Performance analysis of finite-source retrial queues operating in random environments

J. Roszik and J. Sztrik*

University of Debrecen, P.O. Box 12, H-4010, Debrecen, Hungary
E-mail: jroszik@inf.unideb.hu E-mail: jsztrik@inf.unideb.hu
*Corresponding author

J. Virtamo

Helsinki University of Technology,
P.O. Box 3000, FIN-02015 HUT, Finland
E-mail: jorma.virtamo@hut.fi

Abstract: This paper is concerned with the performance analysis of finite-source retrial queues with heterogeneous sources operating in random environments. All random variables involved in the model construction are assumed to be exponentially distributed with a parameter depending on the source index and on the state of the corresponding random environment. The novelty of the investigation is the involvement of the random environments, which makes the system rather complicated. The MOSEL tool is used to formulate and solve the problem and the main performance measures are derived and graphically displayed.

Keywords: retrial queueing systems; finite number of heterogeneous sources; random environments; performance tool; performance measures.

Reference to this paper should be made as follows: Roszik, J., Sztrik, J. and Virtamo, J. (2007) 'Performance analysis of finite-source retrial queues operating in random environments', *Int. J. Operational Research*, Vol. 2, No. 3, pp.254–268.

Biographical notes: János Roszik received his MSc Degree in Computer Science in 2003 at the University of Debrecen. He is currently a PhD student 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 in modelling of telecommunication systems.

János Sztrik is a Full Professor at the Faculty of Informatics, University of Debrecen, Hungary. He received his MSc Degree in 1978 and his PhD in 1980 both in Probability Theory and Mathematical Statistics from the University of Debrecen. 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 the University of Debrecen in 1999, Doctor of the Hungarian Academy of Sciences, Budapest, 2002. His research interests are in the field of production systems modelling and analysis, queueing theory, reliability theory and computer science.

Jorma Virtamo is a Professor in the Networking Laboratory of the Helsinki University of Technology. He received his DSc (Tech.) in Theoretical Physics from the Helsinki University of Technology in 1976. His current research

interests include queueing theory and performance analysis of the internet, ad hoc networks and peer-to-peer networks.

1 Introduction

Queues with repeated attempts have been widely used to model many real situations in telephone registration systems, web access, call centres, telecommunication networks and computer systems. Retrial queueing systems are characterised by the feature that customers who find the server(s) busy do not wait in a queue, but repeat their requests at pre-determined or random intervals. There is a large literature devoted to retrial queues with infinite source since it has been a very active research area for about 20 years owing to mainly telecommunication and computer applications. Retrial queues, including a good bibliography, are extensively discussed in the book of Falin and Templeton (1997) and in papers Kulkarni and Liang (1996), Artalejo (1999a, 1999b) and Wu et al. (2005). Except for a few models, for example, Li and Yang (1995), Falin and Artalejo (1998), Falin (1999), Almási et al. (2004, 2005), Alfa and Isotupa (2004), Sztrik (2005) and Sztrik et al. (2006) which have considered such a system as having a finite source, most of the other models in the literature dealing with this problem have approximated it as an infinite-source model.

Finite-source retrial queueing systems appear in practice, too, as we can illustrate by the following example encountered in the ethernet system, see Alfa and Isotupa (2004). An ethernet system usually consists of a bus type of network to which several terminals are connected. The number of terminals is finite and usually not large enough to approximate the system with infinite-source models. If the terminals are connected to only one bus, then, only one terminal can send messages at a time. A terminal will receive a busy signal if it attempts to send messages when the bus is busy, and will have to retry to send the message again later. The interval between attempts depends on the protocols used in a particular ethernet system. In addition, if several terminals try to send messages at the same time there is usually a collision and only one can send its messages, the others have to wait. The method for handling the collision is called Carrier Sensing Medium Access with Collision Detection (CSMA/CD). In all cases, the unsuccessful terminal has to retry. When a terminal is waiting to retry it does not generate new messages. It simply goes into what we call an orbit where it waits to retry and becomes inactive in terms of generating new messages. This system can be represented by a retrial queueing model with finite source of customers. Another example is from the university life when the student population is of moderate size in a university or college then the telephone registration systems, which normally have multiple lines, should probably be studied as a system with finite source of customers and multiple servers. This is also true for customer call back services.

In many applications of queueing we find systems that evolve under the influence of a random environment. The random environment may model the irregularity of the arrival process (for example when there are rush hour phenomena or a periodically changing arrival stream), the irregularity of the service mechanism (owing to servers' breakdowns, servers' vacations, availability of resources, bandwidth, speeds, etc.) or both. Typical examples are queues with variable arrival or (and) service speeds. Recently, this kind of queues has been attracting considerable attention again, because of its applicability to telecommunication problems where requests arrivals and service speed (i.e., processing rate of unfinished work) are governed by a continuous-time Markov chain with finite states. Several authors have investigated the properties of such models, either in a general setting or in a specific areas such as reliability, mathematical biology, mathematical programming and queueing theory, see latest results and references in Economou (2003, 2005), Özekici and Soyer (2003a, 2003b), Bambos and Michailidis (2004), Mahabhashyam and Gautam (2005) and Takine (2005).

The majority of such models in the applied probability literature are described by stochastic processes evolving in a Markovian random environment, that is, the governing process is a Markov chain. The general model is a stochastic process $((E(t), Y(t)) : t \geq 0)$, where $E(t)$ and $Y(t)$ represent the random environment and the process of interest, respectively. The main assumption is that the evolution of $Y(t)$ does not influence the evolution of $E(t)$, while the evolution of $E(t)$ does influence the evolution of $Y(t)$ in the following way. The rates at which certain transitions in $Y(t)$ occur depend on the environmental state, thus a change in the environment might not immediately trigger a transition of $Y(t)$, but changes its dynamics (indirect interactions). Finite-source queueing systems operating in random environments have been the interest of recent research, see for example, Gaver et al. (1984), Moller and Sztrik (2001), Sztrik (2002) and Almási et al. (2001, 2003).

