Annales Mathematicae et Informaticae
56 (2022) pp. 122–133
DOI: https://doi.org/10.33039/ami.2022.12.002
URL: https://ami.uni-eszterhazy.hu

Sensitivity analysis of a single server finite-source retrial queueing system with two-way communication and catastrophic breakdown using simulation

János Sztrik, Ádám Tóth

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

Abstract. In this paper, a finite-source retrial queueing system with two-way communication is investigated with the help of a simulation program of own. If a randomly arriving request from the finite-source finds the single server idle its service starts immediately, otherwise it joins an orbit from where it generates retrial/repeated calls after a random time. To increase the utilization of the server when it becomes idle after a random time an outgoing request is called for service from an infinity source. Upon its arrival if the server is busy, it goes to a buffer and when the server becomes idle again its service starts immediately. requests arriving from the finite-source and orbit are referred to as primary or incoming ones while requests called from the infinite source are referred to as secondary or outgoing requests, respectively. The service times of the primary and secondary requests are supposed to be random variables having different distributions. However, randomly catastrophic failures may happen to all the requests in the system, that is from the orbit, the service unit, and the buffer. In this case, the primary requests return to the finite-source, and the secondary ones are lost. The operation of the system is restored after a random time. Until the restoration is finished no arrivals and service take place in the system. All the above-mentioned times are supposed to be independent random variables.

The novelty of this paper is to perform a sensitivity analysis of the failure and restoration/repair times on the main characteristics to illustrate the effect of different distributions having the same average and variance value. Our aim is to determine the distribution of the number of requests in the system, the average response time of an arbitrary primary request without successful

Submitted: October 21, 2022
Accepted: December 3, 2022
Published online: December 28, 2022

service, also the average response time of an arbitrary and successfully served primary request, the total utilization of the service unit, or the probability that a primary request leaves the system without successful service because of a catastrophic event. Results are illustrated graphically obtained by our simulation program.

Keywords: finite-source queueing, two-way communication, catastrophic failure, restoration, sensitivity analysis, characteristics, simulation

1. Introduction

Retrial queues with two-way communication arose as stochastic models of call centers, where the operator can provide both inbound/incoming and outbound/outgoing calls. The idea of call blending is to improve the productivity of call centers by reducing the idle time of an operator was investigated among others in [2, 3, 5], and references cited in them.

However, from a practical point of view, it is also important to investigate situations where the server is not always able to serve the requests. There are many models and assumptions about the distribution of the operation and restoration time of the server. In case of a breakdown, there are many options corresponding to the behavior of request under service and the request generation process. In this paper, we deal with catastrophes, sometimes called disasters or negative requests which clear all the requests from the service facility, orbit, buffer, and stop the arrivals of the requests. The interested reader is referred to among others [1, 7, 8] and references cited in them.

In our earlier papers we dealt with finite-source single server two-way communication systems with an unreliable server under different repair options and request generation processes. With the help of simulation, the main characteristics were obtained and sensitivity analysis was carried out corresponding to failure and repair time distributions, see [9–11].

The primary aim of the present paper is to carry out a sensitivity analysis of the time of catastrophe and restoration/repair on the main characteristicsto illustrate the effect of different distributions having the same average and variance value. Our goal is to determine the distribution of the number of requests in the system, the average response time of an arbitrary primary request without successful service, also the average response time of arbitrary and successfully served primary request, the total utilization of the service unit, or the probability that a primary request leaves the system without successful service because of a catastrophic event. Results are illustrated graphically obtained by our simulation program.

2. System model

Figure 1 shows the behavior of the system with the aim that we are interested in investigating the effect of the catastrophes on the main characteristics. That is the reason that we assume exponentially distributed random variables except the distribution of the time of disaster. N sources generate requests after an exponentially distributed time with parameter λ independently of each other. If an arriving request finds the single server idle its service starts immediately, the services time is supposed to be exponentially distributed with parameter μ_1. If the serves is busy the call joins an orbit from where it generates retrial/repeated calls after an exponentially distributed time with parameter ν. To increase the utilization of the server when it becomes idle after an exponentially distributed time with parameter λ_2 an outgoing request is called for service from an infinity source. Upon its arrival, if the server is busy, it goes to a buffer and when the server becomes idle again its service starts immediately. The service time of this type of request is supposed to be exponentially distributed with parameter μ_2. Requests arriving from the finite-source and orbit are referred to as primary or incoming ones while requests called from the infinite source are referred to as secondary or outgoing requests, respectively.

