Simulation of Finite-Source Retrial Queues with Two-Way Communications to the Orbit

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. In this paper we investigate a single-server two-way communication system by the help of retrial queuing systems with finite source. From the finite source incoming primary calls enter into the system according to an exponential distribution. If the server is idle then the service of incoming customer starts immediately. Alternatively, if an incoming customer discovers the server in busy state it is directed towards the orbit, where after some exponentially distributed time retries to reach the server again. As soon as the server becomes idle it can generate an outgoing call to the customers in the orbit after an exponentially distributed time. In case of two-way communication after the service of an outgoing call it returns to the source. In this work we concentrate on emphasizing a phenomena of outgoing call on the mean waiting time of incoming customers. The novelty of this paper is to carry out a sensitivity analysis comparing various distributions of service time of primary customers on the performance measures like utilization of the server or mean waiting time. By the use of simulation several graphical results and comparison of the applied systems are illustrated.

Keywords: Retrial queues · Two-way communication · Sensitivity analysis · Finite-source queuing systems · Simulation

1 Introduction

Finite-source retrial queues are effective and commonly used systems for modeling real life problems arising in main telecommunication systems like cellular mobile communication networks, computer networks, local-area networks with random access protocols, call centers and CSMA-based wire-less networks. With the decrement of the rate of generation of new calls the number of customers in the system increases in case of many practical situation. This can be performed

with the help of finite-source or quasi-random input models. Their importance can be viewed by the reader in the following works, for example [4, 10, 11, 14, 15].

Systems with retrial feature are identified by a specific feature of arriving customers when the server is occupied. These customers stay in the system and spend their time in a virtual waiting room called orbit. Customers in the orbit attempt to be served after a random time. Because the number of calls are finite, the assumption of working with finite-source queuing systems follows real circumstances. In this paper we examine two-way communication retrial queuing system which is quite popular topic in the recent years. This can be explained by the fact that using two-way communication scheme is very helpful in many application fields to model real life problems. Especially in case of call centers where service unit can perform certain other work in idle state like selling, advertising and promoting products including serving incoming calls. In such systems utilization of the service unit is always pivotal, see for example in [1, 2, 6, 9, 13, 16, 20, 22].

Once the server becomes idle it calls for customers inside and outside of the system which is called an outgoing call. This is a typical feature of two-way communication system. In our investigated model the idle service unit can generate a call only from the orbit which arrives after a random time. It will only be served if no customers from the finite source or from the orbit come. Otherwise this outgoing call will be canceled. Papers dealing with two-way communication systems by the help of retrial queues, where the source is infinite, are found in [3, 5, 7, 8, 17–19, 21].

Our aim is to study the operation of the system where the service unit is reliable and can perform outgoing call from the orbit. The novelty of this paper is to compare this system with the common finite source retrial system using various distribution of service on performance measures like mean waiting time of an incoming call or utilization of the server. We are mainly interested in how the different distributions modify the characteristics of the system. To achieve this goal a simulation program has been developed using the base of SimPack [12] which contains a number of C/C++ libraries and executable programs. One of the main reasons for its usage is that the user has the freedom what performance measure are calculated and how the model is built up. SimPack toolkit also provides a set of utilities that demonstrate how to build a working simulation from a model description.

2 System Model

We consider a retrial queuing system of type $M/G/1//N$ with a reliable server which is capable to produce outgoing calls to the customers residing in the orbit. N customers are located in the source, where all of them can generate incoming, primary calls towards the server. The distribution of the inter-request times is exponential with rate λ/N. In default of waiting queue an incoming customer either from the source or orbit finds the server in an idle state then its service begins instantly. The service times of incoming customers are assumed to

be gamma, hypo-exponentially, hyper-exponentially, Pareto and lognormal distributed with different parameters but with the same mean value. Customers return to the source after their service is terminated. If the server is busy, meaning that a request is under service, an incoming customer remains in the system and enters into the orbit. Customers located in the orbit are able to attempt to access the server again after an exponentially distributed time with parameter σ/N. In the other hand, when the server becomes idle it can make outgoing call towards the customers in the orbit. It is performed after an exponentially distributed time with parameter ν. The service time of these outgoing customers follows gamma distribution with parameters α_2 and β_2. In a consecutive paper we aim to investigate the same system by the help of asymptotic methods when N tends to infinity and that is the reason we use λ/N and σ/N parameters. All the random variables involved in the model construction are assumed to be totally independent of each other.

3   Applied Distributions and Its Parameters

In this Section the reader gets an insight of the parameters of the applied distributions and the process how to select them in order to execute a valid comparison. To do so our program is integrated with random number generators according to gamma, hyper-exponential, hypo-exponential, lognormal and Pareto distribution. These random number generators need input parameters which are different in every distribution, thus parameter selection is crucial. For valid comparison we use the same mean and variance in case of every distribution hence we take over every distribution and how the fitting process is accomplished.