This paper combines the above-mentioned two lines, that is, it deals with the performance analysis of a finite-source retrial queue with heterogeneous sources operating in random environments, where the system parameters are subject to randomly occurring fluctuations. It is the unique contribution of the authors in terms of both modelling and enriching the application of the model developed since there was no paper found on this topic by searching *Zentral-blatt MATH, MathSciNet* databases. The Modelling, Specification and Evaluation Language (MOSEL) tool, see Begain et al. (2001), is used to formulate the model and to calculate the performance measures. This tool helps us to avoid the very difficult calculations owing to the large state space of the describing Markov-chain.

The aim of the present paper is to give more realistic models for finite-source retrial queues since the different request arrival, service and retrial rates are subject to random fluctuations allowing to model systems with server's breakdowns and repairs, a single server with variable service speeds, and conditional dependence of arrivals and services upon the random environment.

The rest of the paper is organised as follows. In Section 2, the mathematical description of the model and the performance measures are given, and in Section 3, some numerical examples are presented. Sections 4 and 5 are devoted to the Comments and Conclusions, respectively.

2 The queueing model

Consider a finite-source queue with K sources and a single server, where each source has different parameters and the operation of the sources and the server is influenced by the state of a given random environment.

The server and the sources are collected into M independent groups ($1 \leq M \leq K + 1$). The members of a group operate in a common random environment. The environmental changes are reflected in the values of the new and repeated call generation and in the values of the service rates. The members of group m are assumed to operate in a random environment governed by an ergodic Markov chain ($\xi_m(t); t \geq 0$) with state space ($1, \ldots, r_m$) and with transition density matrix

$$\left(\tau^{(m)}_{i_m j_m}, i_m, j_m = 1, \ldots, r_m, \tau^{(m)}_{i_m i_m} = -\sum_{k \neq i_m} \tau^{(m)}_{i_m k} \right).$$

The server can be in two states: idle and busy, and each of the sources can be in free, sending repeated calls and under service states. If source i (which is a member of group m) is free at time t and the environmental process $\xi_m(t)$ is in state j_m the probability that this source generates a new request during the time interval (t, $t + \mathrm{d}t$) is $\lambda_i(j_m)\mathrm{d}t + o(\mathrm{d}t)$, $m = 1, \ldots, M$. If the server is free at the time of arrival of a call then the call starts to be served, that is, the source moves into the under service state bringing a certain amount of work distributed exponentially with parameter μ_i hence the server moves into the busy state. Assume the server belongs to group 1 and the environmental process $\xi_1(t)$ is in state j_1 then the service speed available is $b_i(j_1)$, that is, the server can do $b_i(j_1)$ amount of work per unit time. Hence the instantaneous service completion rate is $\mu_i b_i(j_1)$ and the probability that the service is completed in time interval (t, $t + \mathrm{d}t$) is $\mu_i(j_1)\mathrm{d}t + o(\mathrm{d}t)$, where $\mu_i(j_1) = \mu_i b_i(j_1)$. In particular, when $b_i(j_1) = b(j_1)$, $i = 1, \ldots, K$, the speed does not depend on the source index but only on the state of the random environment, thus we could model the influence of the random environment on the server. If the server is busy on arrival, then the source starts generation of a Poisson flow of repeated calls with rate $v_i(j_m)$ until it finds the server free. After service the source becomes free, and it can generate a new primary call, and the server becomes idle and it can serve a new call. All random variables and the random environments are supposed to be independent of each other.

Even the model looks quite simple the solution is far from trivial. Markovian retrial queues with finite homogeneous sources without random environments have been treated in the advanced level part of Falin and Templeton (1997), and with heterogeneous sources in Almási et al. (2004) and Sztrik et al. (2006). Heterogeneous finite-source queues evolving in random environments with FIFO, Processor Sharing and Priority disciplines have been discussed in Almási et al. (2001, 2003). The proposed model combines these two directions, differs from the previous models and hence unique in the sense that to the best knowledge of the authors no other paper has been published on this system. Its relevance to the practice is quite natural since it can be applied to many situations including server's breakdowns, deteriorating server, varying arrivals, etc. It enriches the literature on finite-source queueing models with Markov-modulated arrival and service, too.

Our aim is to calculate main performance measures of the system in steady state, such as, utilisations, mean number of sources staying at the orbit or at the service, mean waiting and response times and to investigate the effect of different parameters on them.

Since the involved random variables are distributed exponentially the following process will be a continuous-time Markov chain. The state of the system at time t can be described by the process

$$X(t) = (\xi_1(t)), \ldots, \xi_M(t), \alpha(t), \beta_1(t), \ldots, \beta_{N(t)}(t)),$$

where $\xi_m(t)$ denotes the states of the background processes ($m = 1, \ldots, M$), and $N(t)$ is the number of sources of repeated calls at time t. The index of the source at the server is denoted by $\alpha(t)$, if there is a customer under service, otherwise this value is 0. Because of the heterogeneity of the sources we need to identify the sources in the sending repeated calls state, so we denote their indices by $\beta_k(t)$, $k = 1, \ldots, N(t)$, if there is a customer in this state, otherwise this last component is 0.

Since its state space is finite the process $(X(t), t > 0)$ is ergodic with the following steady-state probabilities.

$$P(j_1,\ldots,j_M,j,0) = \lim_{t\to\infty} \mathrm{P}\{\xi_1(t) = j_1,\ldots,\xi_M(t) = j_M, \alpha(t) = j, N(t) = 0\}$$

$$P(j_1,\ldots,j_M,j,i_1,\ldots,i_k) = \lim_{t\to\infty} \mathrm{P}\{\xi_1(t) = j_1,\ldots,\xi_M(t) = j_M, \alpha(t) = j, \beta_1(t) = i_1,\ldots,\beta_k(t) = i_k\}, k = 1,\ldots,K-1.$$

In the following we can derive the main characteristics in steady state. Based on the limiting probabilities it can be seen that the system performance measures are obtained as:

- *Utilisation of the server with respect to source i*

$$U_{Si} = \sum_{j_1,\ldots,j_M} P(j_1,\ldots,j_M,i,0) + \sum_{j_1,\ldots,j_M} \sum_{k=1}^{K-1} \sum_{i_1,\ldots,i_k \neq i} P(j_1,\ldots,j_M,i,i_1,\ldots,i_k),$$
$$i = 1,\ldots,K.$$