However, randomly catastrophic failures may happen clearing all the requests in the system, that is from the orbit, the service unit, and the buffer. In this case, the primary requests return to the finite-source, and the secondary ones are lost. The operation of the system is restored after an exponentially distributed time with parameter γ_2. Until the restoration is finished no arrivals and service take place in the system. All the above-mentioned times are assumed to be independent random variables. Catastrophes can take place according to gamma, hypo-exponential, hyper-exponential, Pareto and lognormal distribution selecting their parameters to have the same average value.

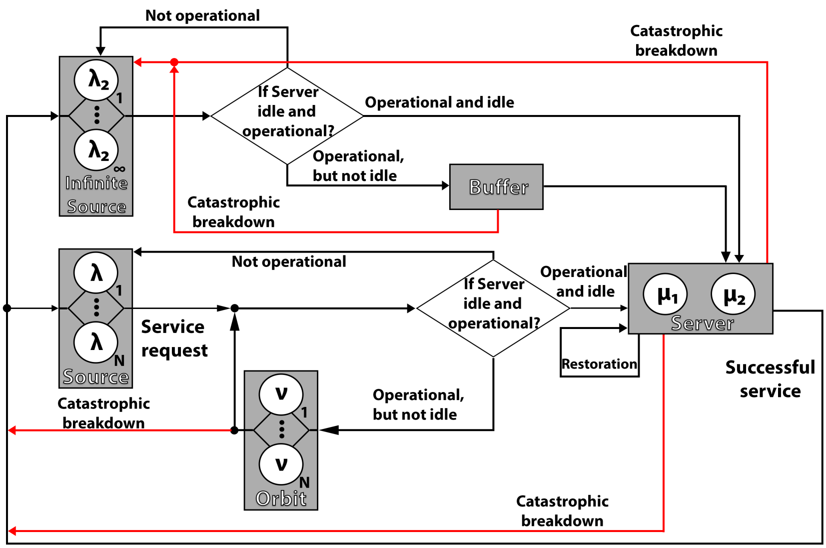

Figure 1. System model.

3. Simulation results and examples

We applied the simulation approach to obtain all the desired characteristics. Due to the simulation, we can deal with generally distributed random variables representing different times that occurred in the model construction. Expect the case when all the random variables are exponentially distributed it is very difficult, if not impossible to get an analytical solution to the characteristics. The estimation is carried out by applying a statistical package in which the method of batch averages is used, see [6]. First, we deal with exponentially distributed failure time with parameter γ_1 and show the effect of the failure rate on the probability that a primary request leaves the system without successful service, see Table 2. Then we turn our attention to generally distributed failure times when the (CV) squared coefficient of variation which is defined as variance/(square of average) is greater or less than one. In both cases we consider distributions with the same average and variances to show the effect of the particular distribution on some of the characteristics.

We must admit by choosing different input parameters our aim is to show how the system behaves and they are not realistic values since we do not have data for this type of system. In this phase the paper is more theoretic than practical.

3.1. Exponentially distributed failure times

In this part, the failure time is assumed to be exponentially distributed with parameter γ_1. The other input parameters are given in Table 1. This model was treated by the help of a software package called MOSEL (MOdeling, Specification and Evaluation Language) and served as a validation for the simulation, see [4].

Table 1. Numerical values of model parameters for exponentially distributed failure time.

N	$\boldsymbol{\lambda}$	$\boldsymbol{\lambda_2}$	$\boldsymbol{\mu_1}$	$\boldsymbol{\mu_2}$	$\boldsymbol{\nu}$	$\boldsymbol{\gamma_2}$
100	0.02	0.5	1	2.5	0.01	1

Table 2. Probability that a primary request departs because of a catastrophic event.

