The Simulation of Finite-Source Retrial Queueing Systems with Two-Way Communication and Impatient Customers

János Sztrik, Ádám Tóth(✉), Ákos Pintér, and Zoltán Bács

University of Debrecen, Debrecen 4032, Hungary
{sztrik.janos,toth.adam}@inf.unideb.hu,
apinter@science.unideb.hu, bacs.zoltan@econ.unideb.hu

Abstract. The aim of the paper is to analyze a M/M/1//N finite-source, two-way communication retrial queueing system with an unreliable server and impatient customers. In this model, every request in the source is eligible to generate customers when the server does not function but they are forwarded immediately to the orbit. Customers may depart from the system during its waiting in the orbit after a random time and they get 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 failure time on the performance measures such as the mean number of customers in the orbit, the mean waiting time of an arbitrary customer, probability of abandonment, etc. With the help of self-developed simulation program, results are illustrated graphically.

Keywords: Simulation · Blocking · Two-way communication · Sensitivity analysis · Finite-source queueing system · Unreliable server · Impatient customers

1 Introduction

Nowadays, network traffic increases in such a way that the design and optimisation of communication systems are required. This phenomenon can be followed both in the industrial sector like in the companies and in our homes due to the quick technological development and the great number of devices capable of IP communication. Therefore, researchers dedicate enough time to create new suitable models of telecommunication systems or adjust the current ones.

Retrial queues play quite an important role to depict real-life problems emerging from main telecommunication systems like telephone switching systems, call centers, computer networks, and computer systems. Investigating the available literature in the Internet many papers address topics related to retrial-queuing

The research was supported by the Thematic Excellence Programme (TKP2020-IKA-04) of the Ministry for Innovation and Technology in Hungary.

V. M. Vishnevskiy et al. (Eds.): DCCN 2021, LNCS 13144, pp. 117–127, 2021.
https://doi.org/10.1007/978-3-030-92507-9_11

systems with repeated calls. In [5,6,11,13] you can see some examples of it. In many areas of science analyzing these models can improve the efficiency of systems or bring about new advantageous features for example in the case of local-area networks with random access protocols and with multiple access protocols [1,10].

Speaking of two-way communication, it possesses favorable impacts on most of the systems. Because similarities can be observed with the operation of certain real-life systems ergo it is no wonder that models based on a two-way communication scheme are introduced in many papers. This is particularly suitable in the case of call centers where the service unit (or agent) performs other actions pertaining to selling, promoting, and advertising products apart from satisfying incoming calls. In our model, the server may perform that action (calling customers residing in the orbit) after some random time when it is functional and no request is under service. Examining such scenarios has a great influence on the utilization of the service unit (or workload of agents) that is an important aspect and extensively examined by several papers like [3,12].

Studying the related articles I found the assumption of having a service unit available all the time which is quite impractical regarding events in real-life applications of systems for example power outages, human error, or other failures. Although companies, providers want to ensure having fault-tolerant devices and services (the intention is to have high-availability scenarios), problems can occur at any time. Not to mention wireless communication where other factors could affect the transmission rate of the wireless channel and the forwarded information prone to undergo failure interruptions throughout transferring the packets. That is why random server breakdowns and repairs are centric topics so the inspection of these features alters undoubtedly the operation of systems, the system characteristics, and the performance measures. Finite-source retrial queues with server breakdowns have been studied in several papers like [4,9,11,19,20].

The main aim of this work is to investigate the operation of such a system containing a non-reliable service unit and customers which may leave the system without obtaining their service needs. The novelty of this investigation is to carry out a sensitivity analysis using different distributions of failure time on performance measures like the mean waiting time of an arbitrary customer, a customer leaving the system through the orbit, or the total utilization of the server. A simulation program is developed to accomplish our goal namely checking the effects of the distributions. Our program is based on SimPack toolkit [7] which is a collection of C and C++ libraries. Several approaches and algorithms are supported providing a set of utilities to build a working simulation from a model description. Simpack contains very basic building blocks, during the coding of the model several functions, random number generator, and features were integrated. With the help of this program, results are illustrated graphically. This paper is the natural continuation of [17].