- *Utilisation of the server*

$$U_S = \sum_{i=1}^{K} U_{Si}.$$

- *Probability of source i is sending repeated calls*

$$N_i = \sum_{j_1,\ldots,j_M} \sum_{j=1}^{K} \sum_{k=1}^{K-1} \sum_{\substack{i_1,\ldots,i_k \neq j \\ i\in\{i_1,\ldots,i_K\}}} P(j_1,\ldots j_M,j,i_1,\ldots,i_k), \quad i = 1,\ldots K.$$

- *Mean number of repeated calls*

$$N = \sum_{i=1}^{K} N_i.$$

- *Utilisation of source i*

$$U_i = 1 - U_{Si} - N_i, \quad i = 1,\ldots,K.$$

- *Probability of source i is free and its background process is in state j_l*

$$F_i(j_l) = \sum_{\substack{p_1,\ldots,p_M \\ p_l = j_l}} \sum_{\substack{j=1 \\ j\neq i}}^{K} \sum_{k=1}^{K-1} \sum_{\substack{i_1,\ldots,i_k \neq j \\ i\notin\{i_1,\ldots,i_k\}}} P(p_1,\ldots,p_M,j,i_1,\ldots,i_k), \quad i = 1,\ldots,K.$$

- *Throughput of source i*

$$\gamma_i = \sum_{j_l=1}^{r_l} F_i(j_l)\lambda_i(j_l), \quad i = 1,\ldots,K.$$

- *Mean response time of source i*

$$T_i = \frac{1-U_i}{\gamma_i}, \quad i = 1,\ldots,K.$$

The traditional way is to derive the related Kolmogorov equations for the steady-state probabilities and using the normalising condition somehow we have to solve the set of equations. Usually it is not so easy, but in our case these two steps are performed by the help of the tool MOSEL as demonstrated in the next subsections.

3 Numerical examples

In this section, we present some validation results and several numerical examples are graphically displayed. The system parameters for the figures are given in Table 1. For the easier understanding only simple cases are considered. We used only one random environment with two states. The tool is able to deal with systems with several environments.

Table 1 System parameters

	K	$\lambda_1(1) \ldots \lambda_5(1)$ $\lambda_1(2) \ldots \lambda_5(2)$	$\mu_1(1) \ldots \mu_5(1)$ $\mu_1(2) \ldots \mu_5(2)$	$\nu_1(1) \ldots \nu_5(1)$ $\nu_1(2) \ldots \nu_5(2)$	$\tau_{12}^{(1)}$	$\tau_{21}^{(1)}$
Figure 1	5	x axis 1e-20	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 1e-20	0.05	0.1
Figure 2	5	x axis x axis	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 0.35, 0.4, 0.45, 0.6, 0.7	0.05	0.1
Figure 3	5	x axis 1e-20	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 1e-20	0.1	0.2
Figure 4	5	x axis $\lambda_1(1)/\lambda_2 \ldots \lambda_5(1)/2$	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 1e-20	0.1	0.2
Figure 5	5	x axis x axis	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 0.35, 0.4, 0.45, 0.6, 0.7	0.1	0.2
Figure 6	5	x axis $\lambda_1(1)/2 \ldots \lambda_5(1)/2$	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 0.35, 0.4, 0.45, 0.6, 0.7	0.1	0.2
Figures 7, 9	5	0.2 1e-20	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 1e-20	x axis	$2\tau_{12}^{(1)}$
Figures 8, 10	5	0.2 0.2	4.1, 4.3, 4.5, 4.7, 4.9 1e-20	0.35, 0.4, 0.45, 0.6, 0.7 0.35, 0.4, 0.45, 0.6, 0.7	x axis	$2\tau_{12}^{(1)}$

