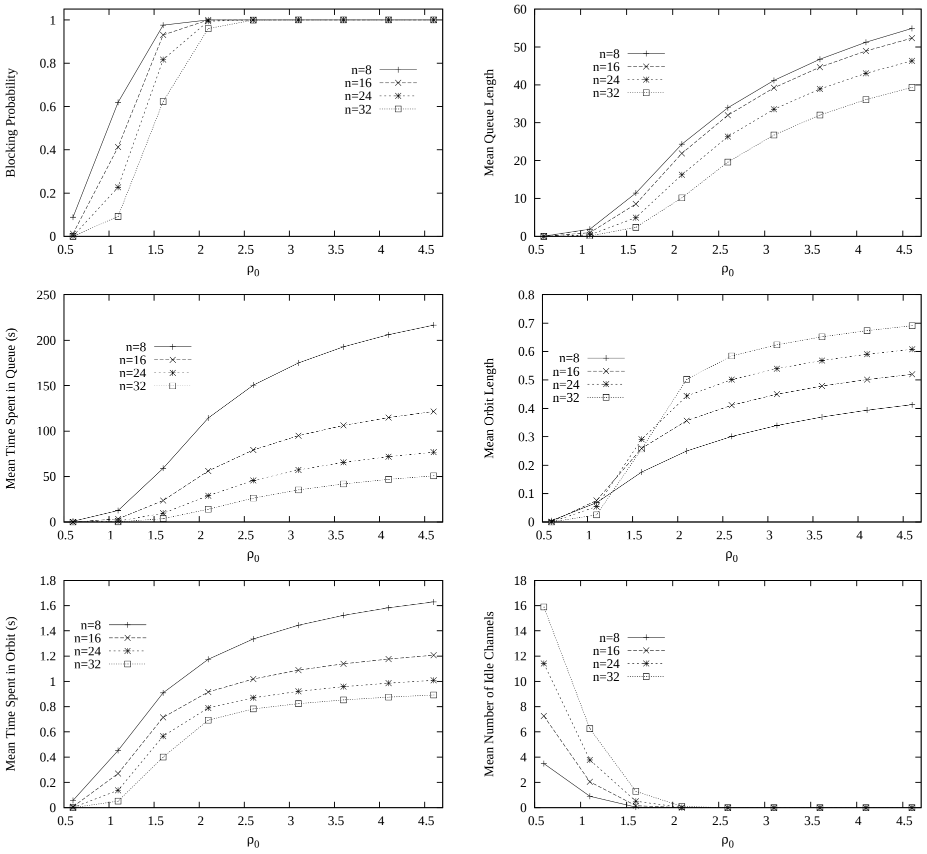

Fig. 2. Performance measures when there is no renting.

We compute the performance measures versus the normalized traffic intensity, t_1 and t_2. It is observed in Figure 3 that the change on t_1 only has a minimal impact on the mean number of rented blocks and on the mean number of busy channels.

Figure 4 also illustrates that t_1 has a negligible influence on the probabilities concerning the utilization of rented blocks. Furthermore, Figure 4 shows that for increasing load, the probability that rented blocks are actually used increases quickly. Hence, the renting of additional blocks is indeed suggested for increasing load.

