Российская академия наук (РАН)

Институт проблем управления им. В.А. Трапезникова Российской академии наук (ИПУ РАН)

Российский университет дружбы народов (РУДН)

Институт информационных и телекоммуникационных технологий Болгарской академии наук (София, Болгария)

Национальный исследовательский Томский государственный университет (НИ ТГУ)

Научно-производственное объединение «Информационные и сетевые технологии» («ИНСЕТ»)

РАСПРЕДЕЛЕННЫЕ КОМПЬЮТЕРНЫЕ И ТЕЛЕКОММУНИКАЦИОННЫЕ СЕТИ: УПРАВЛЕНИЕ, ВЫЧИСЛЕНИЕ, СВЯЗЬ (DCCN-2022)

Материалы XXV Международной научной конференции

Россия, Москва, 26–30 сентября 2022 г.

Под общей редакцией
д.т.н. *В.М. Вишневского* и д.т.н. *К.Е. Самуйлова*

Москва
Российский университет дружбы народов
2022

Russian Academy of Sciences (RAS)
V.A. Trapeznikov Institute of Control Sciences of RAS (ICS RAS)
Peoples' Friendship University of Russia (RUDN University)
Institute of Information and Communication Technologies
of Bulgarian Academy of Sciences (Sofia, Bulgaria)
National Research Tomsk State University (NR TSU)
Research and development company
"Information and networking technologies"

DISTRIBUTED COMPUTER AND COMMUNICATION NETWORKS: CONTROL, COMPUTATION, COMMUNICATIONS (DCCN-2022)

Proceedings of the XXV International Scientific Conference

Russia, Moscow, September 26–30, 2022

Under the general editorship
of D.Sc. *V.M. Vishnevskiy* and D.Sc. *K.E. Samouylov*

Moscow
Peoples' Friendship University of Russia
2022

УДК 004.7:004.4:621.391:007:001(063)
ББК 32.973.202:32.968+32.81
Р24

Под общей редакцией
д.т.н. *В.М. Вишневского* и д.т.н. *К.Е. Самуйлова*

Р24 **Распределенные компьютерные и телекоммуникационные сети : управление, вычисление, связь (DCCN-2022) = Distributed computer and communication networks : control, computation, communications (DCCN-2022)** : материалы XXV Международной научной конференции. Россия, Москва, 26–30 сентября 2022 г. / под общ. ред. В. М. Вишневского и К.Е. Самуйлова. – Москва : РУДН, 2022. – 439 с. : ил.

В научном электронном издании представлены материалы XXV Международной научной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь» по следующим направлениям:

- Алгоритмы и протоколы телекоммуникационных сетей;
- Управление в компьютерных и инфокоммуникационных системах;
- Анализ производительности, оценка QoS / QoE и эффективность сетей;
- Аналитическое и имитационное моделирование коммуникационных систем последующих поколений;
- Эволюция беспроводных сетей в направлении 5G;
- Технологии сантиметрового и миллиметрового диапазона радиоволн;
- RFID-технологии и их приложения;
- Интернет вещей и туманные вычисления;
- Системы облачного вычисления, распределенные и параллельные системы;
- Анализ больших данных;
- Вероятностные и статистические модели в информационных системах;
- Теория массового обслуживания, теория надежности и их приложения;
- Высотные беспилотные платформы и летательные аппараты: управление, передача данных, приложения.

В материалах научной конференции DCCN-2022 обсуждены перспективы развития и сотрудничества в этой сфере.

Издание предназначено для научных работников и специалистов в области управления крупномасштабными системами.

Текст воспроизводится в том виде, в котором представлен авторами.

Конференция организована при поддержке Программы стратегического академического лидерства РУДН.