Figure 1 Mean response time vs. primary request generation rate

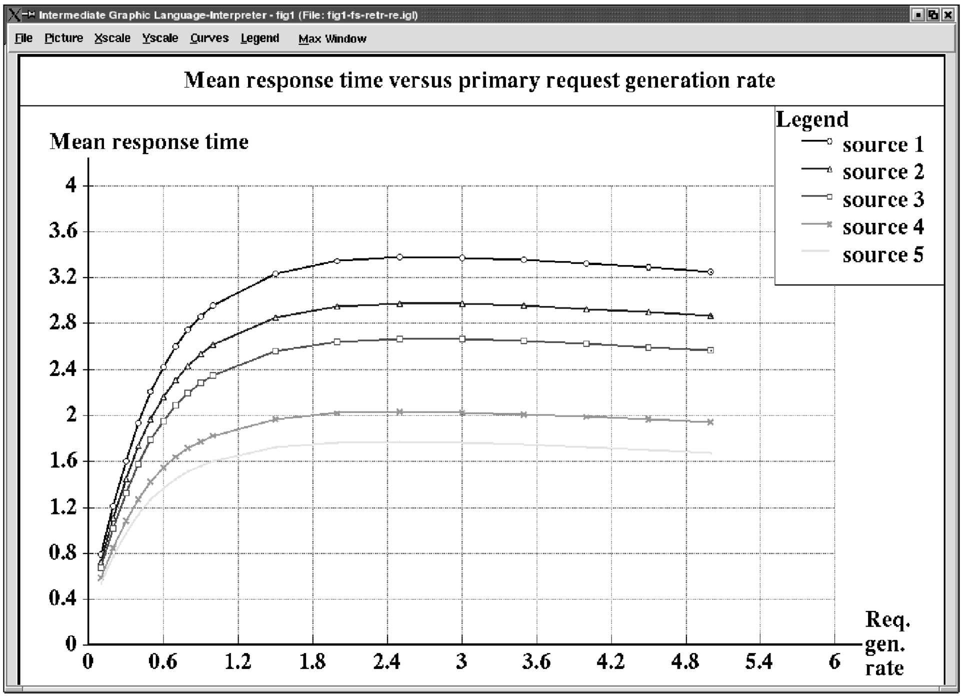


Figure 2 Mean response time vs. primary request generation rate

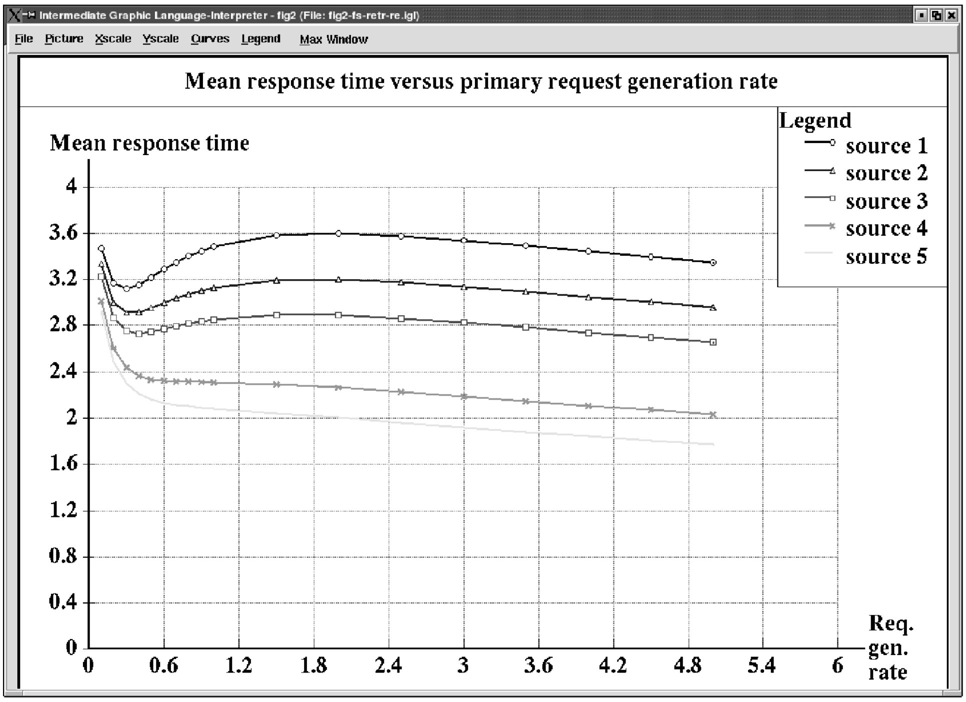


Figure 3 Mean response time vs. primary request generation rate

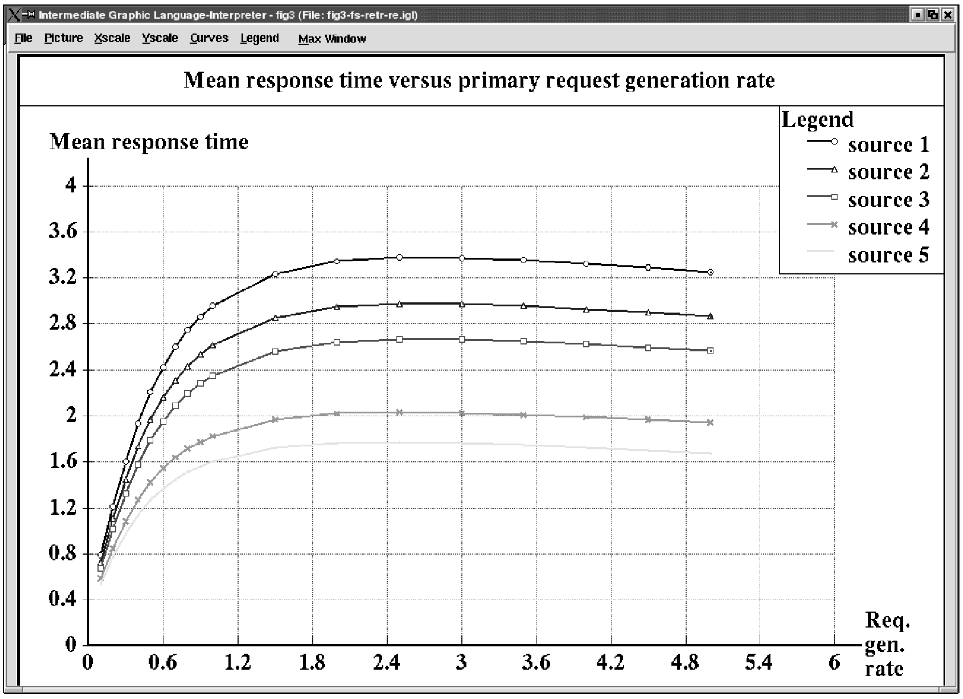