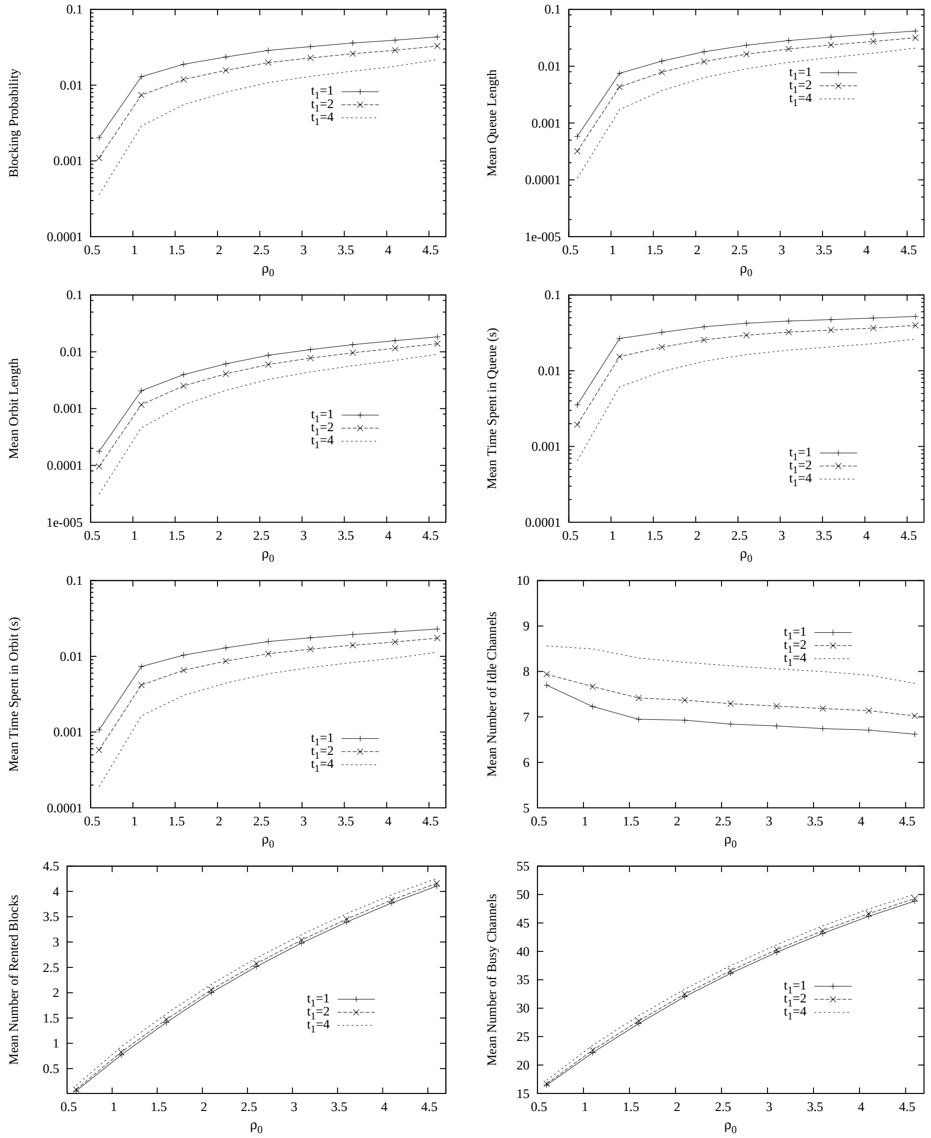

Fig. 3. Performance measures for $t_2 = 6$.

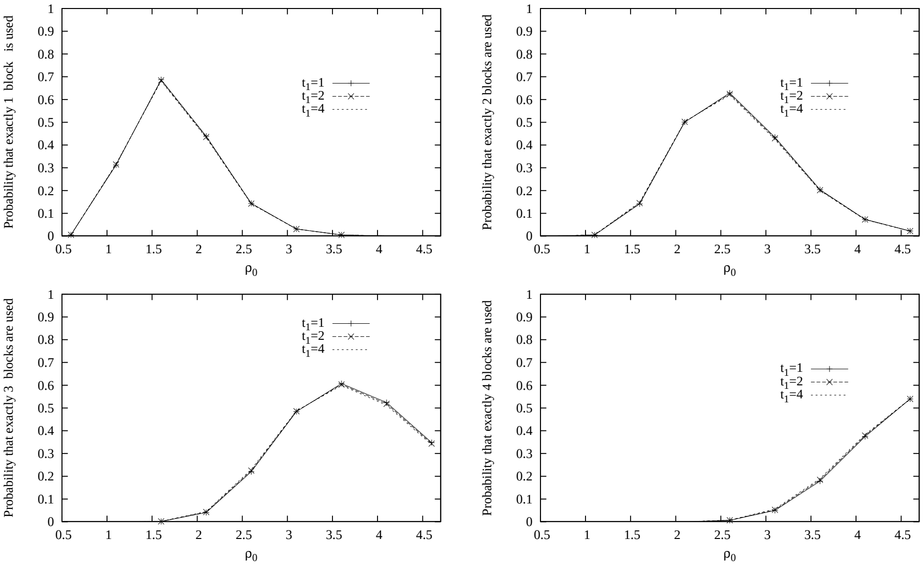

Fig. 4. Further performance measures for $t_2 = 6$.

On the one hand, the higher the value of t_1 (i.e., the earlier the operator initiates a request for an additional frequency band) is, the better the subscribers' "quality of experience" is in the term of the blocking probability and the waiting time of subscribers. The observation is confirmed by the curves in Figure 3. For example at $\rho_0 = 1.1$, the blocking probability is approximately 0.013, 0.007 and 0.003 for $t_1 = 1, 2$ and 4, respectively. On the other hand, high values of t_1 result in a long renting time and a high renting fee an operator needs to pay for additional channel blocks. The same observation can be obtained concerning the impact of t_2 (see Figure 5).

