

Simulation of Finite-Source Retrial Queuing Systems With Collisions, Non-Reliable Server and Impatient Customers in the Orbit

Ádám Tóth, János Sztrik

Faculty of Informatics
University of Debrecen
Kassai Road 26, Debrecen, 4028, Hungary
toth.adam@inf.unideb.hu
sztrik.janos@inf.unideb.hu

Abstract

The goal of the paper is to study a $M/M/1//N$ finite-source retrial queuing system with collisions and customers' impatient behavior in the orbit. The server is not reliable, breakdown can happen either in busy or in idle states. The situation when an incoming customer from the orbit or from the source finds the server busy causes a collision and both requests are directed toward the orbit (including the customer under service, too). It is assumed that every request in the source is eligible to generate customers whenever the server is not working but these requests immediately get into the orbit. A customer after some waiting for the server to be served can depart from the orbit without fulfilling its service requirement these are the so-called impatient customers. In that case it goes back to the source. All random variables involved in the model construction are supposed to be independent of each other. The novelty of the investigation is to carry out a sensitivity analysis comparing various distributions of impatient time of customers on the performance measures such as mean number of customers in the orbit, mean waiting time of an arbitrary customer, mean waiting time of customers who leave the system without service, probability of abandonment, server utilization, etc. By the use of a self-developed simulation program several graphical results and comparisons of the investigated systems are illustrated.

Keywords: impatient customer, retrial queueing system, sensitivity analysis, simulation, non-reliability, collision

1. Introduction

At present with the rapid increase of traffic growth balks the analysis and optimal designing of communication systems. With regard to exchanging of information, a lot of communication process are created by companies from day to day. Consequently, construction of new applicable models of telecommunication systems or modifying of existing ones are necessary. Queuing systems with repeated calls may competently describe major telecommunication systems, such as telephone switching systems, call centers, CSMA-based wireless mesh networks in frame level. Numerous papers and books are devoted to study their importance like [19, 20]. The main feature of retrial queueing system is that customers remain in the system even if it is unable to find idle service unit and after some random time it attempts to reach the service facility again. Impatience of the customers is a natural phenomenon and an interesting topic in queueing theory. The process of renegeing and balking is extensively studied by many researchers for example in [21, 22]. Whenever an arriving customer decides not to enter the system is called balking while in renegeing a customer in the system after waiting for some time leaves the system without being served. In our investigated model renegeing customers are considered.

Speaking of communication systems where the available channels or other facilities are very limited thus users (sources) usually need to fight for these resources. This results a high possibility of conflict because several sources may launch uncoordinated attempts producing collisions. In these cases the loss of transmission takes place and it is necessary to ensure of the process of retransmission. So evolving efficient procedures for preventing conflict and corresponding message delay is essential. Some results on retrial queues with collision that have been published in [1, 2, 3, 4, 5].

In many papers of retrial queueing literature the service unit is assumed to be available steadily. But these assumptions are quite unrealistic because in real life applications of these systems can break down, different types of problems can arise like power outage, human error or other failures. Various factors have effect on the transmission rate of the wireless channel in a a wireless communication scenario and these are apt to suffer transmission failure, interruptions throughout transferring the packets. Investigating retrial queueing systems with random server breakdowns and repairs has a great importance as the operation of non-reliable systems modifies system characteristics and performance measures. Finite-source retrial queues with server breakdowns have been investigated in several papers, for example in [6, 7, 8], where the MOSEL software package (Modeling, Specification and Evaluation Language) is used or in [9] where several homogeneous servers were modeled and analyzed by the help of Generalized Stochastic Petri nets (GSPNs) using retrial systems. There are other papers which studied a finite-source retrial queue where the server is subject to breakdowns like [10, 11, 12, 13, 14].

In this paper we study the operation of the system where the service unit is subject to breakdown and waiting customers can leave the system without spending time at the service unit. The novelty of this investigation is to carry out sensitivity

analysis using assorted distributions of impatience of calls on performance measures like mean waiting time of an incoming call or total utilization of the server. To accomplish this goal a simulation program is evolved based on SimPack toolkit [15] which is a collection of C and C++ libraries. In this several different simulation algorithms are developed and provide the user with a set of utilities to build a working simulation from a model description. This is a frame and contains very basic building blocks, the model and its correct operation was coded by us. With the help of this program graphical results are given.