Figure 4 Mean response time vs. primary request generation rate

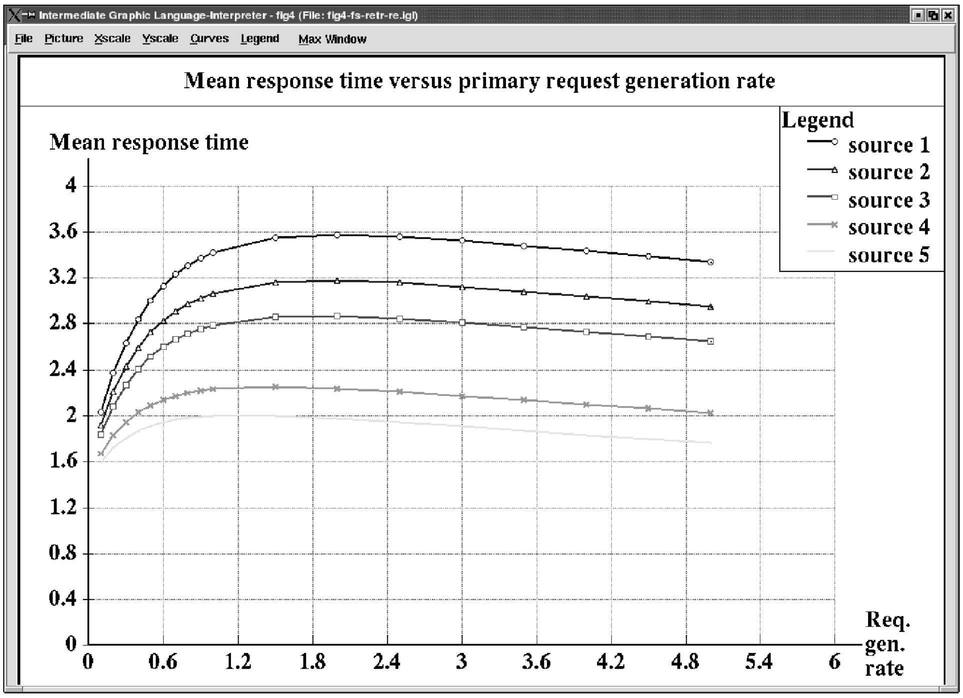


Figure 5 Mean response time vs. primary request generation rate

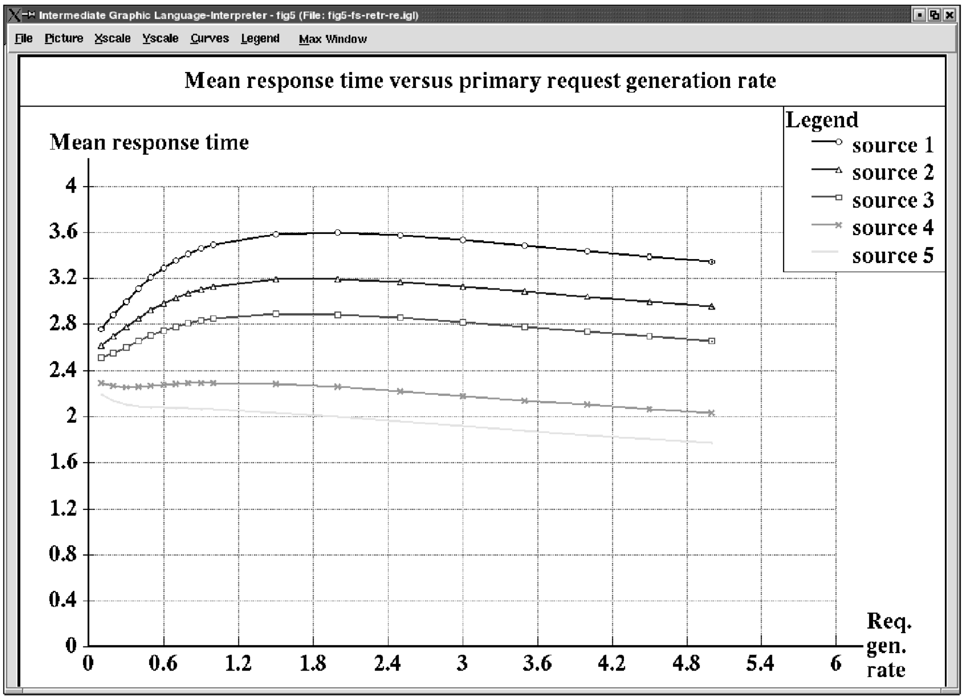


Figure 6 Mean response time vs. primary request generation rate

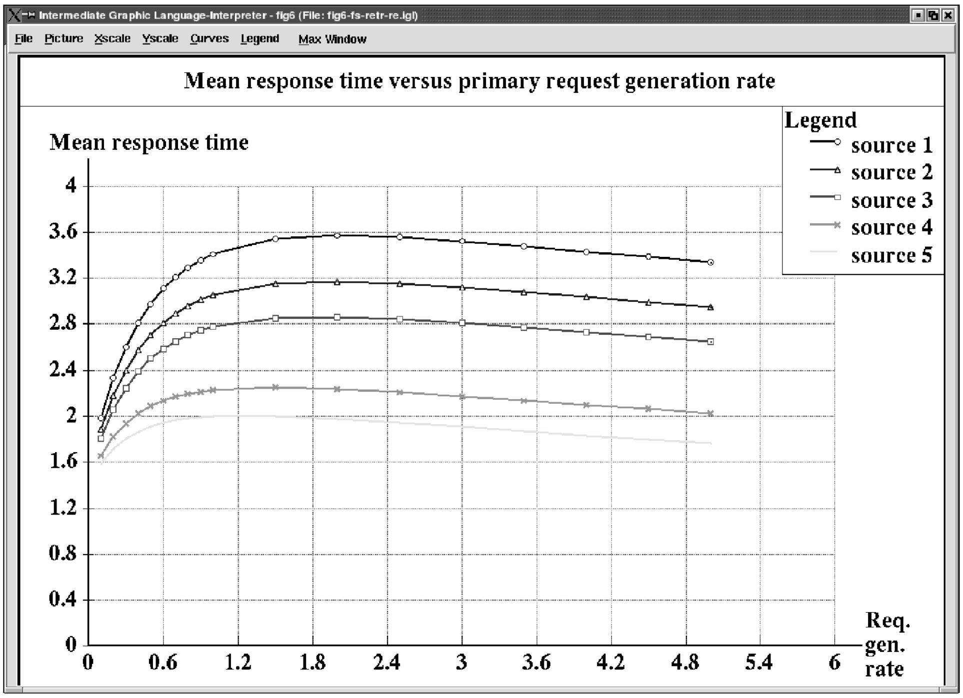