$\boldsymbol{\gamma_1}$	**P(departure)**
0.00001	0.002113
0.01	0.535419
0.1	0.724697

It should be mentioned that even for a small failure intensity the probability of departure is not negligible. In addition, in Figure 2 we can see how the distribution of the number of primary requests changes as the failure rate increases. In the case

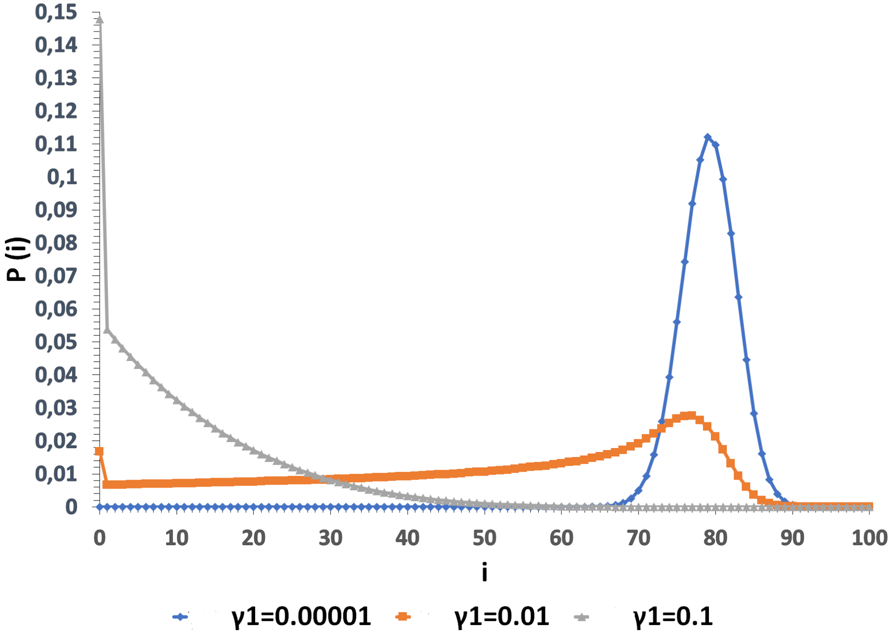

Figure 2. Distribution of the number of primary requests in the system.

of a very small value the distribution graph is similar to a normal distribution, but as we increase the failure rate the distribution is unknown.

3.2. Different distributions of failure time of the system, CV is greater than one

This part is devoted to the sensitivity analysis of the characteristics corresponding to the distribution of failure times. Table 3 shows the used parameter setting and Table 4 collects the values of parameters in the case of gamma, hyper-exponential, lognormal, and Pareto distributions. We assume that $CV > 1$ and to perform a valid comparison both the average value and variance are the same using different parameters' values.

Table 3. Numerical values of model parameters.

N	λ_2	μ_1	μ_2	ν	γ_2
100	0.5	1	2.5	0.01	1

The steady-state probability of the number of primary requests in the systems is presented in Figure 3 when $\lambda = 0.02$. Having the same average and variance,

Table 4. Parameters of failure time.

Distribution	Gamma	Hyper-exponential	Pareto	Lognormal
Parameters	$\alpha = 0.31225$ $\beta = 0.05588$	$p = 0.3619707$ $\lambda_1 = 0.1295528$ $\lambda_2 = 0.2283569$	$\alpha = 2.145538$ $k = 2.9835251$	$m = 1.0027833$ $\sigma = 1.1981970$
average	5.588			
Variance	100			
Squared CV	3.2024857438			

