The Simulation of Finite-Source Retrial Queueing Systems with Two-Way Communications to the Orbit and Blocking

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

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

Abstract. A two-way communication, retrial queueing system is considered with a single server which from time to time is subject to random breakdowns. The investigated model is a M/M/1//N type of system where the number of sources is finite. After the service unit becomes idle it is able to call in customers residing in the orbit (outgoing call or secondary customers). Distribution of the service time of primary and secondary customers is exponential with rates μ_1 and μ_2, respectively. Every used random variable is assumed to be totally independent of each other in the model. Each time the server becoming in faulty state the operation of the system is blocked resulting that throughout this period customers can not enter the system. The novelty of this analysis is to study the effect of blocking in such system on the main performance measures using different distributions of failure time. Results are illustrated graphically with the help of a simulation program developed by the authors.

Keywords: Simulation · Blocking · Sensitivity analysis · Finite-source queueing system · Unreliable server · Retrial queue.

1 Introduction

Because of the increasing number of users and devices mainly due to the rapid development of technology it is not an easy task to cope with the question of designing communication systems or redesigning an existing pattern or scheme. Nowadays, every company possesses some kind of network infrastructure so it is unavoidable that the exchange of information would not take place therefore developing mathematical and simulation models and algorithms play quite an

The research was financed by the Higher Education Institutional Excellence Programme of the Ministry of Human Capacities in Hungary, within the framework of the NKFIH-1150-6/2019 thematic programme of the University of Debrecen.

V. M. Vishnevskiy et al. (Eds.): DCCN 2020, CCIS 1337, pp. 171–182, 2020.
https://doi.org/10.1007/978-3-030-66242-4_14

important role to deal with traffic growth. Applying retrial queues in such scenarios are useful and powerful tools to describe real-life problems emerging from main telecommunication systems like telephone switching systems, call centers, computer networks, and computer systems. Many researchers are dedicated to investigating this topic, some examples are mentioned which study retrial queueing systems with repeated calls like in [4,5]. The applicability of these models is utilized in many areas of science like improving the efficiency of systems for example in the case of local-area networks with random access protocols and with multiple access protocols [1,10].

The characteristics of two-way communication have a beneficial effect on most of the systems consequently its popularity is quite well-founded in recent years. This can be explainable by the fact that the operation of certain real-life systems can be matchable with models based on a two-way communication scheme. In terms of call-centers, this is especially appropriate considering that the service unit (or agent) apart from handling incoming calls may carry out other activities including selling, promoting, and advertising products. In this paper whenever the server gets to idle state after some random time it is capable of calling customers residing in the orbit. In such scenes, the utilization of the service unit (or workload of agents) is crucial and extensively examined by many papers like [3,12].

Scrutinizing the available literature on the internet relatively quite a high number of papers are found where the service facilities are presumed to be available all the time. Reliable operation is quite optimistic and an unrealistic approach because deterioration, power supply failure, or unforeseen circumstances can happen anytime modifying moderately the system characteristics. Regarding wireless communication, several components affect the transmission rate resulting in interruptions that can arise at any time throughout transmitting the packets. It is always a key question of how the property of unreliable operation alters the performance measures and the characteristics of the system. Recently published works about retrial queuing systems with a non-reliable server can be found for example in [8,9,11].

The main aim of this work is to explore the mechanism of blocking of the investigated system and to compare various distributions of failure time on main performance measures like the mean waiting time of an arbitrary customer or the total utilization of the server. The present paper is a natural continuation of [16] and we want to compare the achieved results with each other. Our self-developed simulation program is used to obtain every important performance measure using SimPack [6], which contains C/C++ libraries and executable programs for computer simulation. In this class, numerous algorithms can be found in connection with discrete-event, continuous, and combined (multi-model) simulation. Because of using other distributions apart from exponential and the fact that providing exact formulas is almost impossible we selected stochastic simulation to approximate the desired performance measures and to freely integrate any distribution in our code. The novelty of this paper is to present a sensitivity analysis of failure time on the main measures besides blocking using various distributions. Graphical illustrations are provided depicting an interesting phenomenon of sensitivity problems and comparison with the non-blocking system. This paper is the extended version of [17].

