Conference on
Information Technology and
Data Science

Performance Analysis of Two-Way Communication Retrial Queueing Systems With Non-Reliable Server and Impatient Customers in the Orbit*

Ádám Tóth, János Sztrik

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

Proceedings of the 1st Conference on Information Technology and Data Science
Debrecen, Hungary, November 6–8, 2020
published at http://ceur-ws.org

Abstract

Many models of two-way communication queueing systems have been studied in recent years, they can be utilized in many fields of life like in [7, 28, 30]. Customers have always been characterized by the phenomena of impatience due to the long waiting time for service [4, 14, 15, 27]. In this paper, we consider two-way communication systems with a non-reliable server where primary customers may decide to leave the system after spending a considerable amount of time in the system before getting its proper service. The service unit can break down during its operation or in an idle state, too. Whenever the server becomes idle it may generate requests towards the customers' residing in an infinite source. These requests, the so-called secondary customers, can enter the system after a random time if the service unit is available and functional upon their arrivals. Otherwise, they return to the source without coming into the system. Every primary customer has a property of impatience meaning that an arbitrary request has the ability to

*The research work of Ádám Tóth, János Sztrik was supported by the construction EFOP-3.6.3-VEKOP-16-2017-00002. The project was supported by the European Union, co-financed by the European Social Fund

quit the system after some time while its demand remains unsatisfied. During server failure, every individual may generate requests but these will be forwarded immediately towards the orbit. The source, service, retrial, impatience, operation, and repair times are supposed to be independent of each other.

Keywords: Queueing, impatience, two-way communication system, abandonment, finite-source, stochastic simulation, sensitivity analysis

1. Introduction

Nowadays due to the rapid technology development and the increase of traffic growth, the study of communication systems is crucial and inevitable especially in the topic of optimization. Not just companies or associations have some kind of networking infrastructure but our homes as well, which results in a load of communication processes. Therefore, the creation and modelling of new or existing telecommunication systems are needed. With the help of queuing systems with repeated calls are perfectly useful for modelling such arising issues in main telecommunication systems, such as telephone switching systems, call centers, or computer systems. In the last years, a great number of articles dealt with such scenarios like in [3, 9].

The fundamental characteristic of retrial queueing systems is that when every service unit is busy an incoming customer remains in the system in a virtual waiting room called the orbit. The customers who are located in the orbit try to enter the service unit after a random time.

Impatience is a natural phenomenon and models with this property are nearer to reality and lead us to more precise analysis. Because of the relatively great number of situations that can occur in healthcare applications, call centers, telecommunication networks, it is no wonder that many papers were devoted to examining the effect of impatience like [13, 18, 25]. Impatience can be interpreted in different ways: balking customers decide not to join the queue if it is too long, jockeying customers can move from its queue to another queue if they detect they will get served faster, and reneging customers leave the queue if they have waited a definite time for service. In our investigated model customers have a reneging feature.

In connection with the communication systems users (or sources) typically are coerced to fight for the available channels or facilities. The possibility of conflict is relatively high when several sources initiate random attempts on the channel producing collisions and the loss of transmissions. Developing efficient methods is important for avoiding such collisions and corresponding message delay. The effect of collisions have been published in numerous papers for example in [19–22, 24].

The examination of the availability of the service unit is vital because a lot of studies suppose that the service unit never fails and it is accessible on a permanent basis. But these assumptions are quite unrealistic and in practice, on many occasions, we find the service stations in failed state. This unfortunate behaviour has a considerable influence on the system's characteristics as well on the performance measures thus it is worth investigating such systems with server breakdowns like in the following papers [8, 12, 16, 29, 32].

Recently in a relatively high number of papers make an effort to study two-way communication designs thanks to their usefulness to model real-life examples in the various application field. This is especially true for call centers where the service unit not only processes incoming calls but also operates in idle state executing certain other works like advertising and promoting products. In such systems utilization of the service unit is always pivotal, see for example in [1, 2, 5, 7, 11, 17, 26, 31].

The novelty of the present paper is to achieve a sensitivity analysis using various distributions of service time of customers on the performance measures like the mean waiting time of an arbitrary, successfully served and impatient customer, the total utilization of the service unit, the probability of abandonment, etc. To compare the effect of the different distributions on distinct metrics a stochastic simulation program is developed based on SimPack. In this collection, you may perform various algorithms in connection with discrete event simulation, continuous simulation, and combined (multi-model) simulation. So we built a simulation model to implement every feature of the system and to calculate and estimate the desired measure using various values of input parameters. The obtained results demonstrate the importance of utilized distribution under different parameter settings represented by numerous figures and highlight some interesting specialties of these types of systems.