2 Model Description and Notations

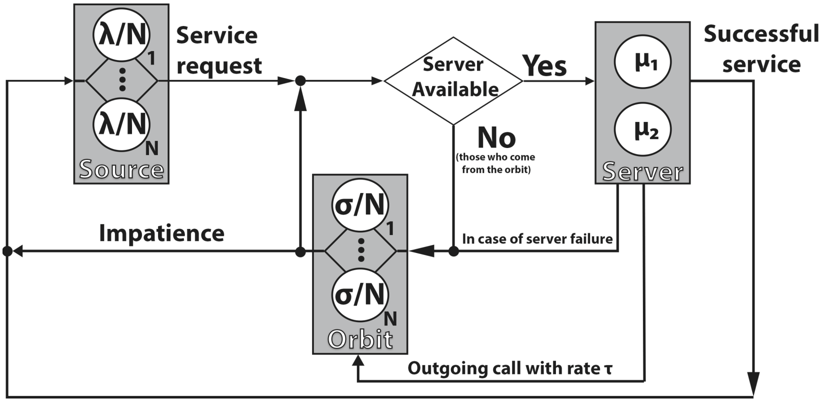

Fig. 1. System model

Figure 1 demonstrates the considered retrial queueing system of type $M/M/1//N$ which contains two-way communication feature and impatient customers. N customers are located in the finite-source where each of them can generate calls towards the server according to an exponential distribution with rate λ/N. In this model, every customer is characterized by an impatience feature that determines the maximum spent time of a customer in the orbit before leaving the system without completing its service requirement. This random variable also follows an exponential distribution with parameter τ. The model does not contain waiting queues therefore if the service unit is idle the service of an incoming customer starts immediately which is exponentially distributed with parameter μ. Upon its completion, request goes back to the source. Otherwise, the incoming customer is delivered to the orbit remaining in the system and after an exponentially distributed time with parameter σ/N they launch another attempt to reach the service facility. Our assumption is that every now and then the server breaks down according to gamma, hypo-exponential, hyper-exponential, Pareto, and lognormal distribution with different parameters but with the same mean value.

Throughout this period customers may proceed to produce their requests but they are transferred to the orbit immediately. The repair process is initiated instantaneously upon the failure of the server, which is also an exponentially distributed random variable with parameter γ_2. When the server breaks down during the service of a customer the execution will be cancelled and the customer returns to the orbit instantly. The feature of two-way communication is when the server becomes idle it may accomplish an outgoing call (secondary customers) after an exponentially distributed random time with rate ν that results in calling a customer in the orbit earlier. The service of these customers follows an exponential distribution with rate μ_2. Rates λ/N and σ/N are used because in [15,16] very similar systems are evaluated by an asymptotic method where N tends to infinity, and was proved that the number of customers in the system follows a normal distribution. All the random variables in the model creation are assumed to be totally independent of each other.

3 Simulation and Results

As mentioned earlier SimPack is the base of our simulation program which consists of a statistic package [8]. The method of batch means is applied and with the help of this tool, it is possible to perform a quantitative estimation of the mean and variance values of the desired variables. The fundamental operation of this method is that in every batch n observations take place and the useful run is divided into numerous batches. For having a valid and correct estimation the batches should be long enough and approximately independent of each other. This is one of the most popular mechanisms among the confidence interval techniques for a steady-state mean of a process. The following works [2,14] comprise very precise description and algorithm about batch means. 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 Scenario 1

Four different distributions of failure time are used to investigate their effects on the main performance measures. To have a valid comparison we selected the parameters in such a way that the mean value and variance would be equal. Before that, a fitting process is necessary to be done to obtain the correct values of parameters. [18] describes in more detail the characteristics of the utilized distributions. In the first scenario, the squared coefficient of variation is greater than one therefore we utilized hyper-exponential, gamma, Pareto, and lognormal distributions and compared them with each other. Table 1 and Table 4 shows every values of the random variables including all the used input parameters of the various distributions of failure time as well (Table 2).