ISBN 978-5-209-11456-7

UDC: 004.94

Performance analysis of a finite-source retrial queueing system with two-way communication, catastrophic breakdown and impatient customers using simulation

János Sztrik [1], Ádám Tóth [1], Ákos Pintér [1], Zoltán Bács[1]

[1]University of Debrecen, Debrecen 4032, Hungary

{toth.adam,sztrik.janos}@inf.unideb.hu, bacs.zoltan@econ.unideb.hu, apinter@science.unideb.hu

Abstract

An M/M/1//N retrial queueing system with two-way communication to the infinite source and impatient customers in the orbit is considered in the paper. There is a finite source in which the primary or regular customers are coming, while requests from the infinite source are the secondary customers. When the service unit is in an idle state, it may call a customer from the infinite source for service. Besides, the server is supposed to break down according to several distributions which have a specialty in removing all the customers located in the system. In the case of a faulty state, blocking is applied not allowing the customers into the system until the service unit fully recovers. This work concentrates on examining the effect of those distributions on several performance measures. The obtained results are graphically realized to show the differences and curiosities among the used parameter settings of the various distributions.

Keywords: Simulation, Catastrophic breakdown, Retrial queuing system, Collision, Impatience, Sensitivity analysis

1. Introduction

In many fields of our life, the phenomenon of waiting appears making inevitably to create queueing systems handling for example increasing network traffic in many info-communication systems. Throughout the years researchers have designed numerous tools and mechanisms which are suitable for modeling various organizations and one prime instance is retrial queueing systems that are capable of depicting arising real-life problems in telecommunication schemes like telephone switching systems, call centers, computer networks, and computer systems. Several publications exist where researchers exploit the advantages of retrial-queuing systems with repeated calls using for their models like in [1].

In the case of a retrial queueing system, a virtual waiting room is taken into consideration which is called the orbit meaning that whenever a service of a job can not start because of failure or occupation of the server it remains in the system. In the orbit, these customers have the opportunity to be at the service facility after a random time. The population of the customers is finite as the probability that a server calls a customer from the orbit is not very small and under such circumstances, it is more suitable to examine models with retrial queues. Exploring the available literature many paper have applied infinite and finite source queueing systems for example in [2].

It is also interesting to observe how the feature of two-way communication is used in papers in many fields of life. Its popularity originates from its usefulness to model applicable systems creating real-life applications. One prime example can be mentioned in the topic of telecommunication, especially in call centres where agents may be occupied with other particular labor during in an inactivity period like selling, advertising, and promoting products besides handling the calls of the customers. Optimizing the utilization of the service units or agents is always pivotal to increasing the efficiency of such systems. To mention some works about applying two-way communication schemes here are some instances [3], [4].

Random breakdowns and malfunctions occur in real-life scenarios caused by a power outage, human negligence, or other catastrophic activity. This greatly changes the operation of the system and the performance measures thus its investigation is necessary. In many cases, the service units are presumed to be accessible all the time which is not realistic. Systems with random failures have been investigated by many researchers for example in [5] [6]. However, there are certain situations where the effect of breakdown is different types of failures can be investigated. For instance, power outages or mechanical failures may cause catastrophic events in which all the customers in the system are removed. This is known as a negative customer and it takes out every other request from service upon its arrival. This eventuates a disaster event because it also breaks down the service unit and in this case, every customer is forwarded back to the source. Papers in connection with negative customers can be found in [7], [8].

The aim of this work is to realize a sensitivity analysis using various distributions of failure time on several performance measures while the departure of customers may happen. The results are obtained by our stochastic simulation program using the basics of SimPack ([9]) which contains the basic building blocks of a simulation model. This gives us the opportunity to model any type of queueing system to create any type of simulation model and we can calculate any performance measure using arbitrary random number generators for the desired random variable. The presented curves highlight both the effect of disaster events and the impatience of the customers applying various parameter settings and these figures concentrate on the interesting phenomena of these systems. The table of input parameters and graphical illustrations of the results are included demonstrating the influence of the used distributions on the main performance metrics.