2. System Model

In this paper, a two-way communication retrial queuing model is considered with a non-reliable server (Figure 1). In this model, N customer populates the finite-source implying that the system will be stable at every moment. Each customer generates a request (primary customers) towards the server with rate λ/N, so the inter-request time follows an exponential distribution with parameter λ/N. Our model does not have a queue at all, thus in the case of an idle server the service of a primary customer starts immediately. The service time of these requests follows gamma, hypo-exponential, hyper-exponential, Pareto, and lognormal distribution with different parameters but with the same mean value. After departing from the system, successful requests go back to the finite-source. When the server is occupied, the incoming primary customers are forwarded to the orbit where they may retry to attain the service unit after an exponentially distributed random time with parameter σ/N. It is assumed that server failures take place during its operation or in idle state according to exponential distribution time with the rate γ_0 when it is busy and with rate γ_1 if it is idle. After server breakdown the repair process begins right away and the repair time of the server is also an exponentially distributed random variable with parameter γ_2.

Every primary customer is characterized by an impatience property and in our investigated model two different types are distinguished:

1. a primary customer after waiting for some time in the orbit leaves the system without receiving its appropriate service,
2. a primary customer after waiting for some time in the system leaves the system without receiving its appropriate service.

This decision is made after a random time which is exponential with rate τ. After the server is becoming idle it has the possibility to produce outgoing calls towards the customers (secondary) from an infinite source which is performed after an exponentially distributed period with parameter γ. The requirement of the service of these customers is that the server will not be in a failed or busy state upon their arrivals, otherwise, they are cancelled and turn back to the infinite source without entering the system. The service time of this type of customer follows a gamma distribution with parameters α_2 and β_2. Whenever the server breaks down during the service of a customer the primary ones are forwarded immediately towards the orbit and the secondary ones depart the system without continuing their service.

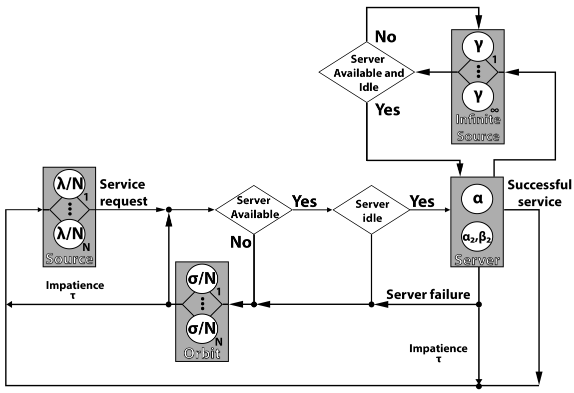

Figure 1. System model.

3. Simulation Results

In Table 1 the various values of input parameters are shown for the simulation. A statistics package is utilized to estimate the desired performance measures in our program which was developed by Andrea Francini in [10]. This code uses the method of batch means to collect a certain number of independent samples (batch means) by amassing consecutive n observations of a steady-state simulation. It is one of the most widespread and common methods to define a confidence interval for the steady-state mean of a process. To have the sample averages approximately independent the size of batches should be carefully selected. By calculating the average of the sample averages of each batch we obtain the final mean value. About this technique you can find detailed information in the following works [6, 23]. The simulations are performed with a confidence level of 99.9%. The relative half-width of the confidence interval required to stop the simulation run is 0.00001.

3.1. First Scenario

Achieving our objective we use hyper-exponential, gamma, lognormal and Pareto distributions to investigate how these distributions of service time of primary customers alter the performance measures. In the first scenario, the squared coefficient of variation is greater than one. The parameters are selected that the mean and variance would be identical and to accomplish an accurate comparison a fitting process has been done. From this process, you can find detailed information in the following paper [28].

Table 1. Numerical values of model parameters.

N	γ_0	γ_1	σ/N	γ	α_2	β_2	τ
100	0.05	0.5	0.01	0.8	1	1	0.001

Table 2. Parameters of service time of primary customers.