2 Model Description and Notations

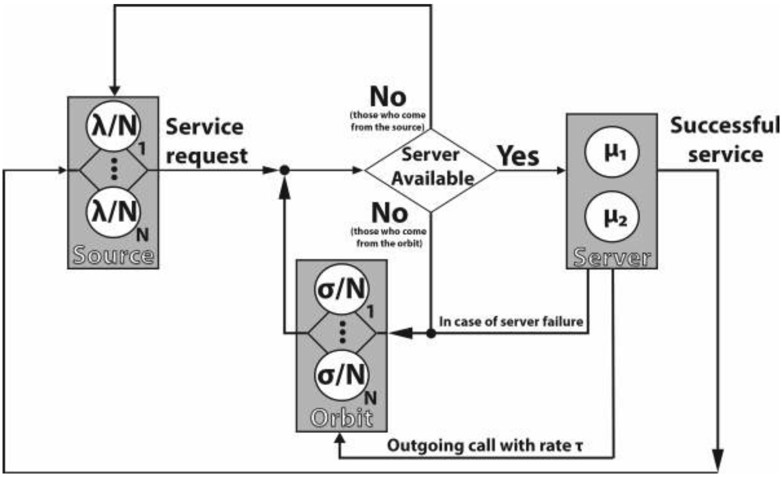

Fig. 1. System model

In Fig. 1 the considered finite-source queueing system is shown by the help of two-way communication with retrials which contains a non-reliable server. We considered a finite-source queueing system with the help of two-way communication with retrials which contains a non-reliable server. The source contains N customers and each of them produces requests (primary or ingoing customers) with rate λ/N resulting exponentially distributed inter-arrival time with parameter λ/N. Our model does not comprise queues thus in case of an idle server the service of an incoming customer starts immediately. The distribution of the service time of these customers is exponentially distributed with parameter μ_1. After being successfully served the customers return to the source. Alternatively, arriving customers from the orbit or source finding the server in a busy state are forwarded instantly to the orbit. Waiting an exponentially distributed time with parameter γ/N in this virtual waiting room customers launches another attempt to occupy the service unit. From time to time failure of the server may arise according to gamma, hypo-exponentially, hyper-exponentially, Pareto, and lognormal distribution with different parameters but with the same mean value. During this period customers can not enter the system because they are rejected in that instant, this is the so-called blocking. The recovery process begins instantaneously upon the failure of the server, which is also an exponentially distributed random variable with parameter γ_2. If the service unit breaks down during the service of a customer then that customer is transferred to the orbit immediately. Whenever the server becomes idle it may perform an outgoing call (secondary customers) towards the customers located in the orbit after an exponentially distributed random time with rate ν. The service of these customers is executed according to an exponential distribution with a rate of μ_2. Rates λ/N and σ/N are used because in [14,15] 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 Results

Our self-written simulation program includes a statistic package that was developed by Andrea Francini in 1994 [7]. Basically, this statistical analysis tool is suitable to make a quantitative estimation of the mean and variance values of the desired variables using the method of batch means. In each batch, there are n observations and the useful run is divided into numerous batches. The batches should be long enough and approximately independent in order that the estimation would work correctly. This method belongs to one of the most popular confidence interval techniques for a steady-state mean of a process. In more detailed information about this method is included in the following works [2,13]. 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 Squared Coefficient of Variation is Greater Than One

To realize the sensitivity analysis four different distributions of failure time are selected to compare the performance measures with each other. The parameters are chosen in such a way that the mean value and variance would be equal, so we applied a fitting process that is necessary to be done. [18] contains a detailed description of the whole process characterizing every used distribution. We differentiated two main scenarios from each other. In the first one, the squared coefficient of variation is greater than one so I utilized hyper-exponential, gamma, Pareto, and lognormal distributions. Table 2 quantifies all the used input parameters of the various distributions of failure time while Table 1 shows the values of other parameters.

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

The steady-state distributions are represented on Fig. 2 when λ/N is $= 0.01$ comparing the effect of all four applied distributions of failure time. It shows the probability that exactly i customers are located in the system. Averagely the same number of customers resides in the system, slight differences can be perceivable especially in the case of Pareto. Taking a closer look at the graphs all the curves correspond to normal distribution despite the characteristics of the various distribution.