To consider the impact of a renting fee, the average profit rate (APR) is computed as follows

$$APR = \alpha \times \overline{S} - \beta \times \overline{B}, \tag{4.9}$$

where α and β are cost coefficients measured in cost units per time unit. Note that $\alpha \times \overline{S}$ can be interpreted as the revenue rate that the operator receives from the realization of calls, while $\beta \times \overline{B}$ is the renting fee due to the usage of rented frequency bands. Because there are r channels in one frequency band, $r \times \alpha$ is the maximum revenue rate obtainable with the rent of one frequency band. It is reasonable to assume that the expected revenue from a rented channel should be higher than or equal to the cost of renting the channel. Furthermore, the operator should take into account the phenomenon that not all rented channels can be fully utilized. Therefore, $r \times \alpha \geq \beta$ holds, which is equivalent to $r \times \alpha = d \times \beta$ for $d \geq 1$. Note that $d = r \times \alpha / \beta$ is interpreted as the discount factor.

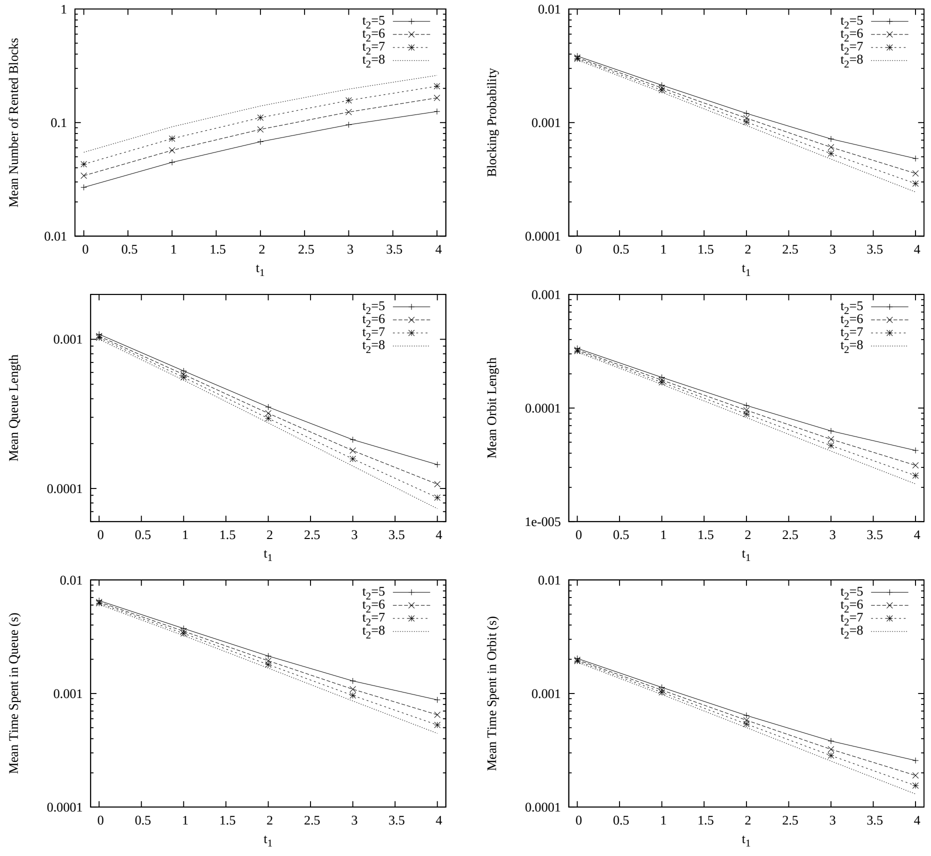

Fig. 5. Performance measures for $\rho_0 = 0.6$.

Figures 6, 7, and 8 illustrate the impact of the discount factor d on the APR. At low loads, the smaller the discount, the more it discourages the operator to rent additional channel blocks. At high loads, $d = 1$ (no discount) still encourages the operator to make an early request for additional bands and release the rented bands as late as possible.

We plot the APR versus ρ_0 in Figure 9 for three parameter sets: no retrials ($p_o \approx 0$, $p_{Io} \approx 0$), modest retrials ($p_o = 0.2$, $p_{Io} = 0.4$), and normal retrials ($p_o = 0.4$, $p_{Io} = 0.8$). It is observed that the APR increases with a distinct retrial behavior because retrials increase the server utilization (see Figure 10). However, retrials are annoying for the users, which may results in subscribers changing their service provider and consequently, in a long-term reduction of profit. This effect, nevertheless, is not yet included in our model.

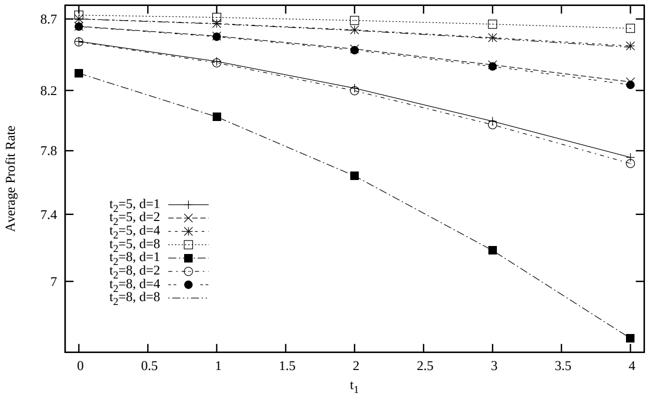

Fig. 6. *APR* vs. t_1 and d for $\rho_0 = 0.6$, $\alpha = 1$ cost unit/s $\beta = r \times \alpha/d$.