2. System model

A finite-source retrial queueing system of type $M/M/1//N$ is regarded with an unreliable service unit and impatient customers. This model has a service unit and exactly N individual resides in the finite-source in which request generation (primary request) is proceeded towards the system according to exponential law with parameter λ. This means that the inter-arrival times are exponentially distributed with mean λ. If an arriving job finds the server in an idle state then its service starts immediately which is an exponentially distributed variable with μ_1. Otherwise, in vain of a queue, jobs are not lost but remain in the system being forwarded to the orbit which is a virtual waiting space. From there these requests after an exponentially random time with parameter ν retry to reach the service unit. After spending futilely an exponentially distributed time with rate τ in the orbit a customer may choose to leave the system without being served so in other words, every request has an impatience characteristic. It is assumed random breakdowns take place according to various distributions like gamma, hyper-exponential, Pareto, hypo-exponential, and lognormal. The parameters are chosen in a way that a real sensitivity analysis would be accomplished. In these occurrences, disaster events develop resulting in interruption of the service of a job and with the customers, in the orbit, they all depart from the system. Blocking is applied during faulty periods so no customers are allowed by the system until the server fully recovers.

The repair process begins to be executed after the failure of the service unit which happens according to an exponential distribution with parameter γ_2. Two-way communication was also introduced in our model, when the server becomes free it may call a request from an infinite source (secondary customer) after an exponentially random time with parameter λ_2. That type of customer occupies instantly the service unit if it is not busy upon its arrival, otherwise, it is forwarded to a special buffer where it waits there until the server turns idle. At that moment a secondary customer automatically enters the service facility. In the case of a catastrophic event, every primary job returns to the finite-source, and every secondary customer exits from the system including the one who is under service. The service time of the secondary customer also follows an exponential distribution with the rate of μ_2. Every appeared arbitrary variable in the model construction is supposed to be independent of each other.

3. Simulation results

As mentioned earlier SimPack was used as a base of our program and we used a statistical package that was responsible for obtaining the desired performance measures.

The confidence level of 99.9% is employed throughout the simulations and 0.00001 is the amount of the relative half-width of the confidence interval to pause the actual simulation sequence. The size of the batch in the initial transient period can not be too small therefore its value is set to 1000.

In Table 1 the used values of input parameters are presented.

Table 1. Numerical values of model parameters

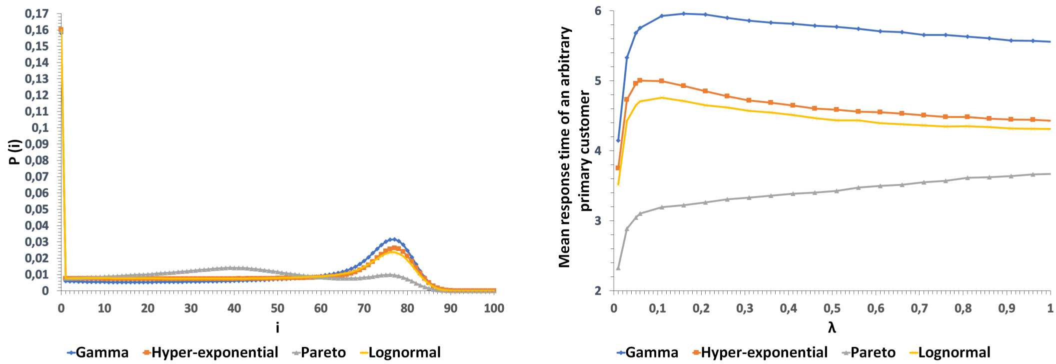
N	ν	μ_1	μ_2	τ	γ_2	λ_2
100	0.01	1	2.5	0.05	1	0.5

The next table (Table 2) consists of the parameters of failure time, every chosen parameter is according to have the same mean and variance value in that way a valid comparison is achieved. The simulation program was tested by many parameter values and the reason for selecting these values is focusing on interesting situations besides that it is worth mentioning that almost the same phenomenon appeared as with this particular setting. The squared coefficient of variation is more than one in this scenario which is totally intentional to check the influence of peculiar random variables.