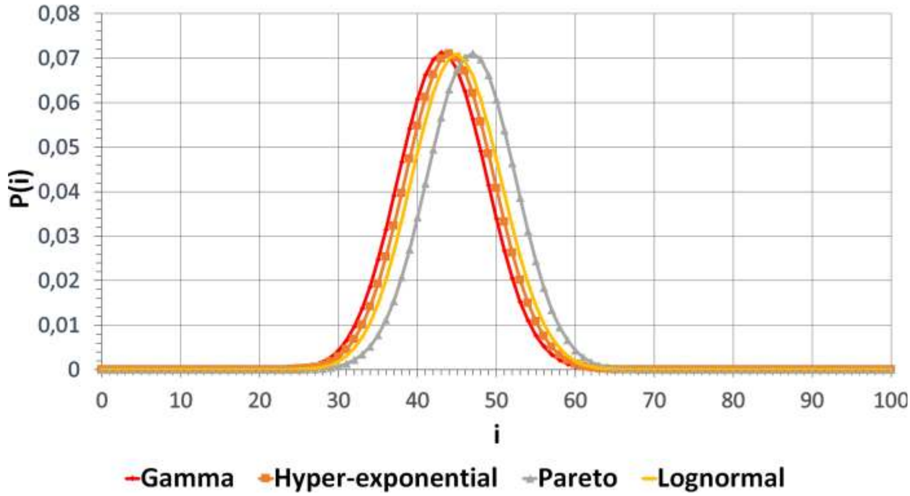

Fig. 2. Comparison of steady-state distributions when $\lambda/N = 0.01$

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			

In Fig. 3 the mean arbitrary response time is demonstrated as the request generation increases. Results clearly illustrate the effect of various distributions which is quite significant even though the first two moments are equal. The highest values are experienced at Pareto distribution while the lowest values at gamma distribution. The maximum property characteristic of a finite-source retrial queueing system arises which under suitable parameter setting occurs in spite of increasing arrival intensity.

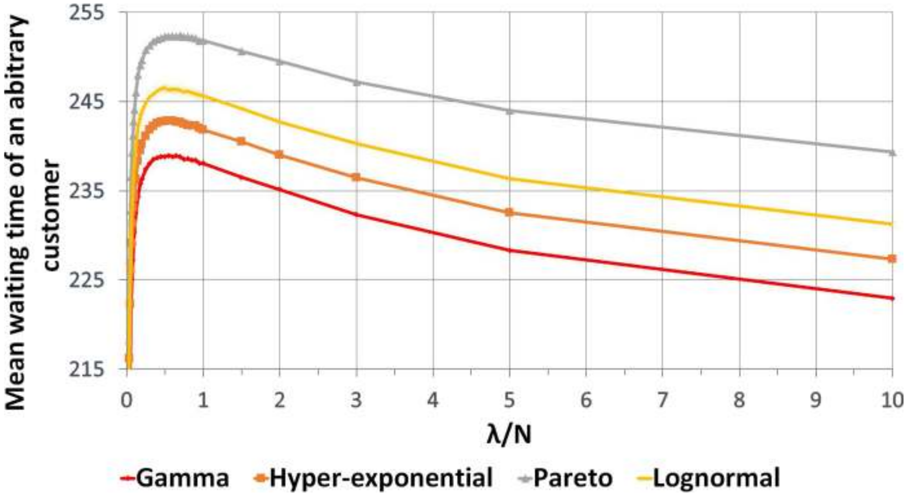

Fig. 3. Mean waiting time vs. arrival intensity

Figure 4 shows how the total utilization of the server escalates applying intensifying arrival intensity. Under total utilization, we mean every single service including the service of primary, secondary customers, and the interrupted ones, too. By examining closely the figure the received values are almost identical but the tendency is counteractive as we have seen in Fig. 3. As more and more customers enter the system the total utilization of the service unit increases.

Figure 5 emphasizes the effect of blocking on the mean waiting time versus arrival intensity. It is observable that in case of blocking the customers spend less time on average because during server failure the incoming customers go back to the source instead of waiting in the orbit. Besides the higher failure rate, the difference is more significant as well. At Fig. 5 the distribution of service time of the incoming customer is gamma, but the same tendency can be found in the case of the other distributions, too.