3.1   Gamma Distribution

Gamma distribution is a general type of statistical distribution and a random variable X has a gamma distribution if its density function is the following:

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{\beta(\beta x)^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)} & \text{if } x \geq 0 \end{cases}$$

where $\beta > 0$ and $\alpha > 0$.

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\,e^{-t}\,\mathrm{d}t$$

This is the so-called complete gamma function, which has two parameters: α is called the shape parameter and β is called the scale parameter. These two parameters are also the input parameters of the random number generator.

The coefficient $C_X^2 = \frac{Var(X)}{(EX)^2}$ is defined as the squared coefficient of variation of random variable X.

The mean value, variation and the squared coefficient of variation can be calculated:

$$\overline{X} = \frac{\alpha}{\beta}, \qquad Var(X) = \frac{\alpha}{\beta^2}, \qquad C_X^2 = \frac{1}{\alpha}$$

For a predetermined mean value and variance to obtain parameters α and β the next calculation has to be done:

$$\alpha = \frac{1}{C_X^2}, \qquad \beta = \frac{\alpha}{\overline{X}}$$

3.2   Pareto Distribution

A random variable X has a Pareto distribution if its density function is the following:

$$f(x) = \begin{cases} 0 & \text{if } x < k \\ \alpha k^\alpha x^{-\alpha-1} & \text{if } x \geq k \end{cases}$$

Hence the distribution function is:

$$F(x) = \begin{cases} 0 & \text{if } x < k \\ 1 - \left(\frac{k}{x}\right)^\alpha & \text{if } x \geq k \end{cases}$$

where $\alpha, k > 0$.

It has two parameters: α is called the shape parameter and k is called the location parameter. These two parameters are the input parameters of the random number generator.

The mean value, variation and the squared coefficient of variation can be calculated as follows:

$$\overline{X} = \begin{cases} \frac{k\alpha}{\alpha-1} & \text{if } \alpha > 1 \\ \infty & \text{if } \alpha \leq 1 \end{cases}$$

$$Var(X) = \frac{k^2\alpha}{\alpha-2} - \left(\frac{k\alpha}{\alpha-1}\right)^2, \qquad C_X^2 = \frac{(\alpha-1)^2}{\alpha(\alpha-2)} - 1, \qquad \alpha > 2.$$

For a predetermined mean value and variance to obtain parameters α and k the following interrelation is used:

$$\alpha = 1 + \frac{\sqrt{1 + C_X^2}}{\sqrt{C_X^2}}, \qquad k = \frac{\alpha-1}{\alpha} \times \overline{X}$$

3.3 Lognormal Distribution

Let $Y \in N(m, \sigma)$ a random variable with normal distribution, lognormal is a continuous distribution in which the logarithm of a variable having a normal distribution, namely $X = e^Y$ has lognormal distribution with parameters (m, σ). Its distribution and density function are the following:

$$F_x(x) = \Phi\Big(\frac{\ln(x) - m}{\sigma}\Big), \qquad x > 0.$$

$$f_x(x) = \frac{1}{\sigma x}\varphi\Big(\frac{\ln(x) - m}{\sigma}\Big), \qquad x > 0.$$

The mean value, variance and the squared coefficient of variation can be calculated:

$$\overline{X} = e^{m + \frac{\sigma^2}{2}}, \qquad Var(X) = e^{2m + \sigma^2}(e^{\sigma^2} - 1), \qquad C_X^2 = e^{\sigma^2} - 1.$$

To obtain the two parameters of the lognormal distribution the following interrelation is applied:

$$\sigma = \sqrt{\ln(1 + C_X^2)}, \qquad m = \ln(\overline{X}) - \frac{\sigma^2}{2}$$

3.4 Hypo-exponential Distribution

Continuous statistical distribution, let $X_i \in Exp(\mu_i)(i = 1, ..., n)$ be independent exponentially distributed random variables. Then $Y_n = X_1 + ... + X_n$ has n-phase hypo-exponential distribution. Its density function is given by

$$f_{Y_n}(x) = \begin{cases} 0 & \text{if } x < 0 \\ (-1)^{n-1}\Big[\prod_{i=1}^{n}\mu_i\Big]\sum_{j=1}^{n}\frac{e^{-\mu_j x}}{\prod_{k=1, k\neq j}^{n}(\mu_j - \mu_k)} & \text{if } x \geq 0. \end{cases}$$

The mean value, variance and the squared coefficient of variation can be calculated:

$$\overline{Y_n} = \sum_{i=1}^{n}\frac{1}{\mu_i}, \qquad Var(Y_n) = \sum_{i=1}^{n}\frac{1}{\mu_i^2}, \qquad C_{Y_n}^2 = \frac{\sum_{i=1}^{n}\Big(\frac{1}{\mu_i}\Big)^2}{\Big(\sum_{i=1}^{n}\frac{1}{\mu_i}\Big)^2}.$$