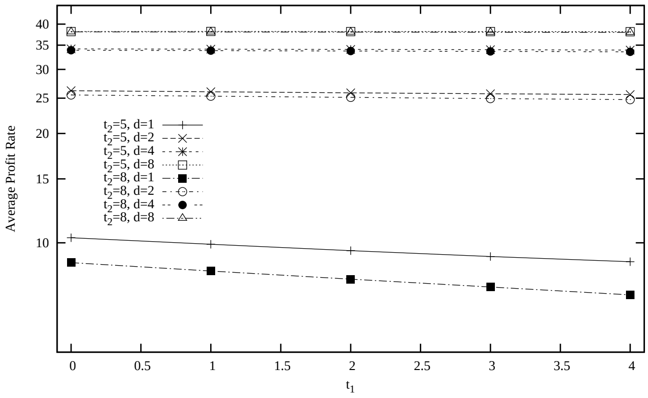

Fig. 7. *APR* vs. t_1 and d for $\rho_0 = 4.6$, $\alpha = 1$ cost unit/s, $\beta = r \times \alpha/d$.

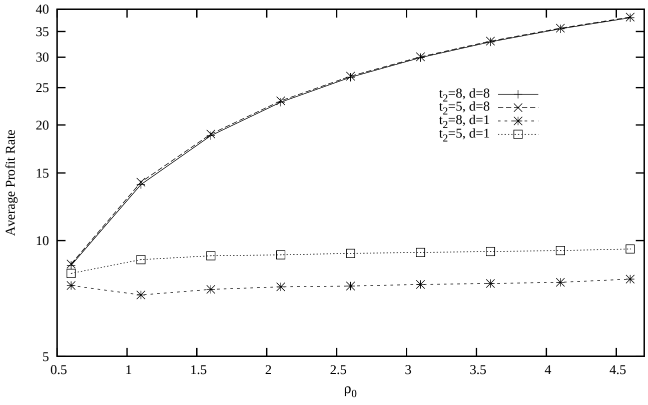

Fig. 8. *APR* vs. ρ_0 and d for $\alpha = 1$ cost unit/s, $\beta = r \times \alpha/d$.

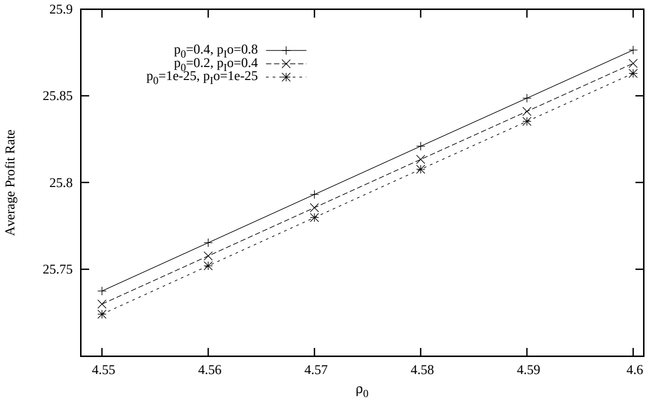

Fig. 9. Impact of retrials on APR for $\alpha = 1$ cost unit/s, $t_1 = 2$ $t_2 = 5$, $d = 2$, $\beta = r \times \alpha/d$.

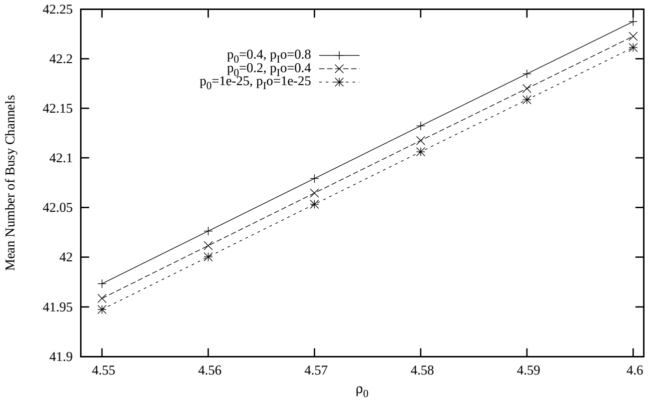

Fig. 10. Impact of retrials on the average number of busy channels for $\alpha = 1$ cost unit/s, $t_1 = 2$ $t_2 = 5$, $d = 2$, $\beta = r \times \alpha/d$.