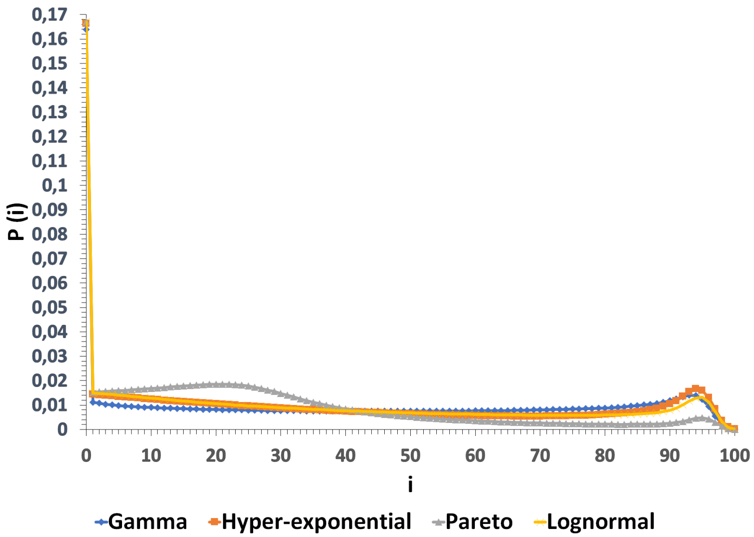

Figure 3. Distribution of the number of primary requests in the system.

the obtained results vary from each other which is especially true in the case of the Pareto distribution. This figure illustrates the impact of the selected distribution on the operation of the system, as was expected.

In Figures 4, 5 the average response time of a primary request and a primary request without successful service can be seen as the function of the arrival rate λ. Essential differences can be observed which is due to the distributions. Naturally, the average response time of requests without successful service should be greater as they leave the system because of catastrophes. Some of them can be in the orbit and one under service. Since the average failure time is 5.588 we expected that all the average response times are less than this value. However, it is true only for the Pareto distribution. It also looks surprising that three averages first increasing then decreasing, while in the Pareto case it is increasing. During several simulation runs, we realized that the behavior of the systems heavily depends on the variance of the failure time and the other input parameters of the system. Our explanation for the unexpected higher average response time is the following.

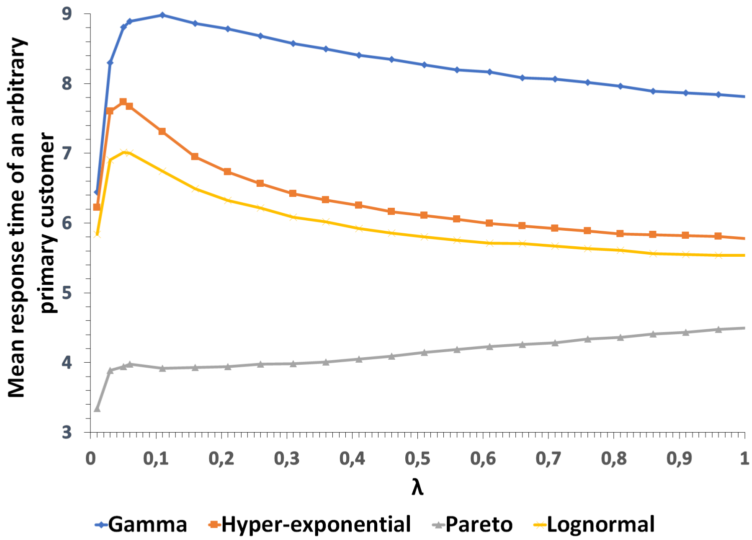

Figure 4. Average response time of a primary request.

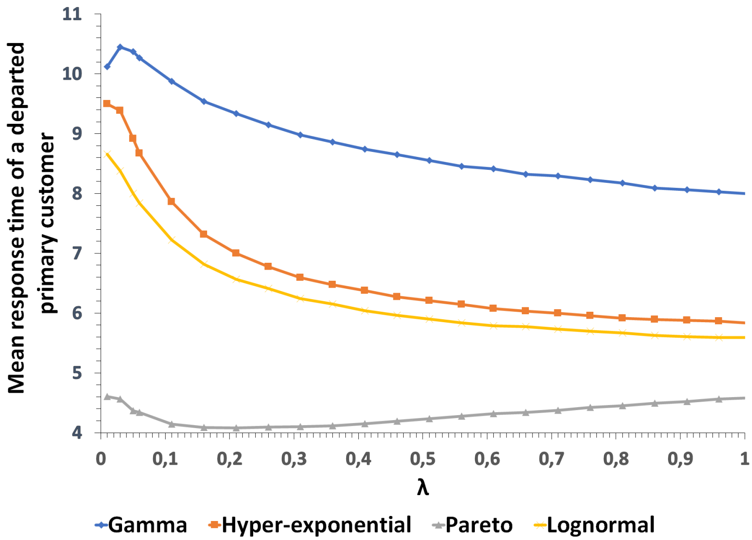

Figure 5. Average response time of a primary request without service.