2. System model

A retrial queueing system of type M/M/1//N is considered with a non-reliable server and impatient customers which Figure 1 represents. In the finite-source N customers reside and each of them is able to generate calls towards the server with rate λ/N so the inter-request time is exponential with parameter λ/N . Every incoming customer has an “impatience” property which determines how much time the customer spend in the orbit without fulfilling its service requirement. Exceeding this time result that the customer no longer waits for the service unit and departs without being served properly. This variable follows gamma, hypo-exponential, hyper-exponential, Pareto and lognormal distribution with different parameters but with the same mean value. In absence of waiting queue if an incoming customer finds the server in idle state its service starts immediately. The service times of the customers is exponentially distributed with parameter μ . After its successfully service customers return to the source. Encountering the service unit in busy state the incoming customer remains in the system entering the orbit. After an exponentially distributed time with parameter σ/N customers located in the orbit make an other attempt to get into the service. The server is not reliable so from time to time it is supposed to break down. This is an exponentially distributed random variable with parameter γ_0 in case of a busy server and γ_1 when the server awaits a customer. The repair process starts immediately upon the breakdown which also follows exponential distribution with parameter γ_2 . If server failure takes place during the service of a customer then it is transferred to the orbit.

All the random variables involved in the model construction are assumed to be totally independent of each other.

3. Simulation results

Table 1 shows the set of parameters used in simulation. To obtain the desired performance measures a statistics package is involved in our simulation program, written by Andrea Francini in [16]. This class applies the method of batch mean to gather a sequence of independent samples (batch means) by aggregating n successive observations of a steady state simulation. It is one of the easiest and common technique for establishing a confidence interval for steady state mean of a process.

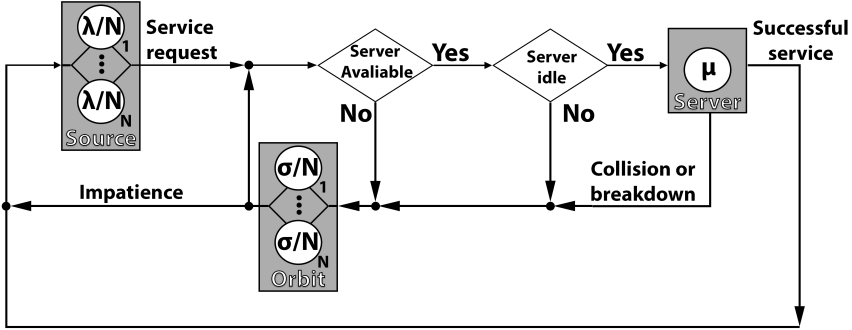


Figure 1: System model

The size of batches should be long enough to guarantee that the sample averages are approximately independent. Taking the average of the sample averages of each batch results the final mean value. More detailed information about this technique see for example [17]. The simulations are performed with the confidence level of 99.9%. Relative half-width of the confidence interval required to stop the simulation run is 0.00001.

Our aim is to examine how the different distributions of impatience of calls have an effect on the performance measure when the mean and variance are equal, Table 2 shows the parameters of distinct distributions. The parameters are chosen in such a way that the squared coefficient of variation would be greater than one. For comparison hyper-exponential, gamma, lognormal and Pareto distributions are utilized besides the case when the mean value is constant. Our simulation program is equipped with random number generators and these functions need input parameters which are different in every distribution. To achieve valid comparison in case of every distribution fitting process is needed to be done which can be viewed in [18].

N	λ/N	γ_0	γ_1	γ_2	σ/N	μ
100	0.01	0.1	0.1	1	0.01	1

Table 1: Numerical values of model parameters

Distribution	Gamma	Hyper-exponential	Pareto	Lognormal
Parameters	$\alpha = 0.390625$ $\beta = 0.0007813$	$p = 0.33098$ $\lambda_1 = 0.00132$ $\lambda_2 = 0.00268$	$\alpha = 2.1792$ $k = 270.56302$	$m = 5.57973$ $\sigma = 1.12684$
Mean	500			
Variance	640000			
Squared coefficient of variation	2.56			