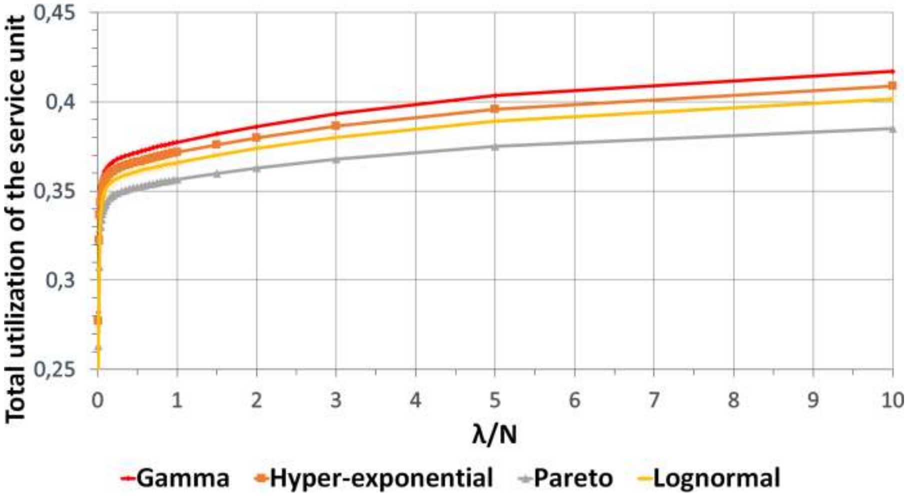

Fig. 4. Total utilization of the server vs. arrival intensity

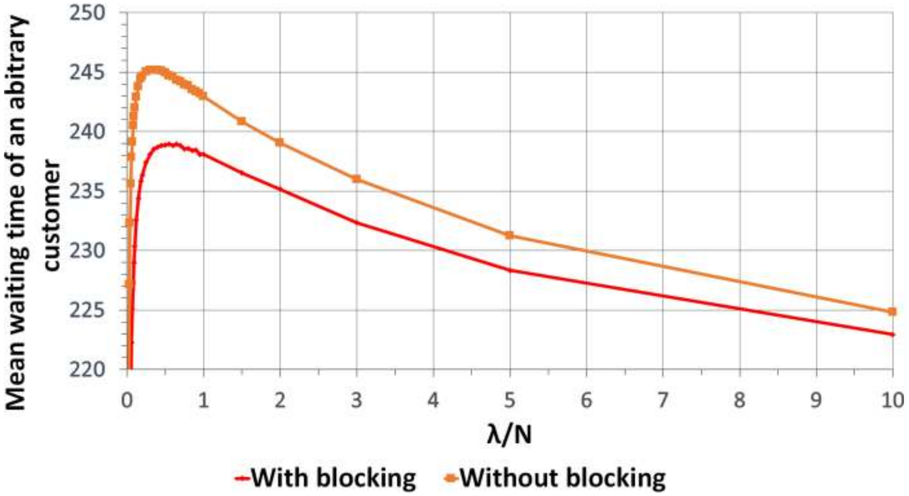

Fig. 5. The effect of blocking on the mean waiting time

3.2 Squared Coefficient of Variation Is Less Than One

Analyzing the results of the previous section we were curiously interested in how the modified parameters of the failure time alter the performance measures. In that case the parameters were chosen so that the squared coefficient of variation should be less than one. Instead of hyper-exponential we utilize hypo-exponential distribution because in the case of hypo-exponential distribution the squared coefficient of variation is always less than one. The same performance measures will be presented graphically as above but with using the new parameters of failure time which is shown in Table 3. The other parameters remain unchanged see Table 1.

Table 3. 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			

Figure 6 presents the comparison of the steady-state distribution using different distributions of failure time. The curves overlap almost each other meaning that on average regardless the utilized distribution of failure time the same number of customers are located. As compared with Fig. 2 even the averages are identical, which is around 65–66. Despite the distinct distributions every graph is tend to be normally distributed as in the previous section.