Since the standard deviation of the operation time is almost two times higher than its average there will be short operation times in which there are no requests in the system, and there are long operation times with high response times. Thus the average response time can be greater than the average operation time. The

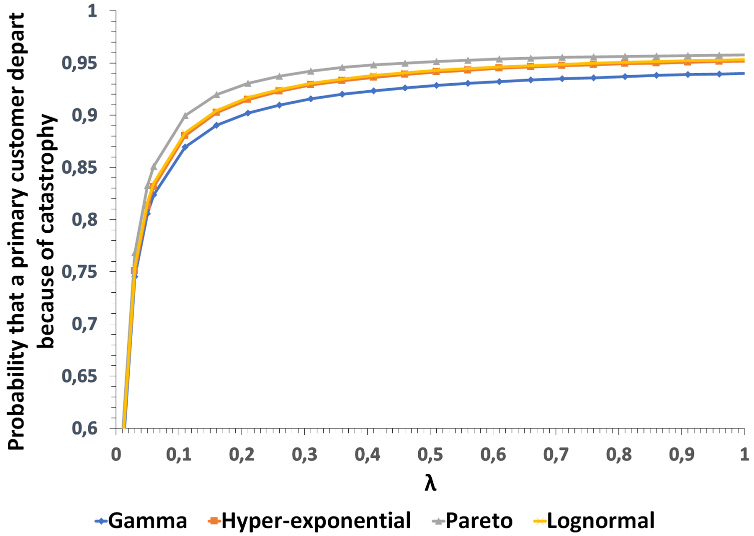

Figure 6. Probability that a primary request departs.

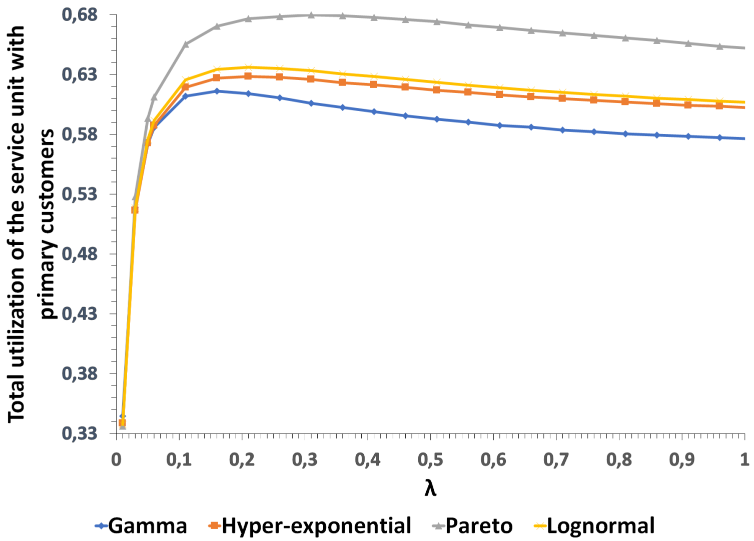

Figure 7. Total utilization w.r. primary requests.

maximum of the average happens only at special parameter setup.

Figure 6 shows the probability that a request departs from the system due to the catastrophe. There are differences between the distributions and of course the probability is an increasing function of the arrival rate from the source since more and more requests are in the system when a catastrophe happens.