Table 2: Parameters of the distribution of impatience of calls

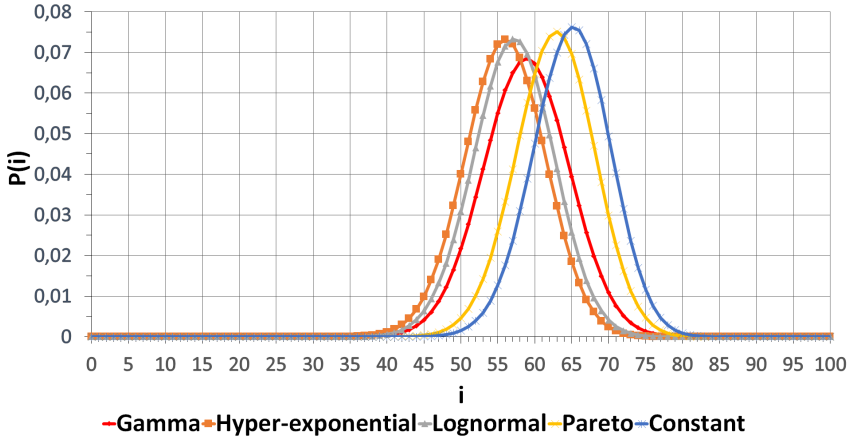


Figure 2: Comparison of steady-state distributions

Figure 2 shows the comparison of steady-state distribution of the customers. It represents the probability of how many customers residing in the system. Taking a closer look on the results all the curves correspond to normal distribution and in the following papers the proof of this phenomenon can be found in [23, 24, 25]. However, this figure clearly displays the contrast among the applied distributions. Although the shape of the curves is almost the same the average number of customers in the system varies a little bit especially in case of Pareto distribution and when the average of impatience of calls is constant the mean is more compared to the others.

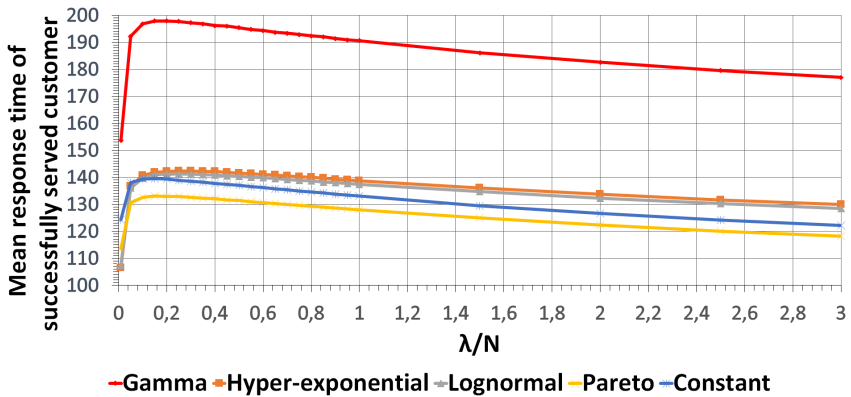


Figure 3: Mean response time vs. arrival intensity using various distributions

The mean response time of a successfully served customer is shown in function

of arrival intensity on Figure 3. Under successfully served customer we mean those customers who leaves the system through the service unit. Interestingly, differences can be observed even though the first two moments are equal, especially in case of applying gamma distribution. Results clearly illustrate the effect of various distributions. Highest values are experienced at gamma distribution. Despite the increasing arrival intensity the maximum property characteristic of finite-source retrieval queueing systems occurs under suitable parameter settings.

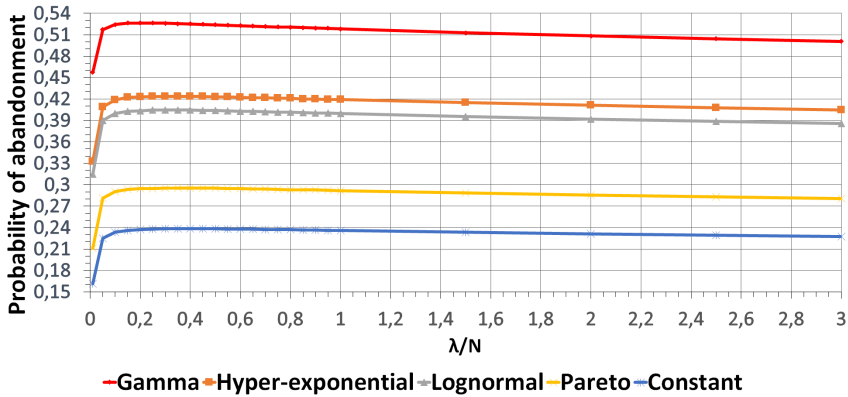


Figure 4: Probability of abandonment of a customer vs. arrival intensity using various distributions

Figure 4 demonstrate how the probability of abandonment of a customer changes with the increment of the arrival intensity. Under probability of abandonment we mean the probability of that a customer leaves the system without getting its full service requirement (through the orbit). After a slow increase of the value of this performance measure it stagnates which is true for every used distributions of impatience of calls but they differ significantly from each other. At gamma distribution the tendency of leaving the system earlier is much higher than the others especially compared to at constant mean of impatience of calls. Here the disparity is much higher among the applied distributions compared to the previous figures.