In our simulation program we used the 2-phase hypo-exponential distribution where the parameters are the parameters of the two independent exponential distribution (μ_1, μ_2). For a predetermined mean value and variance to obtain parameters μ_1 and μ_2 the next equation system has to be solved:

$$\overline{X} = \frac{1}{\mu_1} + \frac{1}{\mu_2}, \qquad Var(X) = \frac{1}{\mu_1^2} + \frac{1}{\mu_2^2}$$

3.5   Hyper-exponential Distribution

Suppose $X_1, X_2, \cdots, X_n$ are independent exponential random variables, where the rate parameter of X_i is λ_i. The random variable X can be one of the n independent exponential random variables $X_1, X_2, \cdots, X_n$ such that X is X_i with probability p_i with $p_1 + \cdots + p_n = 1$. Such a random variable X is said to follow a hyper-exponential distribution. Its density function is given by

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ \sum_{i=1}^{n} p_i \lambda_i e^{-\lambda_i x} & \text{if } x \geq 0. \end{cases}$$

Its distribution function is

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - \sum_{i=1}^{n} p_i e^{-\lambda_i x} & \text{if } x \geq 0. \end{cases}$$

In the case when for a random variable $X, C_X^2 > 1$ then the the following two-moment fit is suggested

$$f_Y(t) = p\lambda_1 e^{-\lambda_1 t} + (1-p)\lambda_2 e^{-\lambda_2 t}.$$

Y is a 2-phase hyper-exponentially distributed random variable. The most commonly used procedure is the balanced mean method, that is

$$\frac{p}{\lambda_1} = \frac{1-p}{\lambda_2}.$$

To obtain the three parameters of the hyper-exponential distribution the following calculation is used:

$$p = \frac{1}{2}\left(\sqrt{\frac{C_X^2 - 1}{C_X^2 + 1}}\right), \qquad \lambda_1 = \frac{2p}{\overline{X}}, \qquad \lambda_2 = \frac{2(1-p)}{\overline{X}}.$$

4   Simulation Results

4.1   Squared Coefficient of Variation is Greater than One

The values of the input parameters are shown in Table 1. In this section these results are in connection with the effect of different service time distributions of incoming customers where the mean and variance are equal, respectively. We use hyper-exponential distribution if the squared coefficient of variation is greater than one, Table 2 shows the exact values of parameters of service time of incoming customers. Besides hyper-exponential, gamma, lognormal and Pareto distributions are also used for comparisons.

Figure 1 shows the mean waiting time in function of arrival intensity of incoming customers. For these values of parameters regardless of the applied distribution a maximum value of the mean waiting time can be seen. This maximum

Table 1. Numerical values of model parameters

N	σ/N	ν	α_2	β_2
100	0.01	0.02	1	1.1

Table 2. Parameters of service time of incoming customers

Distribution	Gamma	Hyper-exponential	Pareto	Lognormal
Parameters	$\alpha = 0.04$	$p = 0.48$	$\alpha = 2.02$	$m = -1.629$
	$\beta = 0.04$	$\lambda_1 = 0.961$	$k = 0.505$	$\sigma = 1.805$
		$\lambda_2 = 1.04$		
Mean	1			
Variance	25			
Squared coefficient of variation	25			
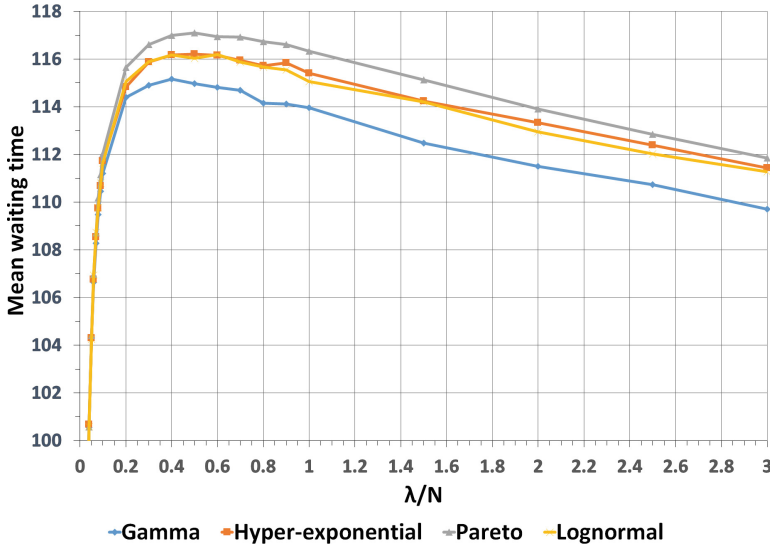
Fig. 1. Mean waiting time vs. arrival intensity using various distributions

feature occurs for finite-source retrial queues, see for example [4,9,10,16]. Differences can be observed among the values of mean waiting time especially in the case of using gamma and Pareto distribution, despite the fact that the mean and variance are the same. On this figure the effect of different distributions is clearly observable.