Figure 7 Mean response time vs. $\tau_{12}^{(1)}$

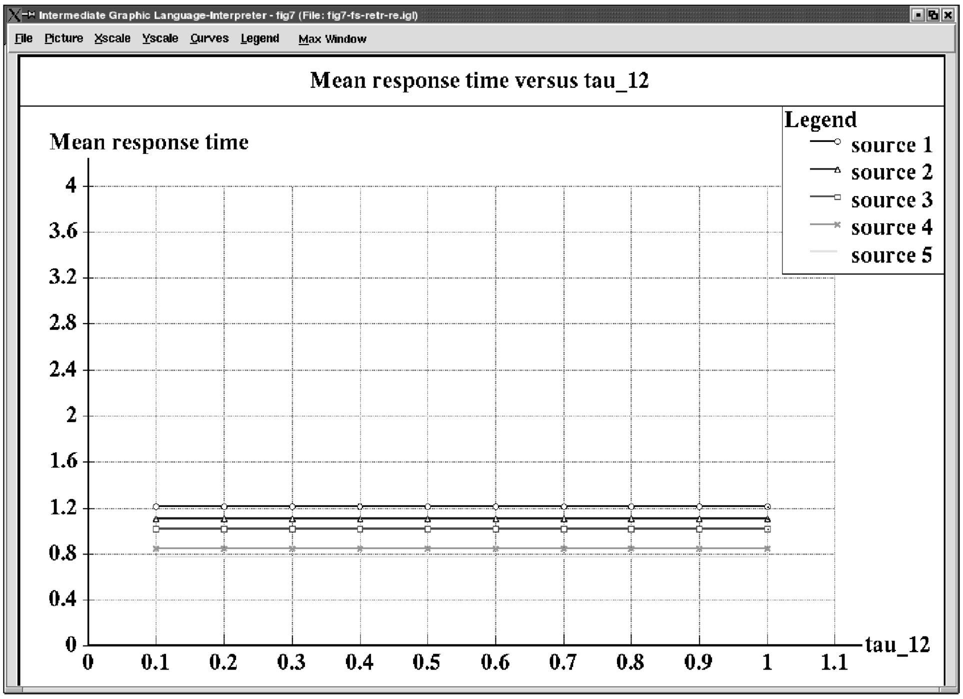


Figure 8 Mean response time vs. $\tau_{12}^{(1)}$

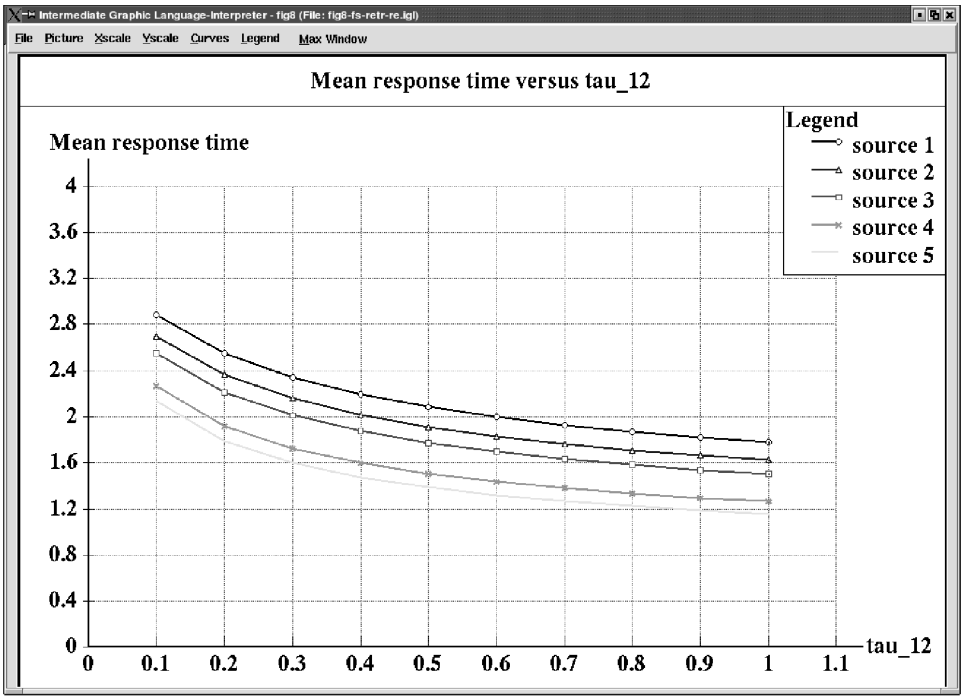


Figure 9 Mean response time vs. $\tau_{12}^{(1)}$

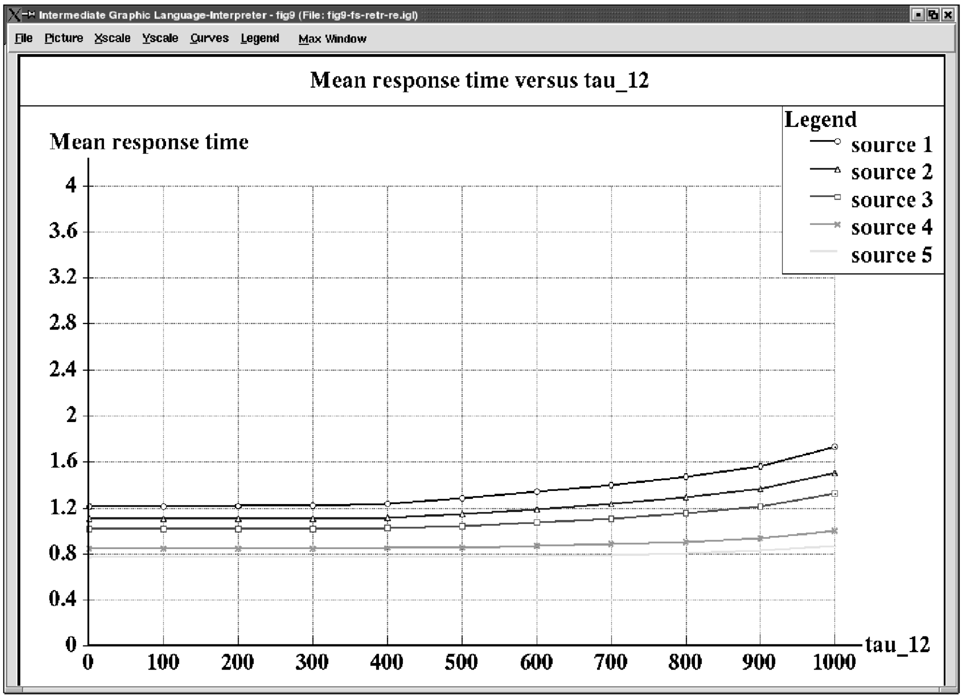
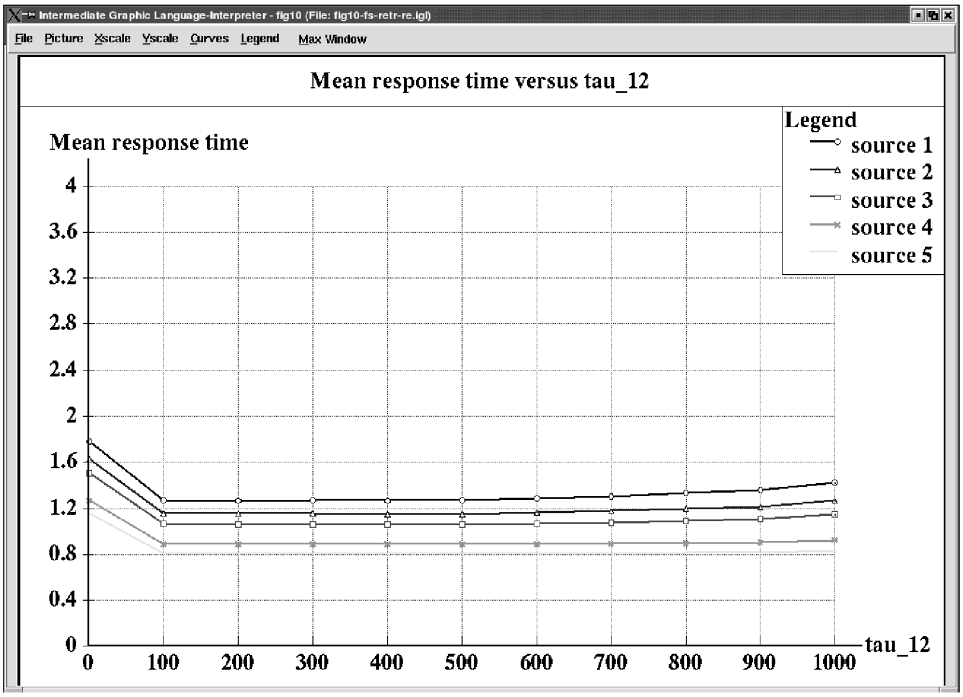