5. Conclusions

We have proposed a finite-source retrial queueing model to evaluate the performance and the impact of spectrum renting in mobile cellular networks. We have investigated the impact of parameters on the performance of the system. The discount factor is introduced to take into account the renting fee. Numerical results show that it is still desirable to initiate a spectrum renting request even if there is no discount from the owners of frequency bands.

To our best knowledge, this is the first proposal for the use of the theory of retrial queues to model spectrum renting. Many variants and generalizations of the of presented model can be defined (e.g., the consideration of handover calls and guard channels, including non-exponential distributions), which will be investigated in our future works. We hope we can raise the attention of researchers to a sub-class of new retrial queueing models by shedding light on its application to the performance evaluation of mobile cellular networks.

Acknowledgement

The publication of T.V. Do, T. Bérczes and J. Sztrik was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

The contributions by University of Passau are partially supported by the European Community's Seventh Framework Program through the SOCIONICAL project (FP7-ICT-2007-3, Grant No. 231288) and through the Network of Excellence EINS (FP7-ICT-2011-7, Grant No. 288021).

The work of Támas Bérczes was funded in the framework of TÁMOP 4.2.4.A/2-11-1-2012-0001 “National Excellence Program Elaborating and operating an inland student and researcher personal support system”. The project was subsidized by the European Union and co-financed by the European Social Fund.

References

Ajmone Marsan, M., Carolis, G. D., Leonardi, E., Lo Cigno, R., and Meo, M. (2001). Efficient Estimation of Call Blocking Probabilities in Cellular Mobile Telephony Networks with Customer Retrials. *IEEE Journal on Selected Areas in Communications*, 19(2):332–346.

Almási, B., Roszik, J., and Sztrik, J. (2005). Homogeneous finite-source retrial queues with server subject to breakdowns and repairs. *Mathematical and Computer Modelling*, 42(5-6):673–682.

Artalejo, J. R. (1998). Retrial queues with a finite number of sources. *J. Korean Math. Soc.*, 35:503–525.

Artalejo, J. R. (1999). Accessible bibliography on retrial queues. *Mathematical and Computer Modelling*, 30:1–6.

Artalejo, J. R. (2010). Accessible bibliography on retrial queues: Progress in 2000-2009. *Mathematical and Computer Modelling*, 51(9-10):1071–1081.

Artalejo, J. R., Economou, A., and Gómez-Corral, A. (2007). Applications of maximum queue lengths to call center management. *Computers & OR*, 34(4):983–996.

Artalejo, J. R. and Gómez-Corral, A. (2008). *Retrial Queueing Systems: A Computational Approach*. Springer Verlag.

Artalejo, J. R. and Lopez-Herrero, M. J. (2010). Cellular mobile networks with repeated calls operating in random environment. *Computers & OR*, 37(7):1158–1166.

Artalejo, J. R. and Lopez-Herrero, M. J. (2012). The single server retrial queue with finite population: a BSDE approach. *Operational Research*, 12(2):109–131.

Artalejo, J. R., Orlovsky, D. S., and Dudin, A. N. (2005). Multi-server retrial model with variable number of active servers. *Comput. Ind. Eng.*, 48(2):273–288.

Bolch, G., Greiner, S., de Meer, H., and Trivedi, K. (2006). *Queueing Networks and Markov Chains*. John Wiley & Sons, New York, 2nd edition.

Buddhikot, M. M. (2007). Understanding dynamic spectrum access: Models,taxonomy and challenges. In *2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, DySPAN 2007.*, pages 649–663.

Ciardo, G., Muppala, J. K., and Trivedi, K. S. (1989). SPNP: Stochastic petri net package. In *PNPM*, pages 142–151. IEEE Computer Society.

Do, T. V. (2010a). An efficient solution to a retrial queue for the performability evaluation of DHCP. *Computers & OR*, 37(7):1191–1198.

Do, T. V. (2010b). A new computational algorithm for retrial queues to cellular mobile systems with guard channels. *Computers & Industrial Engineering*, 59(4):865–872.

Do, T. V. (2011). Solution for a retrial queueing problem in cellular networks with the fractional guard channel policy. *Mathematical and Computer Modelling*, 53(11–12):2058–2065.

Do, T. V. and Chakka, R. (2010). An Efficient Method to Compute the Rate Matrix for Retrial Queues with Large Number of Servers. *Applied Mathematics Letters*, 23:638–643.

Do, T. V., Do, N. H., and Chakka, R. (2012). A new queueing model for spectrum renting in mobile cellular networks. *Computer Communications*, 35(10):1165–1171.