Figures 2 and 3 illustrates how the utilization of the server grows with the increment of the arrival intensity of incoming customers. The highest values can be found at gamma distribution but the differences of the applied distributions
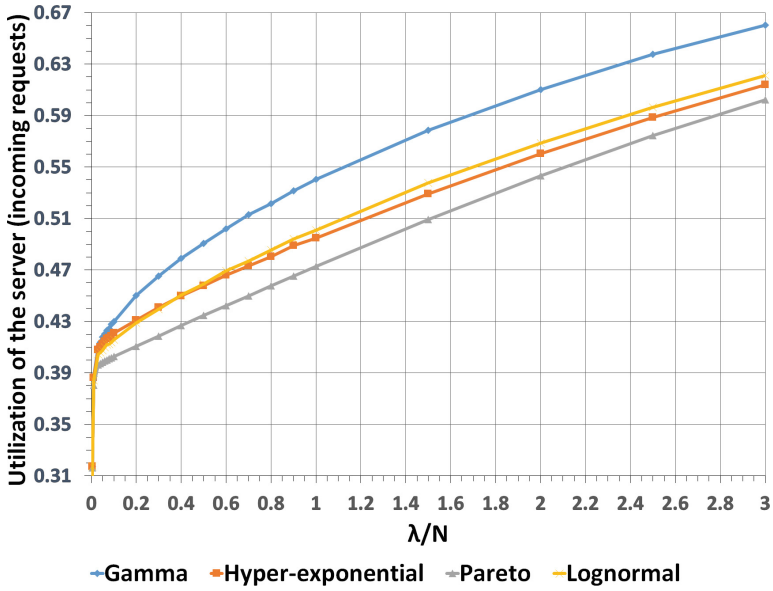
Fig. 2. Utilization of server vs. arrival intensity using various distributions
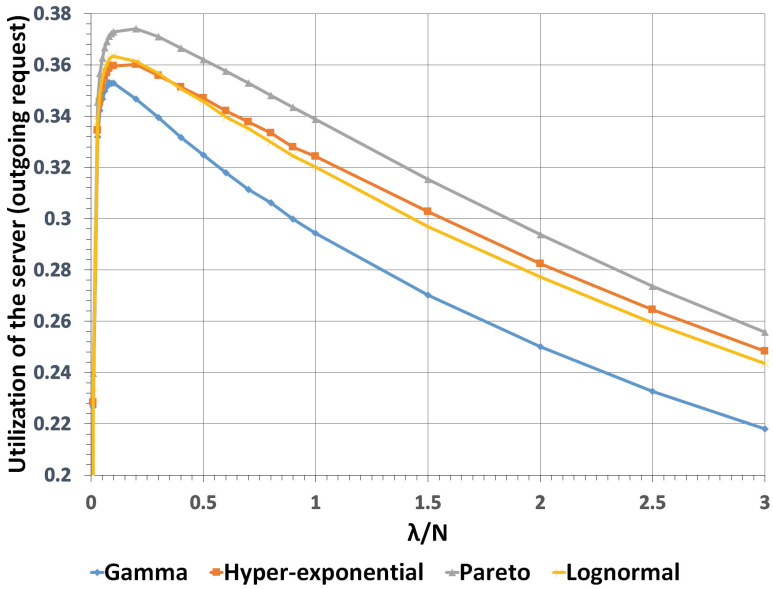

Fig. 3. Utilization of server vs. arrival intensity using various distributions

are as commensurable as in case of Fig. 1. As the arrival intensity increases the
probability of performing outgoing call become less so outgoing requests spend
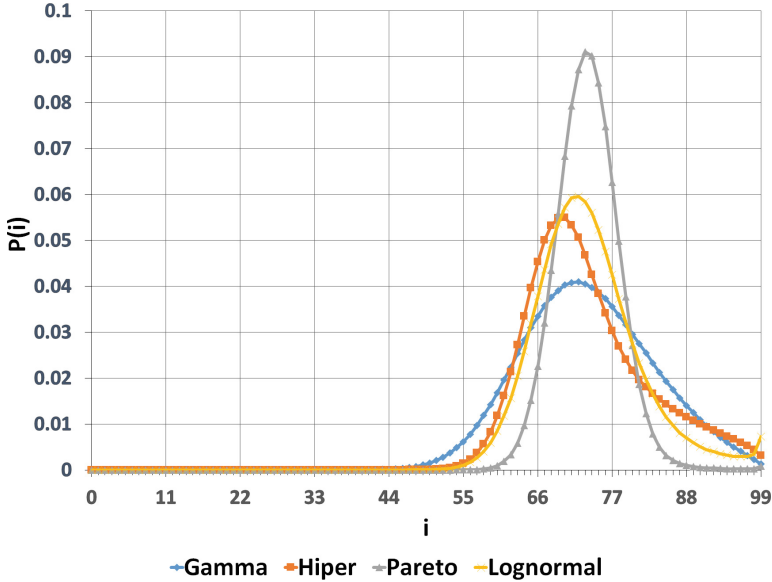less time at the service unit.