Figure 10 Mean response time vs. $\tau_{12}^{(1)}$



The calculated performance measures were validated by the results of Almási et al. (2001), where an First-Come, First-Served (FCFS) queueing model is studied. Some performance measures are collected in Table 2 and can be compared. We can see that we get back the results of the corresponding queueing model with waiting line, since with very high retrial rates the difference between the two models is negligible.

Table 2 Validation by the FCFS model

	FCFS	*Retrial*
Number of sources	5	5
Request's generation rate	0.12, 0.06	0.12, 0.06
Service rate	1, 1	1, 1
Retrial rate	–	1e+20
Rate of environment's change	0.5, 1	0.5, 1
Server queue length	0.6418	–
Requests in the orbit or in service	–	0.6418567007
Utilisation of the server	0.4342	0.4342057023
Utilisation of the sources	0.8716	0.8716286599
Mean response time	1.4782	1.4782319316

The numerical calculations were tested by the results of the retrial model with a non-reliable server investigated in Sztrik et al. (2006), too. We used the model in which the sources are blocked if the server is not operational, and the server continues servicing the interrupted call after it has been repaired. In Table 3, we can see that the results are the same. It can easily be seen that non-reliable model can be considered as a system modulated by a two-state background process. The system failure can be modelled by setting the rates to 10^{-20} in the second state of the background process.

In the following, the effect of the request arrival rate and the transition rate of the random environment on the mean response times are considered. It is easy to see that in the case of a two state governing Markov chain its steady-state probability remains the same for proportional transition rates. By the help of numerical examples we show that the mean response time depends on the transition rates of the governing Markov chain and not only on its steady-state probabilities, as we might expect. In other words, although the mean service speed is the same, the mean response times are different, see Figures 7 and 9. However, at the same time it is also observed that for a range of parameters the mean response times remain the same, see Figures 7 and 10. It means that the changes in the random environment has no effect on the response times. Of course, this region depends on the other parameters setup. These examples showed us that the performance of the systems in steady state is influenced by the changes in the random environment although in some cases and for some range of transition rates the mean response times are the same, see Figures 7 and 10.

Table 3 Validation by the non-reliable retrial model

	Non-reliable retrial	*Retrial in random env.*
Number of sources	5	5
Request's generation rate	0.10, 0.15, 0.17, 0.19, 0.21	0.10, 0.15, 0.17, 0.19, 0.21
Service rate	1.0, 1.1, 1.2, 1.5, 1.6	1.0, 1.1, 1.2, 1.5, 1.6
Retrial rate	0.15, 0.18, 0.21, 0.22, 0.25	0.15, 0.18, 0.21, 0.22, 0.25
Server's failure/repair rate	0.1, 1	–
Rate of env. change	–	0.1, 1
Utilisation of the server	0.404265368271	0.404265368271
Utilisation of the sources		
Source 1	0.673069675362	0.673069675362
Source 2	0.629766856845	0.629766856845
Source 3	0.634124904383	0.634124904383
Source 4	0.622974780687	0.622974780687
Source 5	0.627325387151	0.627325387151
Mean response time		
Source 1	5.34303316033	5.34303316033
Source 2	4.31118758383	4.31118758383
Source 3	3.73337662001	3.73337662001
Source 4	3.50379767493	3.50379767493
Source 5	3.11179039958	3.11179039958

4 Comments

- In Figures 1 and 2, the mean response time is displayed as the primary request generation rate increases. The difference between them is that in Figure 1, all operations are stopped if the background process is in the second state, but in Figure 2, only service is interrupted. The results are in agreement with the results of Sztrik et al. (2006), where the same parameters were used.
- In Figure 3, we can see that we get the same curves if we use the parameters of Figure 1, expect from the parameters $\tau_{12}^{(1)}$ and $\tau_{21}^{(1)}$, which are the double of the original value.
- In Figure 4, we suppose that the primary request generation is not stopped during server repairing. The request generation rates are the half of the rate in the first state of the background process. As we expected, we got a similar curves to the ones in Figure 3, with an increase in the mean response times.
- In Figures 5 and 6, the parameters of Figure 2 were used with the same modifications as in the case of Figures 1, 3 and 4. Contrary to Figures 1 and 3, Figures 2 and 5 are not the same.