Distribution	Gamma	Hyper-exponential	Pareto	Lognormal
Parameters	$\alpha = 0.037$ $\beta = 0.015$	$p = 0.482$ $\lambda_1 = 0.385$ $\lambda_2 = 0.416$	$\alpha = 2.018$ $k = 1.261$	$m = -0.751$ $\sigma = 1.826$
Mean	2.5			
Variance	169			
Squared coefficient of variation	27.04			

Figure 2 shows the steady-state distribution of the number of customers in the orbit of the three investigated cases, using the two different impatient modes, and when the customers are not impatient. From the shape of the curves, it is clearly visible that the steady-state distributions of the cases are seemed to be normally distributed. With this parameter setting big differences between the two impatient modes do not develop but obviously, more customers are located averagely in the system in the case of without impatience.

Figure 3 demonstrates the observed differences between the applied distributions in the case of the steady-state distributions of the number of customers in the orbit with impatience mode 2. Regardless of the distribution, every obtained curve is similar to the normal distribution. In the case of Pareto distribution, the mean number of customers is the highest and the effect of different distributions is clearly observable.

The mean waiting time is presented in function of arrival intensity of primary customers on Figure 4, 5 and 6. First, I compared the effect of impatience. Interestingly, the results are almost identical between the impatient modes whether we talk about the mean waiting time of an arbitrary, successfully served, or impatient customer. However, pronounced differences appear when impatience is not present vs. when there it is.

After that we are curious to see how the effect of different used distributions develop, so Figure 7, 8 and 9 demonstrates that. Even though the mean and variance are the same, results clearly illustrate the effect of various distributions. The highest values are experienced in the case of Pareto distribution and the lowest in the case of the gamma distribution. With suitable parameter settings, we experience the maximum property characteristic of a finite-source retrial queueing system.

Figure 10 demonstrate how the probability of abandonment of a customer changes with the increment of the arrival intensity. By a probability of abandonment, we mean the probability 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 distribution 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 the other distributions.

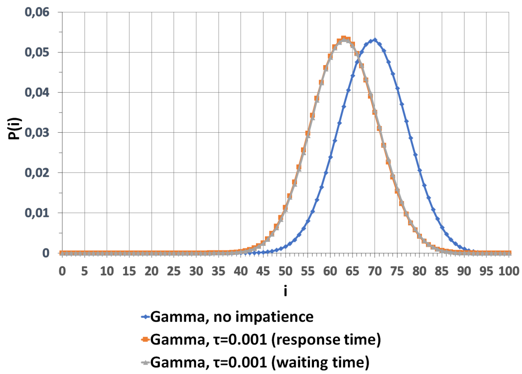

Figure 2. Comparison of steady-state distributions, λ/N=0.01.

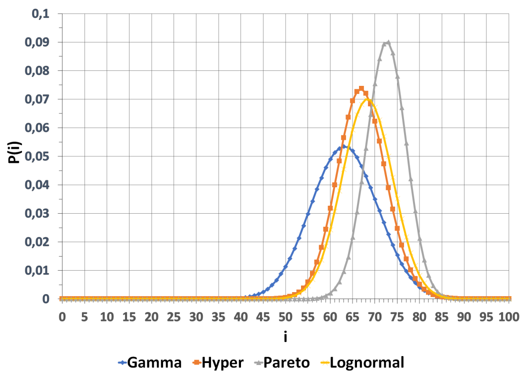

Figure 3. Comparison of steady-state distributions, λ/N=0.01.

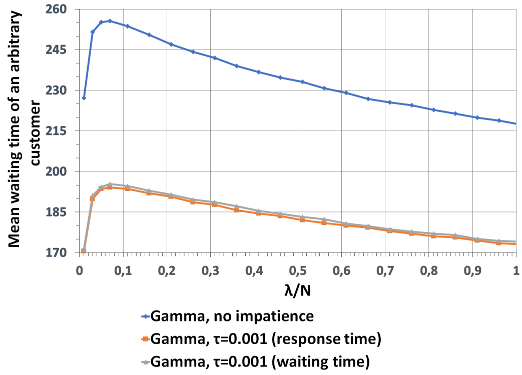

Figure 4. Mean waiting time of an arbitrary customer.