Fig. 4. Comparison of steady-state distributions

On Fig. 4 the comparison of steady-state distribution can be seen when the
distribution of service time of the incoming customers is different. It represents
the probability of how many customers residing in the orbit. Exploring the curves
in more detail they correspond to normal distribution. The same parameter
setting is used what Table 1 demonstrates where λ/N is 0.03.

To emphasize the importance of outgoing calls we compare our investigated
model to the model without outgoing calls. This model is named as the classical
retrial queuing system. On Fig. 5 comparison of the mean waiting time can be
seen and due to the phenomena of outgoing call customers spend less time in
the system, which is obvious looking at the curves. However, in our investigated
model the utilization of the service unit (Fig. 6) is much higher compared to
the classical retrial queuing system therefore it spends less time in idle state.
In this way the efficiency of the server grows such that the mean waiting time
decreases substantively. The distribution of service time of the incoming cus-
tomer is gamma at Figs. 5 and 6, but the ratio of difference is also true for the
other distributions, too. On this figure under total utilization of server we mean
the service of both the incoming and outgoing requests at the curve of with
outgoing call.
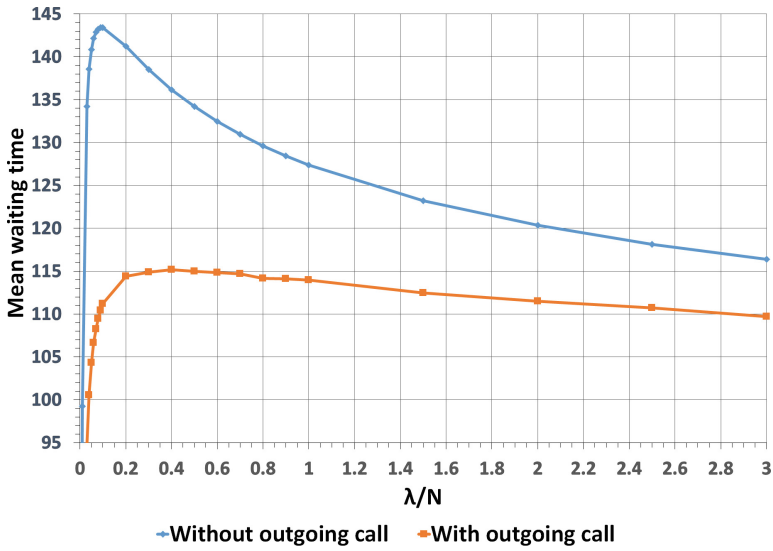
**Fig. 5.** Comparison of our investigated model and the classical retrial queuing model on the mean waiting time
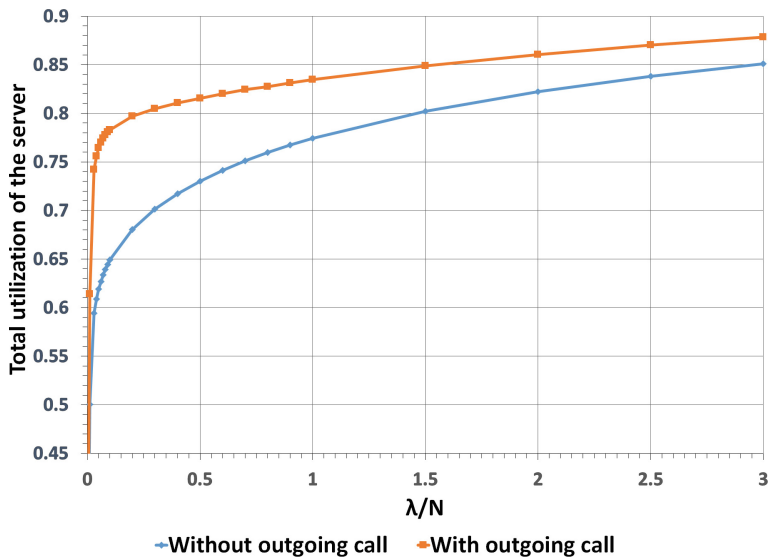
Fig. 6. Comparison of our investigated model and the classical retrial queuing model on the utilization of server

4.2   Squared Coefficient of Variation is Less than One

The same input parameters are used as in the previous section, see Table 1. The results are also in connection with the effect of different service time distributions of incoming customers where the mean and variance are equal. Instead of hyper-exponential distribution hypo-exponential distribution is used if the squared coefficient of variation is less than one. Table 3 illustrates the values of parameters of service time of incoming customers. In addition to hypo-exponential, we apply gamma, lognormal and Pareto distributions to perform sensitivity analysis.