- In Figures 7 and 8, the mean response time is displayed as $\tau_{12}^{(1)}$ increases, and $\tau_{21}^{(1)} = 2\tau_{12}^{(1)}$. The surprising difference between the curves can be explained by the following. If only service is suspended, the requests have to wait longer in the orbit if the background process changes its state slower. For this range of transition rates, the mean response times are unchanged, but this range depends on the other parameters, too.
- In Figures 9 and 10, where the same parameters were used as in Figures 7 and 8, we can see that the mean response time increases with higher $\tau_{12}^{(1)}$ values.
  This means that in both the cases, the mean response time is dependent on the values of $\tau_{12}^{(1)}$ and $\tau_{21}^{(1)}$, and not only on their proportion.

5 Conclusions

In this paper, a finite-source retrial queue with heterogeneous sources operating in random environments was considered. The numerical results were obtained by the MOSEL tool, and several examples were illustrated graphically. We analysed the differences between two non-reliable models. Furthermore, we found that the performance measures of these retrial queuing systems were dependent not only on the proportion of the state change rates of the background process, but also on the values of these rates. Furthermore, there could be some ranges of transition rates where the changes in the random environment has no significant effect on the mean response times.

Acknowledgement

The authors would like to thank the referee for valuable comments and helpful suggestions on the manuscript that led to major improvements of the paper.

Research is partially supported by Finnish–Hungarian Bilateral Scientific Cooperation SF-19/03 and Hungarian Scientific Research Fund-OTKA K60698/2005.

References

Alfa, A.S. and Isotupa, K.P.S. (2004) 'An M/P/H/k retrial queue with finite number of sources', *Computers and Operations Research*, Vol. 31, pp.1455–1464.

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

Almási, B., Bolch, G. and Sztrik, J. (2001) 'Performability modeling a client-server communication system with randomly changing parameter using MOSEL', *5th International Workshop on Performability Modeling of Computer and Communication Systems, Erlan-gen, Germany*, pp.37–41.

Almási, B., Bolch, G. and Sztrik, J. (2003) 'Analysing Markov-modulated finite-source queuing systems', *Annales Univ. Sci. Budapest., Sect. Comp.*, Vol. 22, pp.22–33.

Almási, B., Bolch, G. and Sztrik, J. (2004) 'Heterogeneous finite-source retrial queues', *Journal of Mathematical Sciences*, Vol. 121, pp.2590–2596.

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

Artalejo, J.R. (1999b) 'A classified bibliography of research on retrial: progress in 1990–1999', *TOP*, Vol. 7, pp.187–211.

Bambos, N. and Michailidis, G. (2004) 'Queueing and scheduling in random environments', *Adv. in Appl. Probab.*, Vol. 36, pp.293–317.

Begain, K., Bolch, G. and Herold, H. (2001) *Practical Performance Modeling, Application of the MOSEL Language*, Kluwer Academic Publisher, Boston.

Economou, A. (2003) 'A characterisation of product-form stationary distributions for queueing systems in random environment', *International Journal of Simulation*, Vol. 4, pp.4–11.

Economou, A. (2005) 'Generalised product-form stationary distributions for Markov chains in random environments with queueing applications', *Adv. in Appl. Probab.*, Vol. 37, pp.185–211.

Falin, G.I. (1999) 'A multiserver retrial queue with a finite number of sources of primary calls', *Mathematical and Computer Modelling*, Vol. 30, pp.33–49.

Falin, G.I. and Artalejo, J.R. (1998) 'A finite source retrial queue', *European Journal of Operational Research*, Vol. 108, pp.409–424.

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

Gaver, D.P., Jacobs, P.A. and Latouche, G. (1984) 'Finite birth-and-death models in randomly changing environments', *Advances in Applied Probability*, Vol. 16, pp.715–731.

Kulkarni, V.G. and Liang, H.M. (1996) in Dshalalow, J.H. (Ed.): *Retrial Queues Revisited, in Frontiers in Queueing*, CRC Press, Boca Raton, FL, pp.19–34.

Li, H. and Yang, T. (1995) 'A single server retrial queue with server vacations and a finite number of input sources', *European Journal of Operational Research*, Vol. 85, pp.149–160.

Mahabhashyam, S.R. and Gautam, N. (2005) 'On queues with Markov modulated service rates', *Queueing Systems*, Vol. 51, pp.89–113.

Moller, O. and Sztrik, J. (2001) 'Stochastic simulation of Markov-modulated finite-source queueing systems', *Journal of Mathematical Sciences*, Vol. 105, pp.2615–2625.

Özekici, S. and Soyer, R. (2003a) 'Network reliability assessment in a random environment', *Naval Research Logistics*, Vol. 50, pp.574–591.

Özekici, S. and Soyer, R. (2003b) 'Reliability modeling and analysis in random environments, mathematical reliability: an expository perspective', *Internat. Ser. Oper. Res. Management Set*, Vol. 67, pp.249–273.

Sztrik, J. (2002) 'Markov-modulated finite-source queueing models and their applications', *Journal of Mathematical Sciences*, Vol. 111, pp.3895–3900.

Sztrik, J. (2005) 'Tool supported performance modelling of finite-source retrial queues with breakdowns', *Publicationes Mathematicae*, Vol. 66, pp.197–211.

Sztrik, J., Almasi, B. and Roszik, J. (2006) 'Heterogeneous finite-source retrial queues with server subject to breakdowns and repairs', *Journal of Mathematical Sciences*, Vol. 132, pp.677–685.

Takine, T. (2005) 'Single-server queues with Markov-modulated arrivals and service speed', *Queueing Systems*, Vol. 49, pp.7–22.

Wang, J.T., Cao, J.H. and Li, Q.L. (2001) 'Reliability analysis of the retrial queue with server breakdowns and repairs', *Queueing Systems*, Vol. 38, pp.363–380.

Wu, X., Brill, P. and Hlynka, M. (2005) 'An M/G/1 retrial queue with balking and retrials during service', *Int. J. Operational Research*, Vol. 1, pp.30–51.