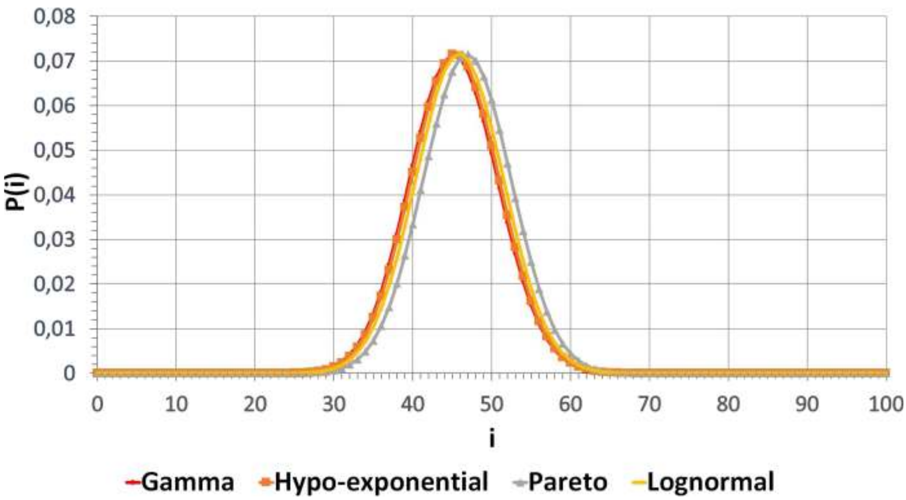

Fig. 6. Comparison of steady-state distributions

In Fig. 7 it can be seen how the mean waiting of a customer develops along with increasing arrival intensity. By examining closer the received graphs it is observable that they are much closer to each other compared to Fig. 3 although they can be differentiated from each other moderately. As in the case of Fig. 3 the highest values are experienced at Pareto distribution. When the squared coefficient of variation is less than one in terms of every distribution the values of mean waiting is higher as in the previous section.

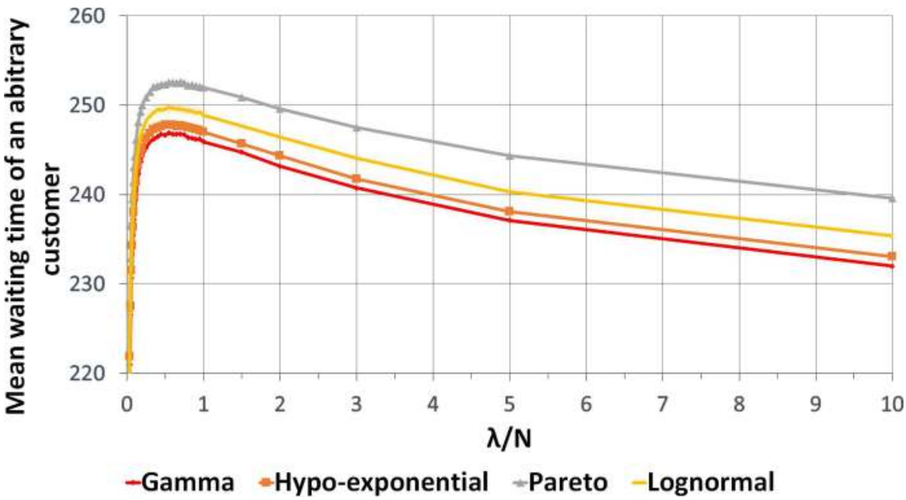

Fig. 7. Mean waiting time vs. arrival intensity using various distributions of operation time

Figure 8 demonstrates the utilization of the service unit in function of arrival intensity. No huge distinction can be discovered in terms of total utilization of the service unit among the various distributions observing the results. Basically we got back almost the same values except Pareto where the utilization is a little bit smaller. The same tendency can be noticeable comparing to Fig. 4 so at every distribution possess approximately same value of utilization. It is not surprising to see that the utilization increases with the increment of the arrival intensity.

Figure 9 represents the effect of blocking on the mean waiting time in the function of arrival intensity. The analyzed system with blocking results in lower average waiting time of an arbitrary customer due to that a customer can not enter the system when the service unit is faulty. Regardless of the used input parameters of the operation time the difference is similar to what we have observed in Fig. 5. The distribution of the operation time follows gamma distribution but it has to be noted that the same tendency can be noticed in the case of the other distributions, too.

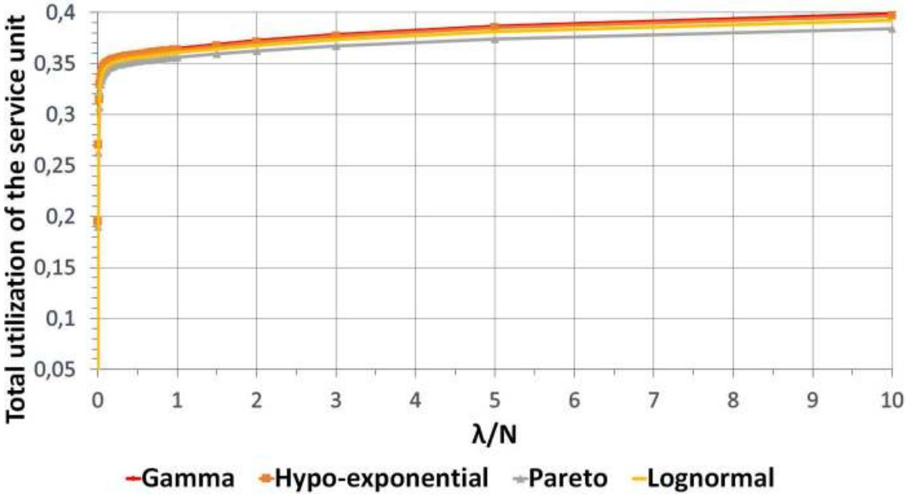

Fig. 8. Total utilization of the server vs. arrival intensity using various distributions of operation time