Table 3. Parameters of service time of incoming customers

Distribution	Gamma	Hypo-exponential	Pareto	Lognormal
Parameters	$\alpha = 1.5504$	$\mu_1 = 1.3$	$\alpha = 2.597$	$m = -0.249$
	$\beta = 1.5504$	$\mu_2 = 4.333$	$k = 0.615$	$\sigma = 0.705$
Mean	1			
Variance	0.6449704142			
Squared coefficient of variation	0.6449704142			
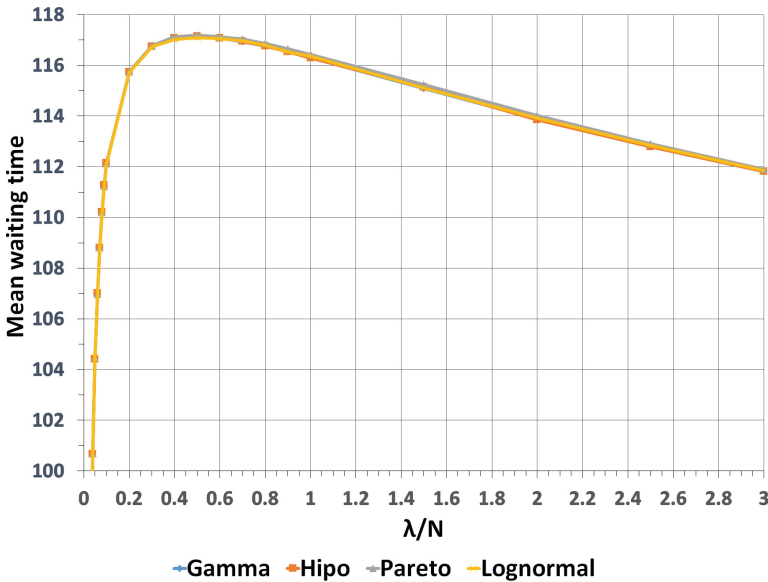
Fig. 7. Mean waiting time vs. arrival intensity using various distributions

Figure 7 demonstrates the mean waiting time in the function of arrival intensity of incoming calls. Taking closer look at the curves it can be stated that

the values of mean waiting time are almost identical regardless of the applied distribution. With this parameter setting the interesting maximum value of the mean waiting time appears as in the previous section.
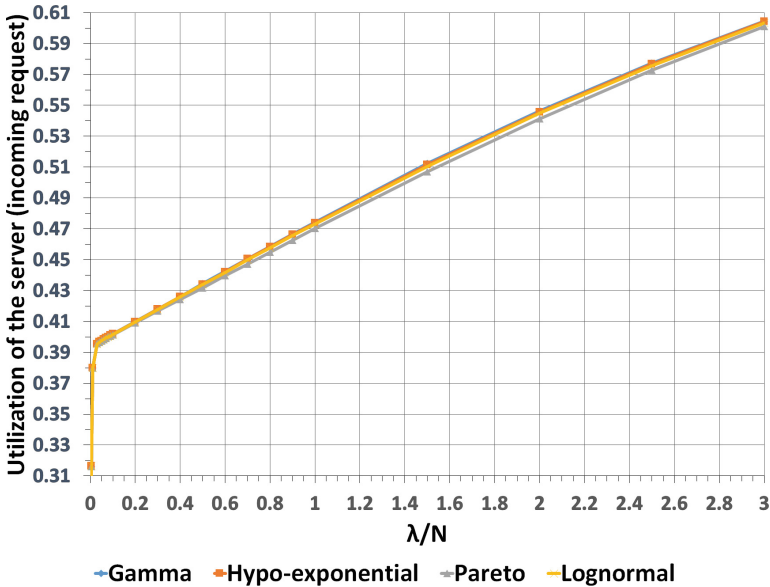

Fig. 8. Utilization of server vs. arrival intensity using various distributions

Figures 8 and 9 illustrates how the utilization of the server increases with the increment of the arrival intensity of incoming customers. As in case of mean waiting time here using different distributions result the same utilization. It seems that when the squared coefficient of variation is less than one using different distributions have no effect on the performance measures and the obtained results are nearly identical.