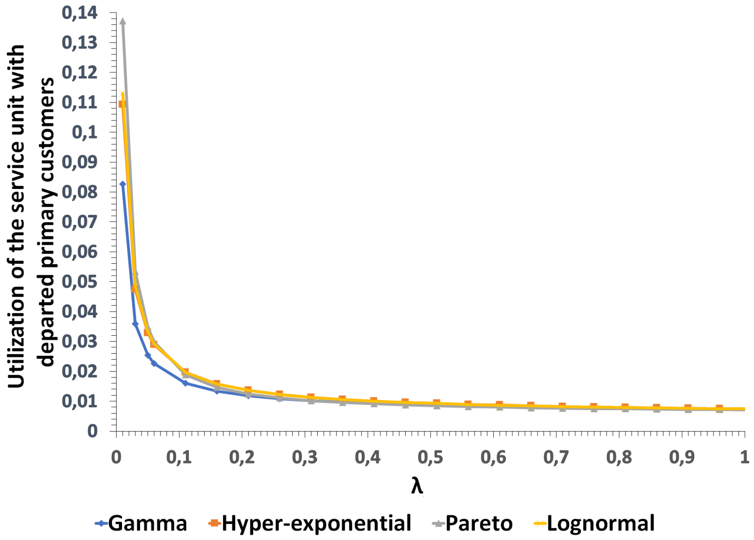

Figure 8. Total utilization w.r. primary requests without service.

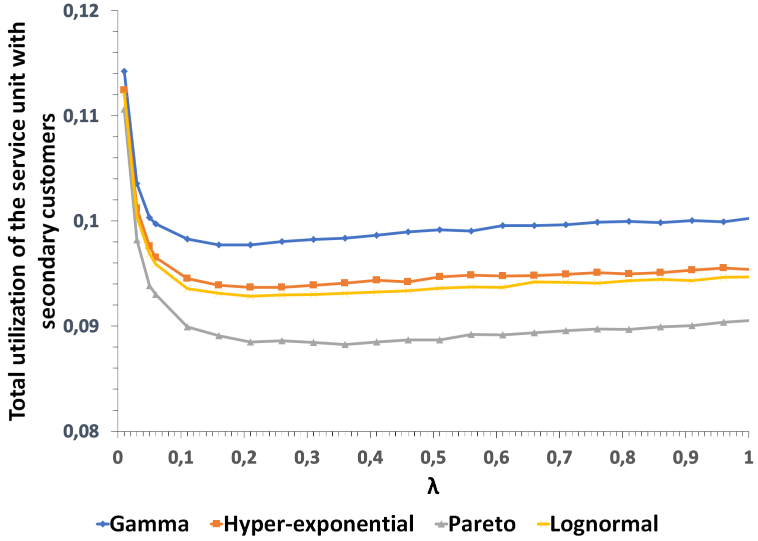

Figure 9. Total utilization w.r. secondary requests.

In Figures 7, 8, 9 utilization of the server corresponding to different types of requests is illustrated. As usual, the utilization of the server with respect to a certain type of request is defined as the probability that the server is busy with that type of request, respectively. There is a very special property of finite-source retrial queues, namely under special parameter setup the mean response time of a customer has maximum as the function of arrival rate from the source. We could find such a parameter setup that there is maximum of the utilization for the primary requests which includes requests with successful and without successful service, see Figure 7. In the first phase due to the increasing number of requests the utilization increases, but after a certain point due to the catastrophes many requests depart from the system and the utilization decreases.

Figure 8 shows the utilization corresponding to the departed requests due to the catastrophes. Since the number of requests in the system increases as the function of the arrival rate λ more and more requests depart the systems because of the failure, hence the utilization decreases. As we can observe this measure is almost the same for all distributions.

Finally, Figure 9 shows the utilization of the server with respect to the secondary requests invited when the server is idle. The behavior can be explained by the catastrophes since the server in this case is idle and there is more chance for a secondary request to occupy the server.

3.3. Different distributions of failure time of the system, CV is less than one

This part is devoted to the sensitivity analysis of the characteristics with respect to the distribution of failure times. Table 3 shows the used parameter setting and Table 5 collects the values of parameters in the case of gamma, hypo-exponential, lognormal, and Pareto distribution. $CV < 1$ and both the average value and variance are the same using different parameters' values.

Table 5. Parameters of failure time.