Table 2. Parameters of failure time

Distribution	Gamma	Hyper-exponential	Pareto	Lognormal
Parameters	$\alpha = 0.31225$ $\beta = 0.05588$	$p = 0.36197$ $\lambda_1 = 0.12955$ $\lambda_2 = 0.22835$	$\alpha = 2.1455$ $k = 2.9835$	$m = 1.00278$ $\sigma = 1.19819$
Mean	5.558			
Variance	100			
Squared coefficient of variation	3.2024857438			

Due to the page limitation we involved just two figures, the others and the results of other scenario will be displayed in the extended version of the paper. In Figure 1a i represents the number of primary customers in the system on the X-axis, and $P(i)$ denotes the probability that exactly i primary customers are situated at the server and in the orbit altogether on the Y-axis. The distribution of the number of primary customers in the system is displayed when λ is 0.11 using various distributions of failure time. The mean number of primary customers in the system differs from each other greatly. In the case of the gamma distribution, customers tend to spend more time in the system compared to Pareto distribution. It is also noticeable that the highest probability is at 0 and this can be explained by the fact that during faulty periods customers are not allowed to enter and for every catastrophic breakdown the system is emptied.



(a) Distribution of the number of primary customers in the system.

(b) Average response time of an arbitrary primary customer.

Fig. 1. The effect of different distributions

The expected response time of an arbitrary primary customer is presented in the function of the arrival intensity of incoming primary request in Figure 1b. Even though the mean and the variance value are equal with each other, huge gaps develop among the applied distributions, at gamma the highest average response times can be observed compared to the others. Also with the increment of the arrival intensity, the expected response time of an arbitrary primary customer starts to increase then after a certain intensity arrival ($\lambda = 0.07$) it decreases except in the case of Pareto.

Although the mean operation time is 5.558 the values of the expected response times are higher than that which is quite interesting. Our intuition is that this can be explained by that the variance is quite high resulting in many small operation times and in most of them no customer can enter because they are so small. But there are several high operation periods in which it is very probable that many jobs enter and spend relatively a high amount of time.

4. Conclusion

A finite-source retrial queueing system is introduced with a non-reliable server, a two-way communication scheme, and impatient customers. We investigated a scenario where different parameters are used to carry out a sensitivity analysis to figure out how the different performance measures develop. The results are obtained by our simulation program and several graphical figures depict the effect of using various distributions of failure time on the expected response time of primary and on the distribution of customers in the system. In our figures, the differences were clearly seen among the values of several performance measures when the squared coefficient of variation is greater than one showing how pivotal applying a distribution. The curves also reveal the impact of impatience on reducing the average response time of a primary customer. In the future we plan to continue our research work, examining other types of finite-source retrial queuing systems with two-way communication or adding another service unit for backup purposes.

REFERENCES

1. J. Kim, B. Kim, A survey of retrial queueing systems, Annals of Operations Research 247 (1) (2016) 3–36.
2. T. Takeda, T. Yoshihiro, A distributed scheduling through queue-length exchange in CSMA-based wireless mesh networks, Journal of Information Processing 25 (2017) 174–181.
3. 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.
4. 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: V. M. Vishnevskiy, K. E. Samouylov, D. V. Kozyrev (Eds.), Distributed Computer and Communication Networks: Control, Computation, Communications, Springer International Publishing, Cham, 2020, pp. 171–182.
5. V. I. Dragieva, Number of retrials in a finite source retrial queue with unreliable server., Asia-Pac. J. Oper. Res. 31 (2) (2014) 23. doi:10.1142/S0217595914400053.
6. A. Tóth, J. Sztrik, A. Pintér, Z. Bács, Reliability analysis of finite-source retrial queuing system with collisions and impatient customers in the orbit using simulation, in: 2021 International Conference on Information and Digital Technologies (IDT), 2021, pp. 230–234. doi:10.1109/IDT52577.2021.9497567.
7. U. C. Gupta, N. Kumar, F. P. Barbhuiya, A Queueing System with Batch Renewal Input and Negative Arrivals, Springer Singapore, Singapore, 2020, pp. 143–157. doi:10.1007/978-981-15-5951-8_10.
8. S. R. Chakravarthy, S. Subramanian, A stochastic model for automated teller machines subject to catastrophic failures and repairs, Industrial & Manufacturing Engineering Publications 1 (1) (2018) 75–94.
9. P. A. Fishwick, Simpack: Getting started with simulation programming in c and c++, in: In 1992 Winter Simulation Conference, 1992, pp. 154–162.