Similarly to the previous section we compare the results between the classical retrial queuing system and our investigated model. On Figs. 8 and 9 the same tendency can be observed like when the the squared coefficient of variation is greater than one, namely values of mean waiting time is lesser when the server can make outgoing calls. But this also affects the utilization of the service unit because with the help of outgoing calls server spend less time without satisfying the needs of the customers. As in the previous section the service time of the incoming customer follows gamma distribution at Fig. 10 and 11, but the ratio of difference is also true for the other distributions, too.

From Figs. 5, 6, 10 and 11 it can be said that the utilization of service unit escalates when outgoing calls are performed, but it also results lesser mean waiting time of incoming customers. With a proper parameter setting in the case of outgoing calls the utilization of the server is much higher in a way that customers
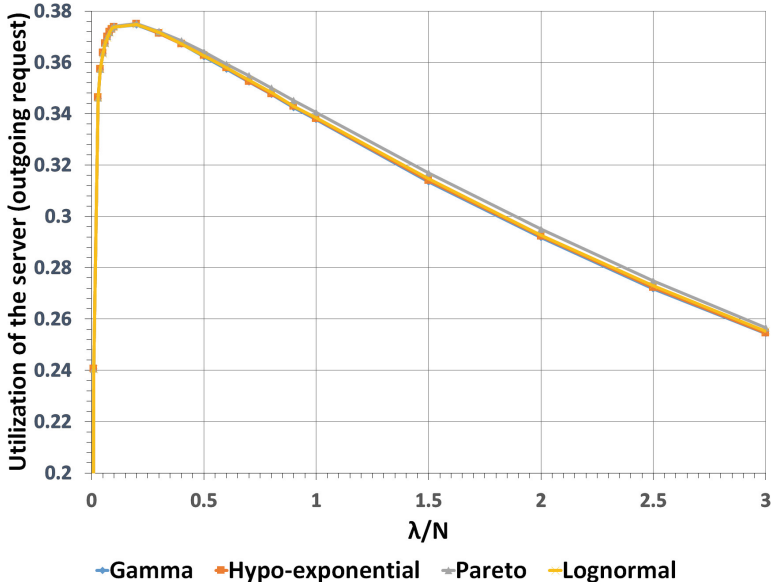
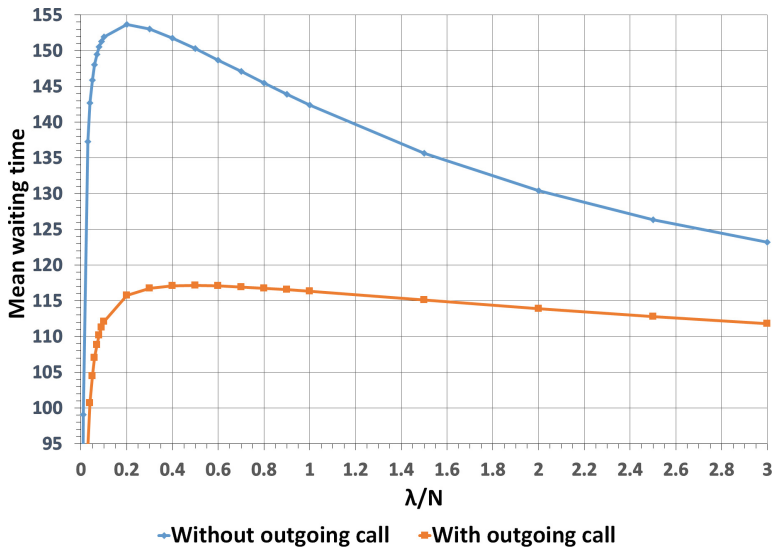**Fig. 9.** Utilization of server vs. arrival intensity using various distributions
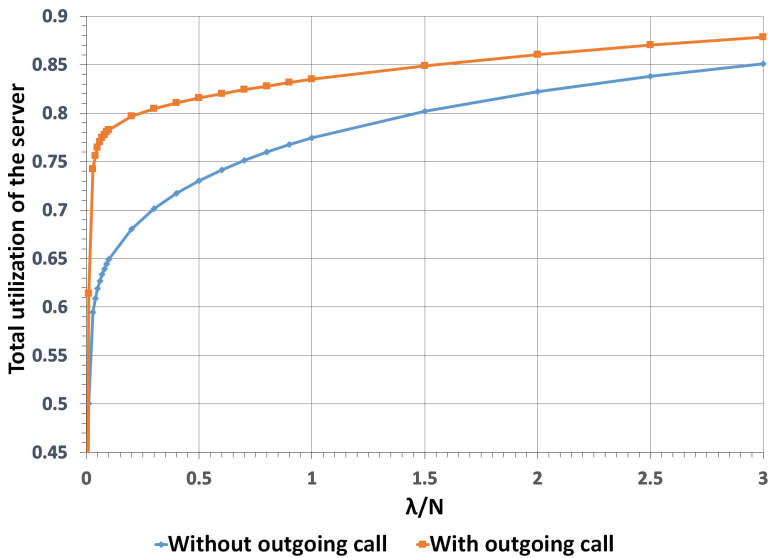

Fig. 10. Comparison of our investigated model and the classical retrial queuing model on the mean waiting time

Fig. 11. Comparison of our investigated model and the classical retrial queuing model on the utilization of server

spend less time in the orbit. In the case of with outgoing calls total utilization of the server includes both incoming and outgoing requests occupying the service unit.