Table 1. Used numerical values of model parameters

N	λ/N	γ_2	σ/N	μ	μ_2	ν
100	0.01	1	0.01	1	1.2	0.02

Table 2. Parameters of failure time

Distribution	Gamma	Hyper-exponential	Pareto	Lognormal
Parameters	$\alpha = 0.6$ $\beta = 0.5$	$p = 0.25$ $\lambda_1 = 0.41667$ $\lambda_2 = 1.25$	$\alpha = 2.2649$ $k = 0.67018$	$m = -0.3081$ $\sigma = 0.99037$
Mean	1.2			
Variance	2.4			
Squared coefficient of variation	1.6666666667			

Figure 2 shows the mean waiting time of an arbitrary customer in the function of arrival intensity. The disparity is quite obvious taking a closer look at the figure that represents the impact on the metrics using various distributions having the same first two moments. Customers spend by far more time in the orbit at Pareto distribution and the least at gamma distribution. Also the interesting maximum property characteristic of a finite-source retrial queueing system occurs despite the increasing arrival intensity.

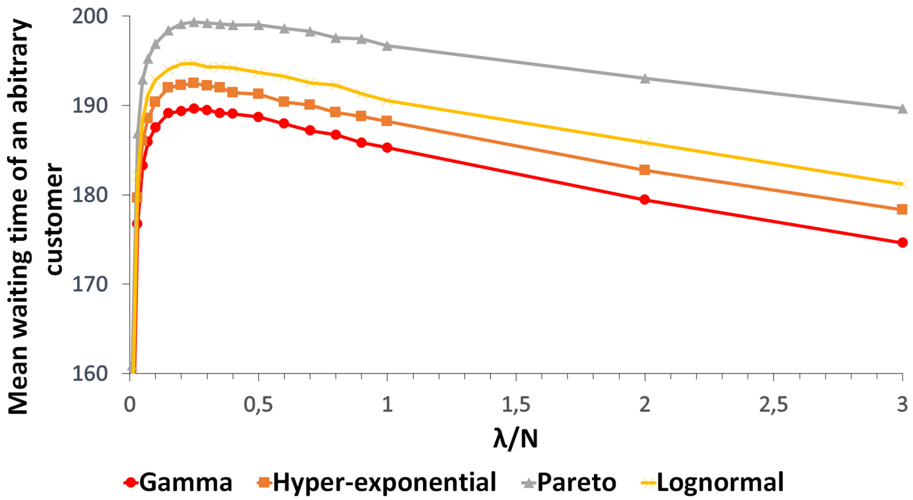

Fig. 2. Mean waiting time of an arbitrary customers

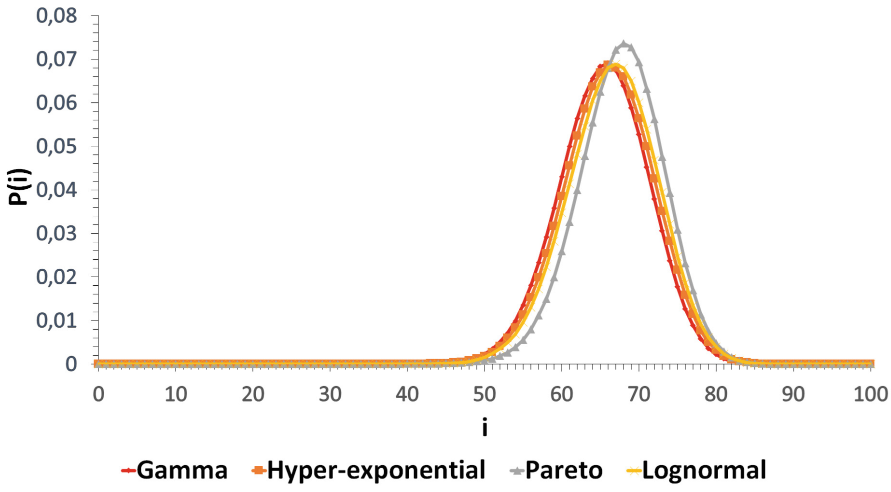

Fig. 3. Distribution of the number of customers in the system, $\lambda/N = 0.01$

Seeing the differences in the previous figure, we wondered what could be the situation of steady-state distribution of the number of customers which is displayed in Fig. 3. In the graph curves are close to each other, basically results of Pareto distribution are separate the others are almost the same. However, the shape of each of them is identical assuming that they follow a normal distribution.