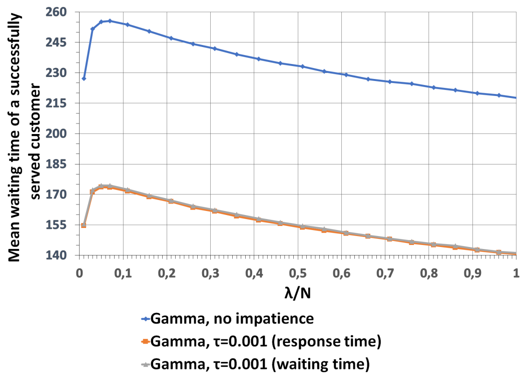

Figure 5. Mean waiting time of a successfully served customer.

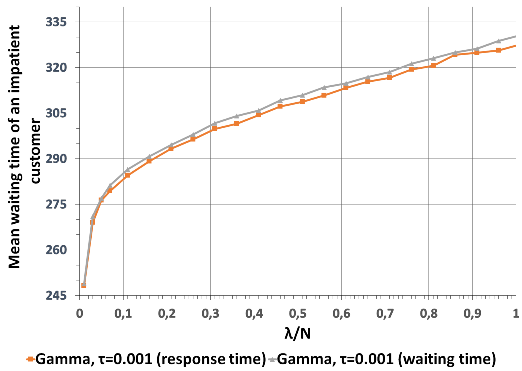

Figure 6. Mean waiting time of an impatient customer.

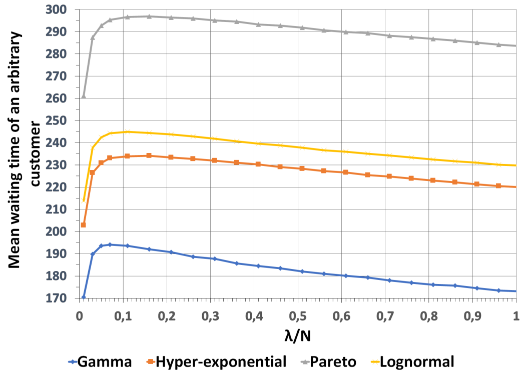

Figure 7. Mean waiting time of an arbitrary customer.

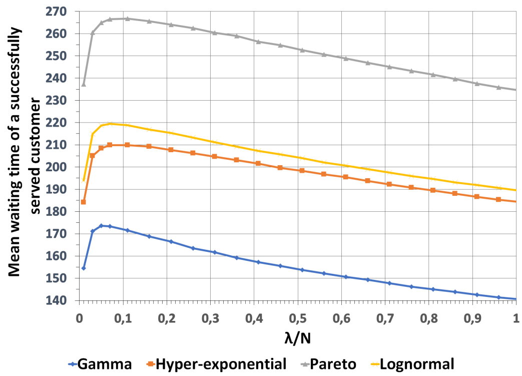

Figure 8. Mean waiting time of a successfully served customer.

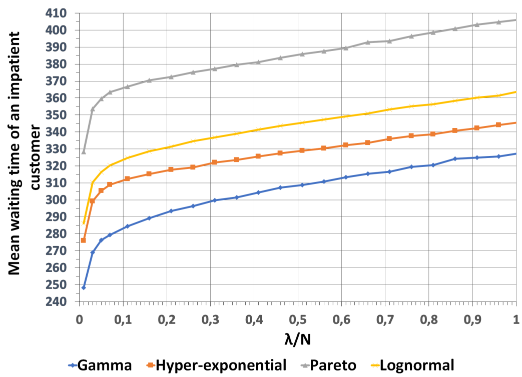

Figure 9. Mean waiting time of an impatient customer.

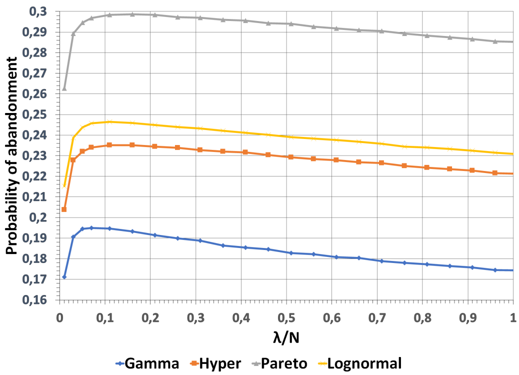

Figure 10. Comparison of probability of abandonment.