5 Conclusion

A finite-source retrial queueing system is introduced where the server can produce outgoing calls towards the customers of the orbit. Several figures present the effect of the applied distributions on the mean waiting time and on the utilization of the server. Using stochastic simulation method results clearly indicate that when the squared coefficient of variation is greater than one then the contrast of the values of the performance measures is quite high having the same mean and variance.

References

1. Aguir, S., Karaesmen, F., Akşin, O.Z., Chauvet, F.: The impact of retrials on call center performance. OR Spectr. **26**(3), 353–376 (2004)
2. Aksin, Z., Armony, M., Mehrotra, V.: The modern call center: a multi-disciplinary perspective on operations management research. Prod. Oper. Manag. **16**(6), 665–688 (2007)
3. Artalejo, J.R., Phung-Duc, T.: Markovian retrial queues with two way communication. J. Ind. Manag. Optim. **8**(4), 781–806 (2012)
4. Artalejo, J., Corral, A.G.: Retrial Queueing Systems: A Computational Approach. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-78725-9

5. Artalejo, J., Phung-Duc, T.: Single server retrial queues with two way communication. Appl. Math. Modell. **37**(4), 1811–1822 (2013)
6. Brown, L., et al.: Statistical analysis of a telephone call center: a queueing-science perspective. J. Am. Stat. Assoc. **100**(469), 36–50 (2005)
7. Dimitriou, I.: A retrial queue to model a two-relay cooperative wireless system with simultaneous packet reception. In: Wittevrongel, S., Phung-Duc, T. (eds.) ASMTA 2016. LNCS, vol. 9845, pp. 123–139. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-43904-4_9
8. Dragieva, V., Phung-Duc, T.: Two-way communication M/M/1 retrial queue with server-orbit interaction. In: Proceedings of the 11th International Conference on Queueing Theory and Network Applications, p. 11. ACM (2016)
9. 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
10. Falin, G., Artalejo, J.: A finite source retrial queue. Eur. J. Oper. Res. **108**, 409–424 (1998)
11. Fiems, D., Phung-Duc, T.: Light-traffic analysis of random access systems without collisions. Ann. Oper. Res. **277**(2), 1–17 (2017). https://doi.org/10.1007/s10479-017-2636-7
12. Fishwick, P.A.: Simpack: getting started with simulation programming in C and C++. In: WSC 1992 Proceedings of the 24th Conference on Winter simulation, pp. 154–162. ACM (1992)
13. Gans, N., Koole, G., Mandelbaum, A.: Telephone call centers: tutorial, review, and research prospects. Manuf. Serv. Oper. Manag. **5**(2), 79–141 (2003)
14. Gómez-Corral, A., Phung-Duc, T.: Retrial queues and related models. Ann. Oper. Res. **247**(1), 1–2 (2016). https://doi.org/10.1007/s10479-016-2305-2
15. Kim, J., Kim, B.: A survey of retrial queueing systems. Ann. Oper. Res. **247**(1), 3–36 (2016). https://doi.org/10.1007/s10479-015-2038-7
16. 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
17. Nazarov, A., Phung-Duc, T., Paul, S.: Heavy outgoing call asymptotics for MMPP/M/1/1 retrial queue with two-way communication. In: Dudin, A., Nazarov, A., Kirpichnikov, A. (eds.) ITMM 2017. CCIS, vol. 800, pp. 28–41. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68069-9_3
18. Nazarov, A.A., Paul, S., Gudkova, I., et al.: Asymptotic analysis of Markovian retrial queue with two-way communication under low rate of retrials condition. In: Proceedings 31st European Conference on Modelling and Simulation (2017)
19. Phung-Duc, T., Rogiest, W.: Two way communication retrial queues with balanced call blending. In: Al-Begain, K., Fiems, D., Vincent, J.-M. (eds.) ASMTA 2012. LNCS, vol. 7314, pp. 16–31. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-30782-9_2
20. Pustova, S.: Investigation of call centers as retrial queuing systems. Cybern. Syst. Anal. **46**(3), 494–499 (2010)
21. Sakurai, H., Phung-Duc, T.: Two-way communication retrial queues with multiple types of outgoing calls. Top **23**(2), 466–492 (2015). https://doi.org/10.1007/s11750-014-0349-5
22. Wolf, T.: System and method for improving call center communications, uS Patent App. 15/604,068. Accessed 30 Nov 2017