Do, T. V., Do, N. H., and Zhang, J. (2013). An enhanced algorithm to solve multiserver retrial queueing systems with impatient customers. *Computers & Industrial Engineering, 10.1016/j.cie.2013.04.008.*

Economou, A. and Lopez-Herrero, M. J. (2009). Mobile networks performance analysis of a cellular mobile network with retrials and guard channels using waiting and first passage time measures. *European Transactions on Telecommunications*, 20(4):389–401.

Falin, G. (1990). A survey of retrial queues. *Queueing Systems*, 7(2):127–167.

Falin, G. and Templeton, J. (1997). *Retrial Queues*. Chapman & Hall.

Gandhi, S., Buragohain, C., Cao, L., Zheng, H., and Suri, S. (2008). Towards real-time dynamic spectrum auctions. *Computer Networks*, 52(4):879–897.

German, R., Kelling, C., Zimmermann, A., Hommel, G., Berlin, T. U., and Robotik, F. P. U. (1995). TimeNET - a toolkit for evaluating non-Markovian stochastic Petri nets. *Performance Evaluation*, 24:69–87.

Gharbi, N. and Charabi, L. (2012). Wireless networks with retrials and heterogeneous servers: Comparing random server and fastest free server discipline. *International Journal on Advances in Networks and Services*, 5(1-2):102–115.

Gharbi, N. and Dutheillet, C. (2011). An algorithmic approach for analysis of finite-source retrial systems with unreliable servers. *Computers & Mathematics with Applications*, 62(6):2535–2546.

Greiner, S. and Bolch, G. (1995). Modeling production lines with blocking, batch processing and unreliable machines using the Markov analyzer MOSES. In *Proc. European Simulation Symp. - ESS '95*, pages 303–309.

Hirel, C., Tuffin, B., and Trivedi, K. S. (2000). SPNP: Stochastic Petri nets. version 6.0. In *Proceedings of the 11th International Conference on Computer Performance Evaluation: Modelling Techniques and Tools*, pages 354–357. Springer-Verlag, London, UK.

Jabbari, B., Pickholtz, R., and Norton, M. (2010). Dynamic spectrum access and management dynamic spectrum management. *IEEE Wireless Communications*, 17(4):6 –15.

Kulkarni, V. G. and Liang, H. M. (1997). *Frontiers in Queueing: Models and Applications in Science and Engineering*, chapter Retrial queues revisited, pages 19–34. CRC Press.

Lin, Y.-B. and Chlamtac, I. (2001). *Wireless and Mobile Network Architectures.* Wiley Computer Publishing, John Wiley & Sons, Inc.

Peha, J. M. (2009). Sharing spectrum through spectrum policy reform and cognitive radio. *Proceedings of the IEEE*, 97:708–719.

Perel, N. and Yechiali, U. (2010). Queues with slow servers and impatient customers. *European Journal of Operational Research*, 201(1):247–258.

Rappaport, T. S. (2002). *Wireless Communications, Principles and Practice.* Prentice-Hall.

Takacs, L. (1962). *Introduction to the theory of queues.* Oxford University Press, New York.

Tran-Gia, P. and Mandjes, M. (1997). Modeling of customer retrial phenomenon in cellular mobile networks. *IEEE Journal on Selected Areas in Communications*, 15(8):1406–1414.

Tzeng, S.-S. (2009). Call admission control policies in cellular wireless networks with spectrum renting. *Computer Communications*, 32(18):1905–1913.

Tzeng, S.-S. and Huang, C.-W. (2010). Threshold based call admission control for QoS provisioning in cellular wireless networks with spectrum renting. In Sobh, T. M., Elleithy, K. M., and Mahmood, A., editors, *Novel Algorithms and*

Techniques in Telecommunications and Networking, pages 17–22. Springer, Berlin Heidelberg.

Wang, J., Zhao, L., and Zhang, F. (2011). Analysis of the finite source retrial queues with server breakdowns and repairs. *Journal of Industrial and Management Optimization*, 7:655–676.

Wüchner, P., Sztrik, J., and de Meer, H. (2009). Finite-source M/M/S retrial queue with search for balking and impatient customers from the orbit. *Computer Networks*, 53(8):1264–1273.

Yang, T. and Templeton, J. G. C. (1987). A survey on retrial queues. *Queueing Systems*, 2:201–233.

Zhang, F. and Wang, J. (2012a). Performance analysis of the retrial queues with finite number of sources and service interruptions. *Journal of the Korean Statistical Society, doi:10.1016/j.jkss.2012.06.002.*

Zhang, F. and Wang, J. (2012b). Stochastic analysis of a finite source retrial queue with spares and orbit search. In Schmitt, J. B., editor, *MMB/DFT*, volume 7201 of *Lecture Notes in Computer Science*, pages 16–30. Springer.