3.2. Second Scenario

The question arises whether it is true for using other parameter settings for example when the squared coefficient of variation of the service time of primary customers is less than one. We use the same parameters as in the previous section (Table 1), and Table 3 contains the changed parameters of service time of primary customers. Instead of hyper-exponential, we use this time hypo-exponential distribution because the squared coefficient of variation is always less or equal to one. We will show some figures, Figure 11 is connected to the steady-state distribution of the number of customers in the orbit. Analyzing the curves in more detail they completely overlap each other. As regards the shape of the curves they correspond to normal distribution. The mean number of customers is higher in the case of every distribution compared to the previous section.

The Figure 12 and 13 are related to the mean waiting of an arbitrary and impatient customer. As you can see very slight differences occur in the case of Pareto distribution the values are a little bit higher. With this parameter setting the interesting maximum value of the mean waiting time appears as in the previous section.

Table 3. Parameters of service time of primary customers.

Distribution	Gamma	Hypo-exponential	Pareto	Lognormal
Parameters	$\alpha = 1.8$ $\beta = 0.72$	$\mu_1 = 0.6$ $\mu_2 = 1.2$	$\alpha = 2.6733$ $k = 1.5648$	$m = 0.695374$ $\sigma = 0.6647$
Mean	2.5			
Variance	3.472			
Squared coefficient of variation	0.555			

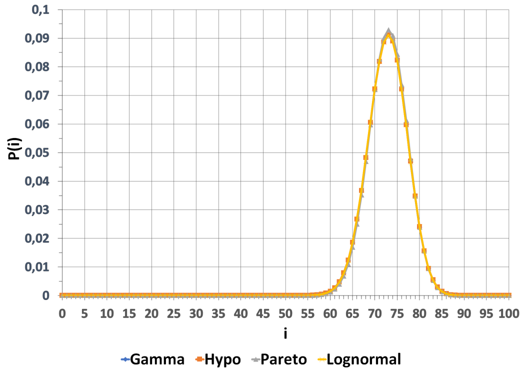

Figure 11. Comparison of steady-state distributions, λ/N=0.01.

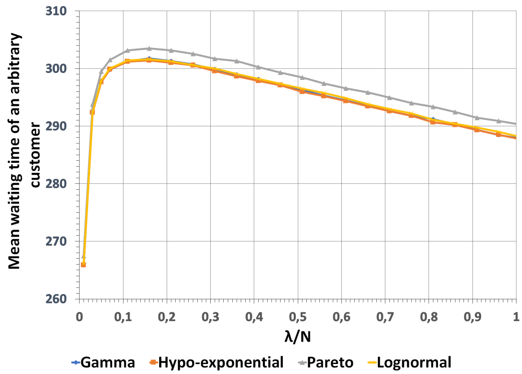

Figure 12. Mean waiting time of an arbitrary customer.

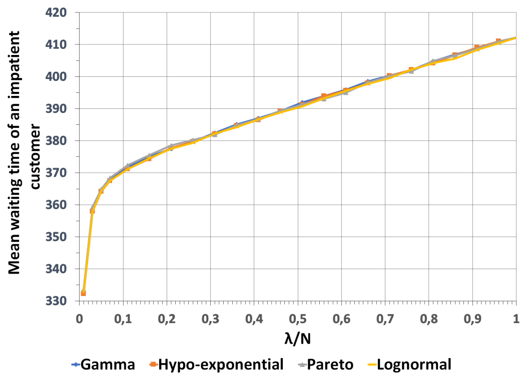

Figure 13. Mean waiting time of an impatient customer.

Figure 14 demonstrates the probability of abandonment of a customer versus arrival intensity. Not surprisingly after seeing the previous figures the difference of achieved values are relatively close to each other. It can be stated that with these parameters customers are more tend to leave the system from orbit.

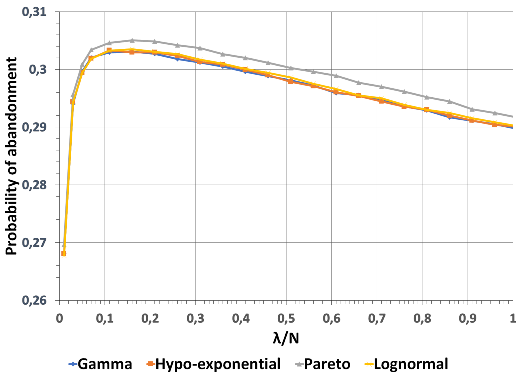

Figure 14. Comparison of probability of abandonment.