Distribution	Gamma	Hypo-exponential	Pareto	Lognormal
Parameters	$\alpha = 1.2320819$ $\beta = 0.2204778$	$\mu_1 = 0.2$ $\mu_2 = 1.7$	$\alpha = 2.4940153$ $k = 3.3475773$	$m = 1.423548$ $\sigma = 0.7708627$
average	5.588			
Variance	25.3460207612			
Squared CV	0.811634349			

Due to the lack of pages, we cannot show the same characteristics as we presented before. We can summarize the findings as follows. The average response times are not greater than the average operation time due to the smaller variance of the operation time. All the other characteristics show similar behavior with fewer differences between the different distributions. In general, performing several simulation runs we observed that the variance of the response times of requests behave similar way as the variance of the operation time either $CV > 1$ or $CV < 1$. One of the advantages of the simulation approach is that we can estimate any of the characteristics giving not only expected values but variances, too.

4. Conclusion

A finite-source retrial queueing system with two-way communication was investigated with the help of simulation. We were interested in carrying out a sensitivity analysis of the failure and restoration/repair times on the main characteristics to illustrate the effect of different distributions having the same average and variance value. We aimed to determine the distribution of the number of requests in the system, the average response time of an arbitrary primary request without successful service, also the average response time of arbitrary and successfully served primary request, the total utilization of the service unit, or the probability that a primary request leaves the system without successful service because of a catastrophic event. Results were illustrated graphically and some explanations were given. The scientific message of the this paper is following: from earlier papers published in different high level journals it can be seen that systems with catastrophic failures are important and needs investigations. The authors are not aware of any papers with two-way communications with this type of failures. In our opinion allowing non-exponentially distributed operation times the analytic solution is hopeless. The only way is the simulation method. It is a natural question to ask how the characteristics of the system depends on the distribution of the operation time assuming the same first two moments, respectively. That was our strong motivation and we are confident that this paper is a valuable contribution to this topic.

Acknowledgment The authors are grateful to the reviewers for their valuable comments and suggestions which improved the quality of the paper.

References

[1] S. R. Chakravarthy: *A catastrophic queueing model with delayed action*, Applied Mathematical Modelling 46 (2017), pp. 631–649.

[2] 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.

[3] J. Kim, B. Kim: *A survey of retrial queueing systems*, Annals of Operations Research 247.1 (2016), pp. 3–36, issn: 1572-9338.

[4] A. Kuki, T. Bérczes, J. Sztrik: *Modeling two-way communication systems with catastrophic breakdowns*, in: Information Technologies and Mathematical Modelling. Queueing Theory and Applications, in Press, Springer, 2022.

[5] 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.

[6] A. M. Law, W. D. Kelton: *Simulation modeling and analysis*, McGraw-Hill New York, 1991.

[7] K. Li, J. Wang: *Equilibrium balking strategies in the single-server retrial queue with constant retrial rate and catastrophes*, Quality Technology & Quantitative Management 18.2 (2021), pp. 156–178.

[8] S. Subramanian et al.: *A Stochastic Model for Automated Teller Machines Subject to Catastrophic Failures and Repairs*, Queueing Models and Service Management 1.1 (2018), pp. 75–94.

[9] J. Sztrik, Á. Tóth, Á. Pintér, Z. Bács: *Simulation of finite-source retrial queues with two-way communications to the orbit*, in: International Conference on Information Technologies and Mathematical Modelling, Springer, 2019, pp. 270–284.

[10] J. Sztrik, Á. Tóth, Á. Pintér, Z. Bács: *The Simulation of Finite-Source Retrial Queueing Systems with Two-Way Communications to the Orbit and Blocking*, in: International Conference on Distributed Computer and Communication Networks, Springer, 2020, pp. 171–182.

[11] Á. Tóth, J. Sztrik, A. Kuki, T. Bérczes, D. Efrosinin: *Reliability analysis of finite-source retrial queues with outgoing calls using simulation*, in: 2019 International Conference on Information and Digital Technologies (IDT), IEEE, 2019, pp. 504–511.