Figure 4 highlights the property of impatience under different parameter setting showing how the mean waiting of an arbitrary customer develops beside increasing arrival intensity. Actually the expected behaviour happens namely as the probability of leaving the system earlier increases fewer customers will be located in the system. This is logical and the results confirm our suspicion. However, impatience does not change the maximum property characteristic, it is clearly visible that every curve has a maximum value (Fig. 5).

The last figure in this section is about the utilization of the service unit by outgoing customers. This includes all the time spent serving clients called by the idle server. By examining closely the figure we find lower values at Pareto distribution meaning that fewer number of outgoing customers are under service and regarding the others the received values are near to each other. With the increment of arrival intensity the utilization of the service unit by outgoing customers increases as well but after 0.01 it slowly decreases which is true for every investigated cases.

3.2 Scenario 2

After analysing the results of Scenario 1, we wondered if the same phenomena that occurred in the previous section would also show up with other parameter settings. Here the parameters of each distribution, which can be seen in Table 4, have been chosen so that the squared coefficient of variation would be less than one. Instead of hyper-exponential distribution we utilize hypo-exponential distribution to carry out a sensitivity analysis. The other parameters remain untouched (see Table 3), only the parameters of failure time differ between the two scenarios.

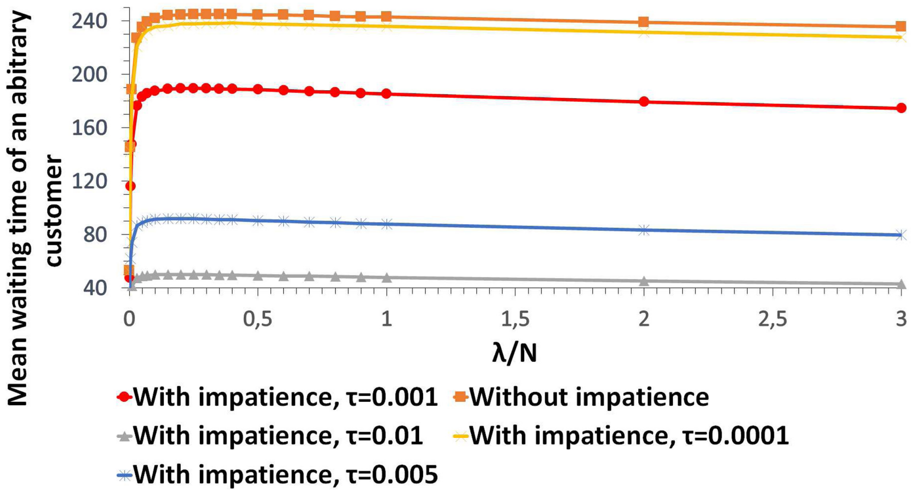

Fig. 4. The effect of impatience on the mean waiting time

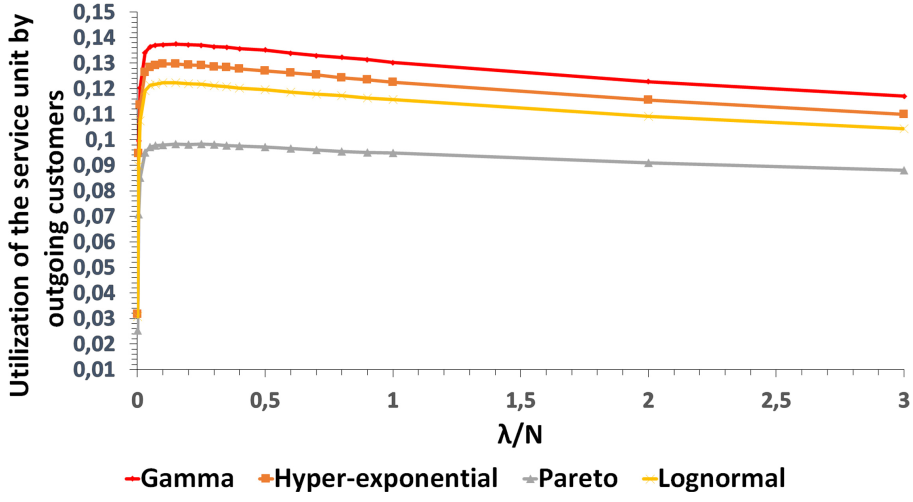

Fig. 5. Utilization of the service unit by outgoing customers

Table 3. Used numerical values of model parameters

N	λ/N	γ_2	σ/N	μ	μ_2	ν
100	0.01	1	0.01	1	1.2	0.02

Table 4. Parameters of failure time

Distribution	Gamma	Hypo-exponential	Pareto	Lognormal
Parameters	$\alpha = 1.3846$ $\beta = 1.1538$	$\mu_1 = 1$ $\mu_2 = 5$	$\alpha = 2.5442$ $k = 0.7283$	$m = -0.08948$ $\sigma = 0.7373$
Mean	1.2			
Variance	1.04			
Squared coefficient of variation	0.72222222			

To truly see the difference between the two scenarios let's first look at the mean waiting time of an arbitrary customer which is demonstrated by Fig. 6. The achieved results are nearly identical no significant differences appear even in the case of Pareto distribution. Otherwise, which is similar to Fig. 2 that the mean waiting time has maximum value. This is a common phenomena of retrial queuing systems having finite number of customers in the source.

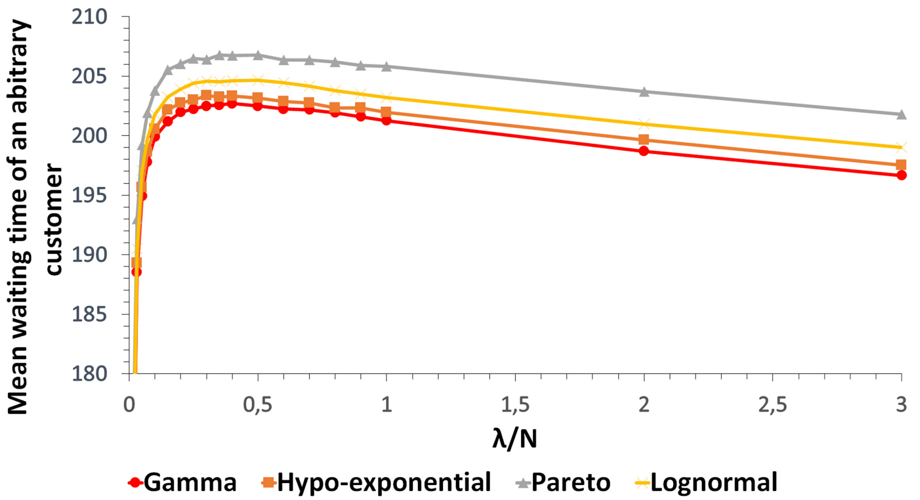

Fig. 6. Mean waiting time of an arbitrary customer

The next Figure (Fig. 7), as in the previous scenario, exhibits the distribution of the number of customers in the system with $\lambda/N = 0.01$ so what is the probability ($P(i)$) having exactly i customer in the system. Compared to Fig. 3, the curves are almost totally identical even in the case of Pareto distribution