4. Conclusion

We investigated a queueing system of type $M/G/1//N$ with impatient customers in the orbit and in the system and an unreliable server capable of calling in requests from an infinite source in this paper. Simulation has been implemented, it is shown that stationary probability distribution of the number of customers in the orbit and in the system correlates to the Gaussian distribution regardless of the used distribution of service time of the primary customers. Under different scenarios, it was displayed when the squared coefficient of variation is greater than one the applied distributions of service time have a significant influence on the performance measures like the mean waiting time of an arbitrary, successfully served and impatient customers or the probability of abandonment even though the mean and variance are equal in the case of every distribution. Results also indicate that there is almost no gap between the obtained values of measures when the squared coefficient of variation is less than one. The authors plan to continue their research work, examining the obtained phenomenon in more detail and expand their model with other features like collisions, outgoing calls toward the customers from the orbit, or carrying out other sensitivity analysis on other random variables.

References

[1] S. Aguir, F. Karaesmen, O. Z. Akşin, F. Chauvet: *The impact of retrials on call center performance*, OR Spectrum 26.3 (2004), pp. 353–376.

[2] Z. Aksin, M. Armony, V. Mehrotra: *The modern call center: A multi-disciplinary perspective on operations management research*, Production and operations management 16.6 (2007), pp. 665–688.

[3] J. Artalejo, A. G. Corral: *Retrial Queueing Systems: A Computational Approach*, Springer, 2008.

[4] J. R. Artalejo, V. Pla: *On the impact of customer balking, impatience and retrials in telecommunication systems*, Computers & Mathematics with Applications 57.2 (2009), pp. 217–229.

[5] L. Brown, N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, L. Zhao: *Statistical analysis of a telephone call center: A queueing-science perspective*, Journal of the American statistical association 100.469 (2005), pp. 36–50.

[6] E. J. Chen, W. D. Kelton: *A Procedure for Generating Batch-Means Confidence Intervals for Simulation: Checking Independence and Normality*, SIMULATION 83.10 (2007), pp. 683–694.

[7] V. Dragieva, T. Phung-Duc: *Two-Way Communication $M/M/1//N$ Retrial Queue*, in: International Conference on Analytical and Stochastic Modeling Techniques and Applications, Springer, 2017, pp. 81–94.

[8] V. I. Dragieva: *Number of retrials in a finite source retrial queue with unreliable server.* English, Asia-Pac. J. Oper. Res. 31.2 (2014), p. 23.

[9] D. Fiems, T. Phung-Duc: *Light-traffic analysis of random access systems without collisions*, Annals of Operations Research 277 (2019), pp. 311–327.

[10] A. Francini, F. Neri: *A comparison of methodologies for the stationary analysis of data gathered in the simulation of telecommunication networks*, in: Proceedings of MASCOTS '96 - 4th International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, Feb. 1996, pp. 116–122.

[11] N. Gans, G. Koole, A. Mandelbaum: *Telephone call centers: Tutorial, review, and research prospects*, Manufacturing & Service Operations Management 5.2 (2003), pp. 79–141.

[12] N. Gharbi, B. Nemmouchi, L. Mokdad, J. Ben-Othman: *The impact of breakdowns disciplines and repeated attempts on performances of small cell networks*, Journal of Computational Science 5.4 (2014), pp. 633–644.

[13] N. Gupta: *Article: A View of Queue Analysis with Customer Behaviour and Priorities*, IJCA Proceedings on National Workshop-Cum-Conference on Recent Trends in Mathematics and Computing 2011 RTMC.4 (May 2012).

[14] C. Kim, S. Dudin, A. Dudin, K. Samouylov: *Analysis of a Semi-Open Queuing Network with a State Dependent Marked Markovian Arrival Process, Customers Retrials and Impatience*, Mathematics 7.8 (2019), pp. 715–734,
doi: `https://doi.org/10.3390/math7080715`.

[15] L. Kleinrock: *Creating a mathematical theory of computer networks*, Operations research 50.1 (2002), pp. 125–131.

[16] A. Krishnamoorthy, P. K. Pramod, S. R. Chakravarthy: *Queues with interruptions: a survey*, TOP 22.1 (2014), pp. 290–320.

[17] A. Kuki, J. Sztrik, Á. Tóth, T. Bérczes: *A Contribution to Modeling Two-Way Communication with Retrial Queueing Systems*, in: Information Technologies and Mathematical Modelling. Queueing Theory and Applications, Springer, 2018, pp. 236–247.