Tien Van Do is a full professor of the Budapest University of Technology and Economics, and a leader of Communications Network Technology and Internetworking Laboratory. He has participated in a number of European research projects. He led various projects on network planning and software implementations that results are directly used for industry such ATM & IP network planning software for Hungarian Telekom, GGSN tester for Nokia, performance testing program for the performance testing of the NOKIAs IMS product, automatic software testing framework for Nokia Siemens Networks. His research interests are queuing theory, telecommunication networks, cloud computing, performance evaluation and planning of ICT Systems.

Patrick Wüchner received his computer science Diploma in 2004 from the University of Erlangen-Nurnberg, Germany. Since then, he is research and teaching assistant as well as PhD student at the University of Passau, Germany. He is interested in research on mathematical performance and reliability modeling of computer systems, communication networks, and self-organizing systems.

Támás Bérczes received his MSc Degree in Mathematics in 2000, at the University of Debrecen, Hungary. He received a Ph.D. degree in 2011. He is currently assistant professor at the Department of Informatics Systems and Networks of the same university. His primary research interests are performance analysis of retrial queues and their application.

János Sztrik is a Full Professor at the Faculty of Informatics, University of Debrecen, Hungary. He obtained the Candidate of Mathematical Sciences Degree in Probability Theory and Mathematical Statistics in 1989 from the Kiev State University, Kiev, USSR, Habilitation from University of Debrecen in 1999, Doctor of the Hungarian Academy of Sciences, Budapest, 2002. His research interests are in the field of production systems modeling and analysis, queueing theory, reliability theory, and computer science.

Hermann de Meer is currently appointed as Full Professor at the University of Passau, Germany, and as Honorary Professor at University College London, UK. He is director of the Institute of IT-Security and Security Law (ISL) at the University of Passau, Germany. He had been an Assistant Professor at Hamburg University, Germany, a Visiting Professor at Columbia University in New York City, USA, and a Reader at University College London, UK. His research interests include peer-to-peer systems, self-organization, quality of service, Internet protocols, home networking, IT security, and mobile computing.

Appendix

A model described in the MOSEL-2 modeling language usually consists of the following parts:

(1) *Declaration part:* In the first part (see Listing 1) of the MOSEL-2 code, we have to define all the system parameters, which we will use in the numerical explorations. We can declare a list of values in the PARAMETER lines, which is a powerful feature of the language, since this makes experimenting with the model very easy.

Listing 1. MOSEL-2 Constant and parameter part.

```
/*** CONSTANTS AND PARAMETERS ***************************************************/

// User Behavior
PARAMETER rho    := 0.6 .. 4.6 STEP 0.5;   // normalized traffic intensity
CONST     K      := 100;                   // population size
CONST     nu     := 1;                     // retrial rate
CONST     p_b    := 0.1;                   // prob. that user gives up (balking; -> sources)
CONST     p_q    := 0.5;                   // prob. that user presses button (-> queue)
CONST     p_o    := 1 - p_b - p_q;         // prob. that user retries later  (-> orbit)
CONST     eta    := 1/300;                 // rate of queueing users getting impatient
CONST     p_Io   := 0.8;                   // prob. that impatient user retries
                                           //   later (-> orbit)
CONST     p_Is   := 1-p_Io;                // prob. that impatient user gives
                                           //   up (-> sources)

// Block Rental
CONST     r      := 8;                     // number of servers/channels per block
CONST     m      := 5;                     // maximum number of blocks that can be rented
CONST     n      := 2*r;                   // minimum number of servers/channels

PARAMETER t1     := 1,2,4;                 // block renting threshold
CONST     lam_r  := 1/5;                   // block renting rate

CONST     p_r    := 0.8;                   // block rental success probability
CONST     p_f    := 1-p_r;                 // block rental failure probability
CONST     nu_r   := 1/7;                   // block rental retrial rate

CONST     t2     := 6;                     // block release threshold
CONST     mu_r   := 1;                     // block release rate


// Service
CONST     C      := n+r*m;                 // maximum number of simultaneous calls possible
CONST     mu     := 1/53.22;               // service rate

// Arrival
CONST lambda = rho * n * mu / K;           // call generation rate
```

(2) *Nodes part:* The node part (see Listing 2) defines the nodes of the system. Each node can hold an integer-valued number of jobs up to its capacity. For every component in the model, a name, capacity, and initial occupancy has to be given. Our queueing network model contains nine nodes: one node each for the active sources (Sources), the incoming call (Request), the outbound service (Queue), the orbit (Orbit), the servers (Servers), and for impatient calls (Impatient), and three further nodes for describing the renting mechanism: the rented channel blocks (Blocks), the renting orbit (RentOrb), and the incoming renting request (Trial).