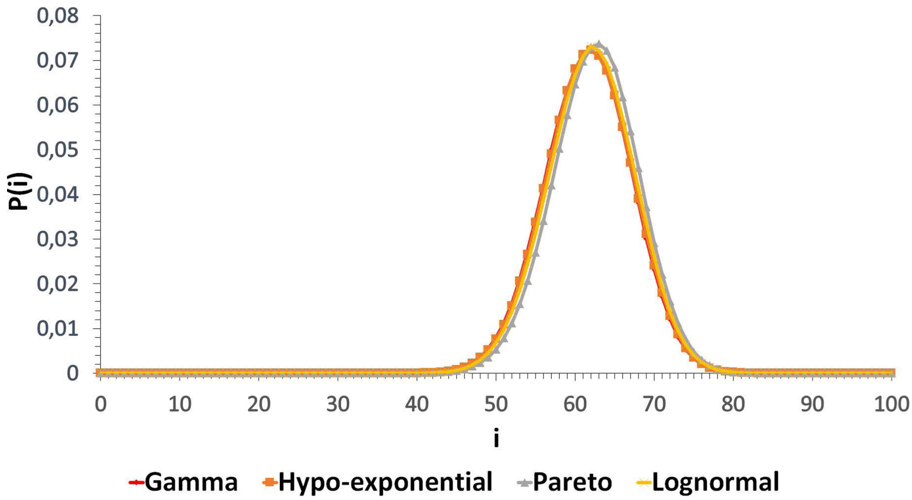

Fig. 7. Distribution of the number of customers in the system, $\lambda/N = 0.01$

and it can be also said that from the look of them they might follow normal distribution under this parameter setting as well.

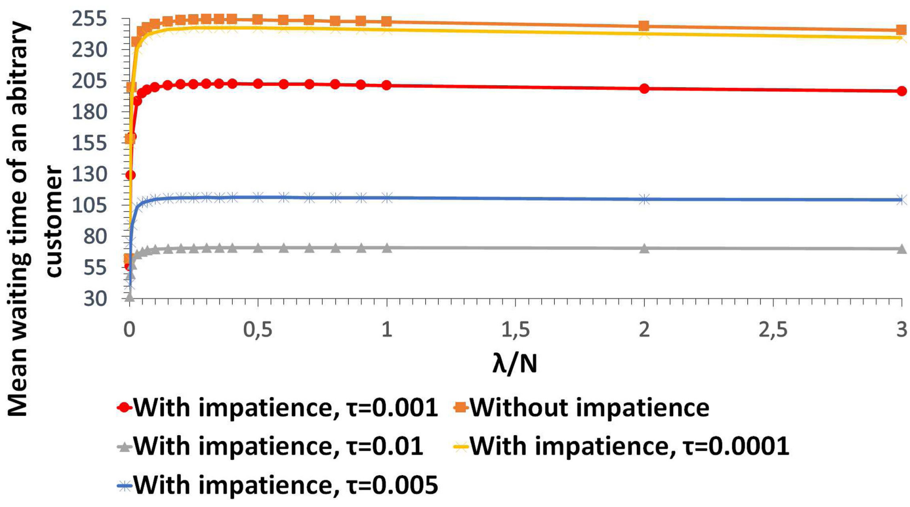

Fig. 8. The effect of impatience on the mean waiting time

To emphasize the effect of impatience Fig. 8 illustrates the mean waiting of an arbitrary customer using different values of impatience intensities. The obtained results are not surprising as this intensity grows the customers tend to spend less time in the system averagely which is quite logical. In comparison to Fig. 4, the tendency of difference seems to be the same, but the values in this scenario are higher.

4 Conclusion

We presented a finite-source retrial queueing system with an unreliable server that may call in requests residing in the orbit (two-way communication property) and impatient customers. The obtained results demonstrate the effect of impatience on the visualized performance measures indicating that customers with less patience depart much earlier without reaching the service facility. Results also display the influence of various distributions of failure time on the performance measures when the squared coefficient of variation was greater than one despite the fact that mean and variance are equal. In the case of less than one significant differences do not appear, curves almost totally overlap each other. In the future we plan to complete that system including other features like trying out other distributions or introducing disaster failure, or including more capacity of service.