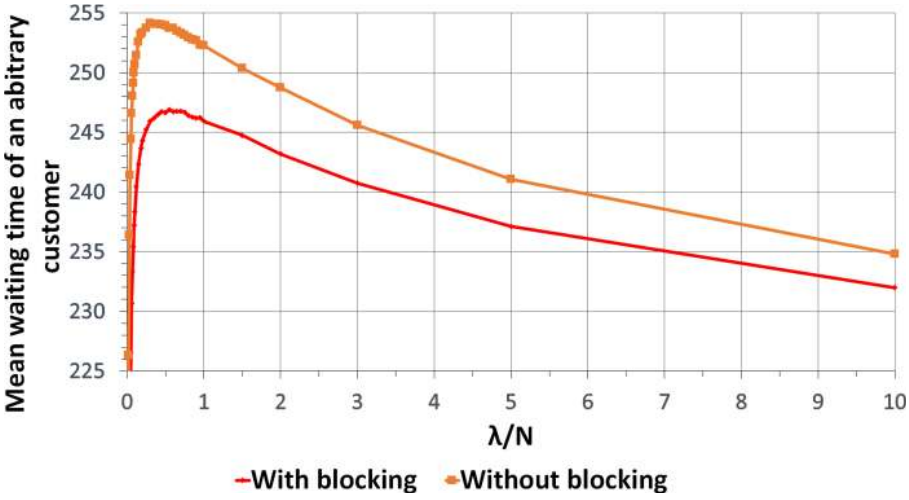

Fig. 9. The effect of blocking on the mean waiting time

4 Conclusion

A finite-source retrial queueing system with the help of two-way communication is introduced with applying blocking and an unreliable server which can make outgoing calls towards the customers of the orbit. The effect of the used distributions and blocking is illustrated by several figures on the mean arbitrary waiting time and the total utilization of the server. With the aid of stochastic simulation, the obtained results clearly revealed that in case the squared coefficient of variation is greater than one the disparity among the values of displayed performance measures is significant having the same mean and variance. In the future we would like to complete this system with other features like experimenting with more distributions, introducing some kind of impatience of the customers, or including more capacity of service.

References

1. Artalejo, J., Corral, A.G.: Retrial Queueing Systems: A Computational Approach. Springer, Cham (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. Falin, G., Artalejo, J.: A finite source retrial queue. Eur. J. Oper. Res. **108**, 409–424 (1998)
5. Fiems, D., Phung-Duc, T.: Light-traffic analysis of random access systems without collisions. Ann. Oper. Res. **277**, 311–3277 (2019). https://doi.org/10.1007/s10479-017-2636-7
6. Fishwick, P.A.: SimPack: Getting started with simulation programming in C and C++. In: In 1992 Winter Simulation Conference, pp. 154–162 (1992)
7. 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 1996 - 4th International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, pp. 116–122 (February 1996)
8. Gharbi, N., Dutheillet, C.: An algorithmic approach for analysis of finite-source retrial systems with unreliable servers. Comput. Math. Appl. **62**(6), 2535–2546 (2011)
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)
11. Krishnamoorthy, A., Pramod, P.K., Chakravarthy, S.R.: Queues with interruptions: a survey. TOP **22**(1), 290–320 (2012)
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. Law, A.M., Kelton, W.D.: Simulation Modeling and Analysis. McGraw-Hill Education, New York (1991)
14. 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

15. 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)
16. 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.) ITMM 2019. CCIS, vol. 1109, pp. 270–284. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-33388-1_22
17. Sztrik, J., Tóth, Á., Pintér, Á., Bács, Z.: Reliability analysis of finite-source retrial queueing systems with two-way communications to the orbit and blocking using simulation. In: Proceedings of International Conference on Distributed Computer and Communication Networks, DCCN 2020 (2020)
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)