Listing 2. MOSEL-2 Nodes part.

```
/*** NODES ****************************************************************************/

// User and Servers
NODE Sources[K]     := K;                                          // the sources
NODE Request[1]     := 0;                                          // incoming/retrying request
NODE Orbit[K-n]     := 0;                                          // the orbit
NODE Queue[K-n]     := 0;                                          // the queue
NODE Impatient[1]   := 0;                                          // impatient request
NODE Servers[C]     := 0;                                          // the servers/channels

// Block Rental
NODE Trial[1]      := 0;                                           // block renting trial
NODE RentOrb[1]    := 0;                                           // unsuccessful block rental
NODE Blocks[m]     := 0;                                           // rented blocks
```

(3) *Functions part:* In this part (see Listing 3), we define the ServAvail function,

Listing 3. MOSEL-2 Functions part.

```
/*** FUNCTIONS **********************************************************************/

FUNC ServAvail := n+Blocks*r; // returns the number of currently available servers/channels
```

which returns the overall number of currently available channels.

(4) *Transition part:* The transition part (see Listing 4) describes how the system works. A rule may change the state by setting some nodes to a specific value or by incrementing or decrementing the values of some nodes. In our model, we have two main parts of the rules: One part describes the spectrum renting procedure and one specifies how the states of the nodes could be changed.

Listing 4. MOSEL-2 Transition part.

```
/*** RULES ***********************************************************************/

// Block Renting Subnet
IF ((ServAvail-Servers)<=t1 AND RentOrb==0) FROM EXTERN TO Trial RATE lam_r;  // trying block renting
FROM Trial TO Blocks  WEIGHT p_r;                                              // renting successful
FROM Trial TO RentOrb WEIGHT p_f;                                              // renting unsuccessful
FROM RentOrb TO Trial RATE nu_r;                                               // renting retrial
IF (ServAvail-Servers)>t1    FROM RentOrb TO EXTERN;                           // retrial interrupted
IF (ServAvail-Servers)>=t2+r FROM Blocks  TO EXTERN RATE mu_r;                 // releasing block

// Call Admission, Balking, and Retrials
FROM Sources                TO   Request RATE Sources*lambda;        // primary requests
IF (Servers  < ServAvail) FROM Request TO   Servers;                 // server available
IF (Servers == ServAvail) FROM Request TO   Orbit   WEIGHT p_o;      // no servers available; patient customer
IF (Servers == ServAvail) FROM Request TO   Sources WEIGHT p_b;      // no servers available; balking customer
IF (Servers == ServAvail) FROM Request TO   Queue   WEIGHT p_q;      // no servers available; customer pressed button
FROM Orbit                  TO   Request RATE Orbit*nu;              // retrials

// Queueing and Impatience
IF (Servers  < ServAvail) FROM Queue TO Servers;  // server available
FROM Queue     TO Impatient RATE Queue*eta;       // impatience
FROM Impatient TO Orbit     WEIGHT p_Io;          // impatient customer retries later
FROM Impatient TO Sources   WEIGHT p_Is;          // impatient customer gives up

// Service
FROM Servers TO Sources RATE Servers*mu;  // service
```

(5) *Results part:* Finally, the result part (see Listing 5) calculates the requested output performance measures. Several state dependent measures can be specified, e.g. the mean value MEAN and the probability that a component of the system is in certain states.

Listing 5. MOSEL-2 Results part.

```
/*** RESULTS *********************************************************************/

PRINT mM       := MEAN(Orbit)+MEAN(Servers)+MEAN(Queue);   // mean # active req.
PRINT mK       := K-mM;                                    // mean # active sources
PRINT ml       := mK*lambda;                               // mean throughput (served and unserved)
PRINT mlgood   := MEAN(Servers)*mu;                        // mean goodput
PRINT Pgood    := mlgood/ml;                               // prob. that arriving customer gets served
PRINT mT       := mM/ml;                                   // mean response time (served and unserved)
PRINT mB       := MEAN(Blocks);                            // mean number of rented blocks
PRINT mC       := MEAN(Servers);                           // mean number of active calls
PRINT mS       := n+mB*r;                                  // mean number of available servers
PRINT mAS      := mS-mC;                                   // mean number of idle servers
PRINT Sutil    := mC/mS;                                   // utilization of available servers
PRINT Pblock   := PROB(Servers==ServAvail);                // blocking probability
PRINT mQ       := MEAN(Queue);                             // mean number of calls in queue
PRINT mO       := MEAN(Orbit);                             // mean number of calls in orbit
@<1..m>{                                                   // probabiliy that # blocks are partly utilized
 PRINT Pb_#    := PROB((n+r*(#-1) < Servers) AND (Servers <= n+r*#));
}
PRINT mTQ      := mQ/ml;                                   // mean time spent in queue
PRINT mTO      := mO/ml;                                   // mean time spent in orbit
```