References

1. Artalejo, J., Corral, A.G.: Retrial Queueing Systems: A Computational Approach. Springer, Heidelberg (2008)
2. Chen, E.J., Kelton, W.D.: A procedure for generating batch-means confidence intervals for simulation: checking independence and normality. Simulation **83**(10), 683–694 (2007)
3. Dragieva, V., Phung-Duc, T.: Two-way communication M/M/1//N retrial queue. In: Thomas, N., Forshaw, M. (eds.) ASMTA 2017. LNCS, vol. 10378, pp. 81–94. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61428-1_6
4. Dragieva, V.I.: Number of retrials in a finite source retrial queue with unreliable server. Asia-Pac. J. Oper. Res. **31**(2), 23 (2014). https://doi.org/10.1142/S0217595914400053
5. Falin, G., Artalejo, J.: A finite source retrial queue. Eur. J. Oper. Res. **108**, 409–424 (1998)
6. Fiems, D., Phung-Duc, T.: Light-traffic analysis of random access systems without collisions. Ann. Oper. Res. **277**(2), 311–327 (2017). https://doi.org/10.1007/s10479-017-2636-7
7. Fishwick, P.A.: Simpack: getting started with simulation programming in C and C++. In: In 1992 Winter Simulation Conference, pp. 154–162 (1992)
8. Francini, A., Neri, F.: 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, pp. 116–122 (Feb 1996)
9. Gharbi, N., Nemmouchi, B., Mokdad, L., Ben-Othman, J.: The impact of breakdowns disciplines and repeated attempts on performances of small cell networks. J. Comput. Sci. **5**(4), 633–644 (2014)

10. Kim, J., Kim, B.: A survey of retrial queueing systems. Ann. Oper. Res. **247**(1), 3–36 (2015). https://doi.org/10.1007/s10479-015-2038-7
11. Krishnamoorthy, A., Pramod, P.K., Chakravarthy, S.R.: Queues with interruptions: a survey. TOP **22**(1), 290–320 (2012). https://doi.org/10.1007/s11750-012-0256-6
12. Kuki, A., Sztrik, J., Tóth, Á., Bérczes, T.: A contribution to modeling two-way communication with retrial queueing systems. In: Dudin, A., Nazarov, A., Moiseev, A. (eds.) ITMM/WRQ -2018. CCIS, vol. 912, pp. 236–247. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97595-5_19
13. Kumar, B.K., Vijayalakshmi, G., Krishnamoorthy, A., Basha, S.S.: A single server feedback retrial queue with collisions. Comput. Oper. Res. **37**(7), 1247–1255 (2010)
14. Law, A.M., Kelton, W.D.: Simulation Modeling and Analysis. McGraw-Hill Education (1991)
15. Nazarov, A., Sztrik, J., Kvach, A.: A survey of recent results in finite-source retrial queues with collisions. In: Dudin, A., Nazarov, A., Moiseev, A. (eds.) ITMM/WRQ -2018. CCIS, vol. 912, pp. 1–15. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-97595-5_1
16. 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. **277**(2), 213–229 (2018). https://doi.org/10.1007/s10479-018-2894-z
17. Sztrik, J., Tóth, Á., Pintér, Á., Bács, Z.: The simulation of finite-source retrial queueing systems with two-way communications to the orbit and blocking. In: Vishnevskiy, V.M., Samouylov, K.E., Kozyrev, D.V. (eds.) DCCN 2020. CCIS, vol. 1337, pp. 171–182. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-66242-4_14
18. Toth, A., Sztrik, J., Kuki, A., Berczes, T., Efrosinin, D.: Reliability analysis of finite-source retrial queues with outgoing calls using simulation. In: 2019 International Conference on Information and Digital Technologies (IDT), pp. 504–511, June 2019
19. Tóth, A.., Bérczes, T.., Sztrik, J.., Kuki, A.., Schreiner, W..: The simulation of finite-source retrial queueing systems with collisions and blocking. J. Math. Sci. **246**(4), 548–559 (2020). https://doi.org/10.1007/s10958-020-04759-4
20. Zhang, F., Wang, J.: Performance analysis of the retrial queues with finite number of sources and service interruptions. J. Korean Stat. Soc. **42**(1), 117–131 (2012). https://doi.org/10.1016/j.jkss.2012.06.002