The last figure is related to the total utilization of server versus arrival intensity. Total utilization contains every service time including the interrupted ones no matter whether a call departed from the service unit or from the orbit. By examining closely the figure we find prominent results when gamma distribution is applied and regarding the others the received values are almost identical. With the increment of arrival intensity the total utilization of the service unit increases as well.

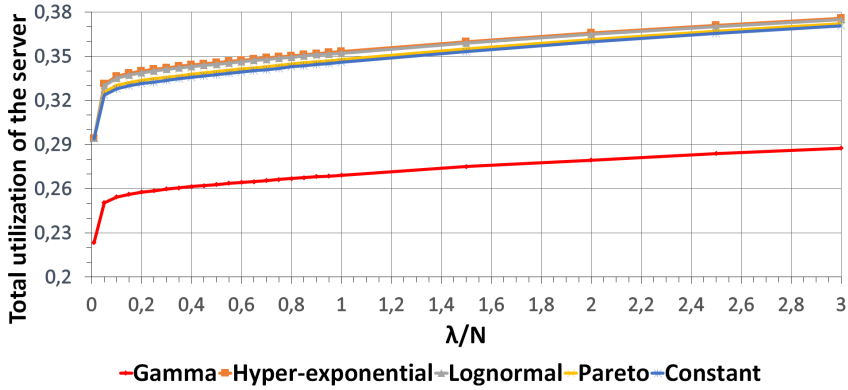


Figure 5: Total utilization of server vs. arrival intensity using various distributions

3.1. Different parameter setting

After viewing the above outcomes and figures we are intrigued to know how the operation of the system changes if another parameter setting is used. To do so we modify the parameters in order the squared coefficient of variation to be less than one so hyper-exponential is exchanged for hypo-exponential distribution. Table 3 contains the modified parameter setting of distribution of impatience of calls. Other parameters remain unchanged (see Table 1).

Distribution	Gamma	Hypo-exponential	Pareto	Lognormal
Parameters	$\alpha = 1.47059$ $\beta = 0.002941$	$\mu_1 = 0.01$ $\mu_2 = 0.0025$	$\alpha = 2.5718$ $k = 305.5844$	$m = 5.9552$ $\sigma = 0.72027$
Mean	500			
Variance	170000			
Squared coefficient of variation	0.68			

Table 3: Parameters of impatience of calls

We take over the same performance measures and figures but with the parameters of Table 3.

First, Figure 6 represents the steady state distribution. Analyzing the curves in more detail they are much closer to each other as on Figure 2. Differences appear among the applied distributions with this parameter setting, too. As regard to the values with these parameters the mean number of customer is higher in case of every distribution.

The next figure shows the mean response time of a successfully served customer in function of arrival intensity. Examining Figure 7 the same tendency can be seen as on the previous figure but differences can still be discovered especially in case of gamma distribution. This figure also reveals that customers averagely spend less time in the system compared to the previous parameter setting.