[18] R. Kumar, N. Jain, B. Som: *Optimization of an $M/M/1/N$ Feedback Queue with Retention of Reneged Customers*, Operations Research and Decisions 24 (Dec. 2014), pp. 45–58,
doi: `https://doi.org/10.5277/ord140303`.

[19] A. Kvach: *Numerical research of a Markov closed retrial queueing system without collisions and with the collision of the customers*, in: Proceedings of Tomsk State University. A series of physics and mathematics. Tomsk, vol. 295, Materials of the II All-Russian Scientific Conference, (In Russian), TSU Publishing House, 2014, pp. 105–112.

[20] A. Kvach, A. Nazarov: *Numerical research of a closed retrial queueing system* $M/GI/1//N$ *with collision of the customers*, in: Proceedings of Tomsk State University. A series of physics and mathematics. Tomsk, vol. 297, Materials of the III All-Russian Scientific Conference, (In Russian), TSU Publishing House, 2015, pp. 65–70.

[21] A. Kvach, A. Nazarov: *The research of a closed RQ-system* $M/GI/1//N$ *with collision of the customers in the condition of an unlimited increasing number of sources*, in: Probability Theory, Random Processes, Mathematical Statistics and Applications: Materials of the International Scientific Conference Devoted to the 80th Anniversary of Professor Gennady Medvedev, Doctor of Physical and Mathematical Sciences, (In Russian), 2015, pp. 65–70.

[22] A. Kvach, A. Nazarov: *Sojourn Time Analysis of Finite Source Markov Retrial Queuing System with Collision*, in: Information Technologies and Mathematical Modelling - Queueing Theory and Applications, ed. by A. Dudin, A. Nazarov, R. Yakupov, Cham: Springer International Publishing, 2015, pp. 64–72.

[23] A. M. Law, W. D. Kelton: *Simulation Modeling and Analysis*, McGraw-Hill Education, 1991.

[24] A. Nazarov, A. Kvach, V. Yampolsky: *Asymptotic Analysis of Closed Markov Retrial Queuing System with Collision*, in: Information Technologies and Mathematical Modelling: 13th International Scientific Conference, ITMM 2014, named after A.F. Terpugov, Anzhero-Sudzhensk, Russia, November 20-22, 2014. Proceedings, Cham: Springer International Publishing, 2014, chap. 1, pp. 334–341.

[25] G. Panda, V. Goswami, A. Datta Banik, D. Guha: *Equilibrium balking strategies in renewal input queue with Bernoulli-schedule controlled vacation and vacation interruption*, Journal of Industrial and Management Optimization 12 (Sept. 2015), pp. 851–878, doi: https://doi.org/10.3934/jimo.2016.12.851.

[26] S. Pustova: *Investigation of call centers as retrial queuing systems*, Cybernetics and Systems Analysis 46.3 (2010), pp. 494–499.

[27] K. Rakesh, S. Sapana: *Transient performance analysis of a single server queuing model with retention of reneging customers*, Yugoslav Journal of Operations Research 28.3 (2018), pp. 315–331,
doi: https://doi.org/10.2298/YJOR170415007K.

[28] J. Sztrik, Á. Tóth, Á. Pintér, Z. Bács: *Simulation of Finite-Source Retrial Queues with Two-Way Communications to the Orbit*, in: Information Technologies and Mathematical Modelling. Queueing Theory and Applications, ed. by A. Dudin, A. Nazarov, A. Moiseev, Cham: Springer International Publishing, 2019, pp. 270–284.

[29] Á. Tóth, T. Bérczes, J. Sztrik, A. Kvach: *Simulation of Finite-Source Retrial Queueing Systems with Collisions and a Non-reliable Server*, in: International Conference on Distributed Computer and Communication Networks, Springer, 2017, pp. 146–158.

[30] Á. Tóth, J. Sztrik: *Simulation of Finite-Source Retrial Queuing Systems With Collisions, Non-Reliable Server and Impatient Customers in the Orbit*, in: Proceedings of 11th International Conference on Applied Informatics, CEUR-WS, pp. 408–419,
url: http://ceur-ws.org/Vol-2650/#paper42.

[31] T. Wolf: *System and Method for Improving Call Center Communications*, US Patent App. 15/604,068, Nov. 2017.

[32] F. Zhang, J. Wang: *Performance analysis of the retrial queues with finite number of sources and service interruptions*, English, Journal of the Korean Statistical Society 42.1 (2013), pp. 117–131.