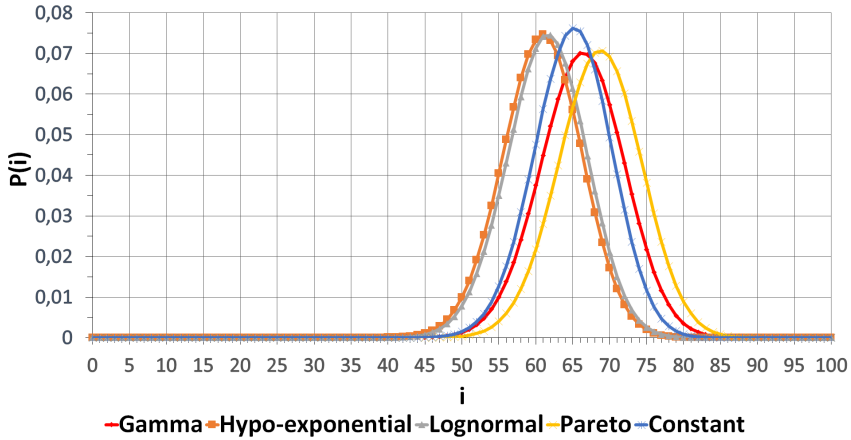


Figure 6: Comparison of steady-state distributions

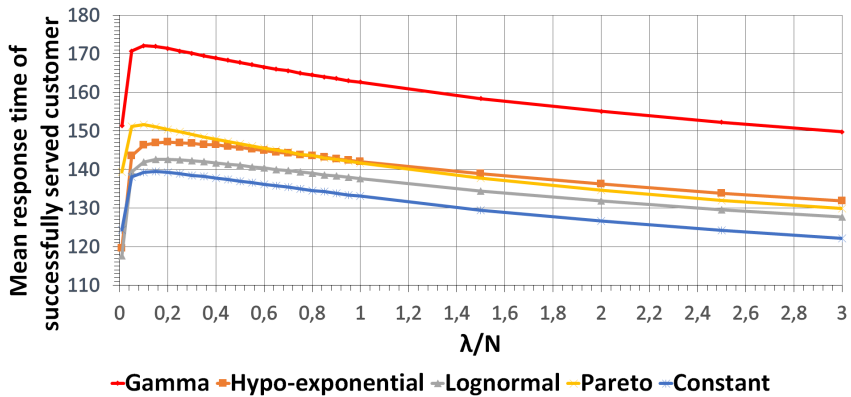


Figure 7: Mean response time vs. arrival intensity using various distributions

Figure 8 demonstrates the probability of abandonment of a customer versus arrival intensity. Not surprisingly after seeing the previous two figures the difference of achieved values are relatively far from each other, disparity is still present among the applied distributions. It can be stated that with these parameters customers are more tend to leave the system from the orbit.

Lastly, on Figure 9 the running parameter (value of x-axis) is the arrival intensity and value of y-axis is the total utilization of the server. Among the lines there are not so significant differences, they coincide with each other meaning that the utilization is almost the same except in case of Pareto and gamma distribution where the utilization of service unit is significantly less.

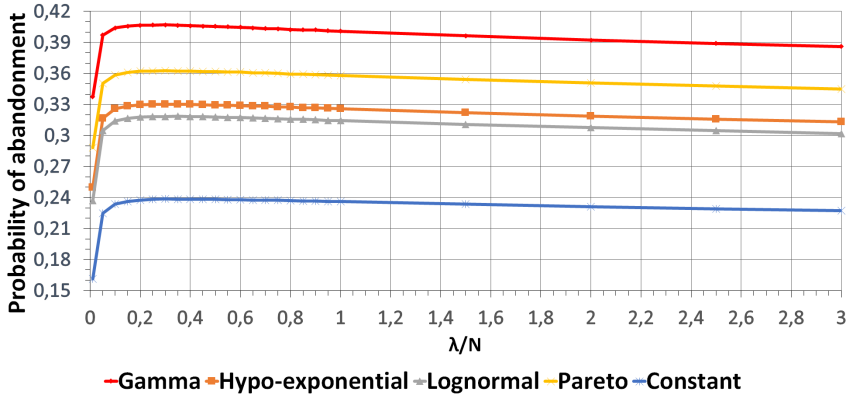


Figure 8: Probability of abandonment of a customer vs. arrival intensity using various distributions

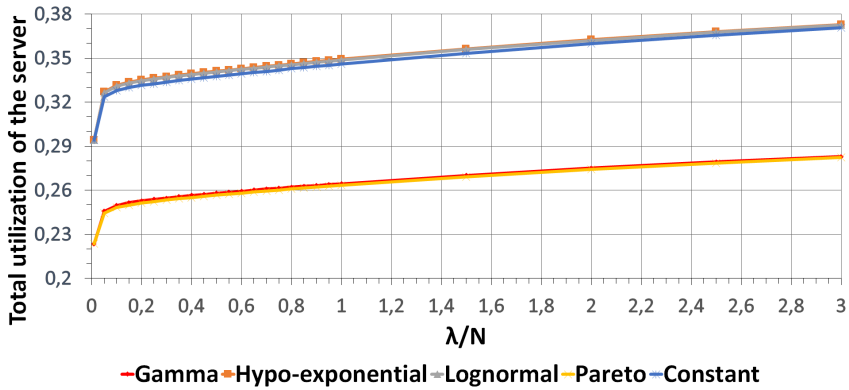


Figure 9: Total utilization of server vs. arrival intensity using various distributions

4. Conclusion

In this paper a finite-source retrial queueing system is presented with a non-reliable server, collisions and impatient customers. The obtained results fully demonstrated how pivotal selecting a distribution of impatience of calls because it has a great influence on the system characteristics despite the fact that the mean and the variance are exactly the same. Figure in connection with probability of abandonment clearly assure this phenomenon. Results evidently indicated the distinction is noticeable and significant among the performance measures having the same mean and variance of different distributions when the squared coefficient of variation is greater than one and moderate when it is less than one. In the future we would

- [10] DRAGIEVA VELIKA, I. Number of retrials in a finite source retrial queue with unreliable server, *Asia-Pacific Journal of Operational Research* Vol. 31, No. 2 (2014), 23
- [11] GHARBI, N., NEMMOUCHI, B., MOKDAD, L., BEN-OTHTMAN, J. The impact of breakdowns disciplines and repeated attempts on performances of small cell networks, *Journal of Computational Science* Vol. 5, No. 4 (2014), 633–644
- [12] KRISHNAMOORTHY, A., PRAMOD, P. K., CHAKRAVARTHY, S. R. Queues with interruptions: a survey, *TOP* Vol. 22, No. 1 (2014), 290–320
- [13] ZHANG, F., WANG, J. Performance analysis of the retrial queues with finite number of sources and service interruptions, *Journal of the Korean Statistical Society* Vol. 42, No. 1 (2013), 117–131
- [14] TÓTH, Á., BÉRCZES, T., SZTRIK, J., KVACH, A. Simulation of Finite-Source Retrial Queueing Systems with Collisions and a Non-reliable Server, In: *Vishnevskiy V., Samouylov K., Kozyrev D. (eds) Distributed Computer and Communication Networks. DCCN 2017. Communications in Computer and Information Science* Vol. 700 (2017), 146–158
- [15] FISHWICK, PAUL A. SimPack: Getting Started With Simulation Programming In C And C++, In: *Proceedings of the 24th Conference on Winter Simulation* (1992), 154–162
- [16] FRANCINI, A., NERI, F. SimPack: Getting Started With Simulation Programming In C And C++, In: *Proceedings of MASCOTS '96 - 4th International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunication Systems* (1996), 116–122
- [17] CHEN, E., KELTON, D. A Procedure for Generating Batch-Means Confidence Intervals for Simulation: Checking Independence and Normality, *Simulation* Vol. 83 (1996), 683–694
- [18] SZTRIK, J., TÓTH, Á., PINTÉR, Á., BÁCS, Z. Simulation of Finite-Source Retrial Queues with Two-Way Communications to the Orbit, In: *Dudin A., Nazarov A., Moiseev A. (eds) Information Technologies and Mathematical Modelling. Queueing Theory and Applications. ITMM 2019. Communications in Computer and Information Science* Vol. 1109 (2019), 270–284
- [19] FIEMS, D., PHUNG-DUC, D. Light-traffic analysis of random access systems without collisions, *Annals of Operations Research* Vol. 277 (2017), 311–327
- [20] ARTALEJO, J. R., GOMEZ CORRAL, A. Retrial Queueing Systems: A Computational Approach, *Springer* (2008)
- [21] KUMAR, S., SHARMA, S. Transient performance analysis of a single server queueing model with retention of renegeing customers, *Yugoslav Journal of Operations Research* Vol. 28, Issue 3 (2018), 315–331
- [22] KIM, C. S., DUDIN, S., DUDIN, A., SAMOUYLOV, K. Analysis of a Semi-Open Queueing Network with a State Dependent Marked Markovian Arrival Process, Customers Retrials and Impatience, *Mathematics* Vol. 7 (2019), 715–734
- [23] NAZAROV, A., SZTRIK, J., KVACH, A., TÓTH, Á. Asymptotic sojourn time analysis of finite-source M/M/1 retrial queueing system with collisions and server subject to breakdowns and repairs. *Ann Oper Res* Vol. 288, (2020), 417–434

- [24] NAZAROV, A., SZTRIK, J., KVACH, A., BÉRCZES, T. Asymptotic analysis of finite-source M/M/1 retrial queueing system with collisions and server subject to breakdowns and repairs. *Ann Oper Res* Vol. 277, (2019), 213–229
- [25] SUDYKO, E., NAZAROV, A., SZTRIK, J. Asymptotic waiting time analysis of a finite-source M/M/1 retrial queueing system. *Probability in the Engineering and Informational Sciences* Vol. 33 Issue 3, (2019), 387–403