

Asymptotic sojourn time analysis of finite-source M/M/1 retrial queueing system with collisions and server subject to breakdowns and repairs

Anatoly Nazarov, János Sztrik, Anna Kvach & Ádám Tóth

Annals of Operations Research

ISSN 0254-5330

Ann Oper Res

DOI 10.1007/s10479-019-03463-0



Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.



Asymptotic sojourn time analysis of finite-source $M/M/1$ retrial queueing system with collisions and server subject to breakdowns and repairs

Anatoly Nazarov¹ · János Sztrik² · Anna Kvach¹ · Ádám Tóth²

© The Author(s) 2019

Abstract

The aim of the present paper is to investigate the steady-state distribution of response and waiting time in a finite-source $M/M/1$ retrial queueing system with collision of customers where the server is subjects to random breakdowns and repairs depending on whether it is idle or busy. An asymptotic method is applied under the condition that the number of sources tends to infinity, the primary request generation rate, retrial rate tend to zero while service rate, failure rates, repair rate are fixed. As the result of the analysis it is shown that the steady-state probability distribution of the number of transitions/retrials of the customer into the orbit is geometric with a given parameter, and the normalized sojourn time of the customer in the system follows a generalized exponential distribution. It is also proved that the limiting distributions of the normalized sojourn time of the customer in the system and the normalized sojourn/waiting time of the customer in the orbit coincide. The novelty of this investigation is the introduction of failure and repair of the server. Approximations of prelimit distributions obtained with the help of stochastic simulation by asymptotic one are considered and several illustrative examples show the accuracy and range of applicability of the proposed asymptotic method.

Keywords Finite-source queueing system · Retrial queue · Collision · Server breakdowns and repairs · Asymptotic analysis · Limiting distribution · Approximation of distributions · Accuracy and area of applicability of approximations

✉ János Sztrik
sztrik.janos@inf.unideb.hu

Anatoly Nazarov
nazarov.tsu@gmail.com

Anna Kvach
kvach_as@mail.ru

Ádám Tóth
toth.adam@inf.unideb.hu

¹ National Research Tomsk State University, 36 Lenina Ave., Tomsk, Russia 634050

² University of Debrecen, Debrecen, Hungary

1 Introduction

Retrial queues, that is queues with repeated attempts have been effectively used to model many problems arising in telephone switching systems, telecommunication networks, computer networks and computer systems, call centers, wireless communication systems, etc. For a systematic account of the fundamental methods and the latest results, furthermore an accessible classified bibliography on this topic the interested reader is referred to, for example Artalejo and Gómez-Corral (2008), Falin and Templeton (1997), Gómez-Corral and Phung-Duc (2016), Kim and Kim (2016), and references therein.

In many practical cases it is essential to take into account that the rate of generation of new primary calls decreases as the number of customers in the system increases. This can be done by means of finite-source, or quasi-random input models. Retrial queues with quasi-random input are recent interest in modeling magnetic disk memory systems, cellular mobile networks, computer networks, and local-area networks with non-persistent CSMA/CD protocols, with star topology, with random access protocols, and with multiple-access protocols, see, for example Alfa and Isotupa (2004), Ali and Wei (2015), Almási et al. (2016), Do et al. (2014), Dragieva (2016), Ikhlef et al. (2016) and Wüchner et al. (2010).

Since in real life some components of the systems are subject to random breakdowns it is important to study reliability of retrial queues with server breakdowns and repairs because of limited ability of repairs and heavy influence of the breakdowns on the performance measures of the system. Finite-source retrial queues with an unreliable server have been investigated in several recent papers, for example Almási et al. (2005), Dragieva (2014), Gharbi and Dutheillet (2011), Gharbi and Ioualalen (2006), Sztrik (2005), Sztrik et al. (2006), Wang et al. (2010, 2011) and Zhang and Wang (2013).

In many communications involving data transmission from diverse sources there can be conflict for a limited number of channels or other facilities. Recent results on retrial queues with collisions can be found in, for example Ali and Wei (2015), Balsamo et al. (2013), Choi et al. (1992), Gómez-Corral (2010), Kim (2010), Kumar et al. (2010a, b), Nazarov et al. (2018) and Peng et al. (2014).

It is well-known that the analysis of waiting/response time and the number of retrials of a customer is much more complicated than the distribution of number of customers in the system. There exist analytic, numerical, and asymptotic methods to help this research, see for example Amador (2010), Artalejo (1998), Artalejo and Gómez-Corral (2008), Dragieva (2013, 2014, 2016), Falin and Artalejo (1998), Falin and Templeton (1997), Falin (1977, 1984, 1986, 1988), Wang et al. (2011) and Zhang and Wang (2013).

The aim of the present paper is to investigate such systems which has the above mention properties, that is finite-source, retrial, collisions, and non-reliability of the server. The present model is a generalization of the systems treated in Kvach and Nazarov (2015), Nazarov et al. (2014), Nazarov and Sudyko (2010), Nazarov and Moiseeva (2006) and Sudyko et al. (2018) and it is the natural continuation of the paper (Nazarov et al. 2018) in which the asymptotic distribution of the number of customers in the system has been investigated.

The rest of the paper is organized as follows. In Sects. 2 and 3 the full description of the model is given by the help of the corresponding two-dimensional Markov chain and the corresponding Kolmogorov equations are derived. Sections 4 and 5 are devoted to sojourn time analysis in the systems and the distribution of number of transitions the tagged customer carries until the successful completion of its service. In Sect. 6 the limiting distributions of the normed sojourn time of a request spends in the systems and in the orbit is treated, respectively. Section 7 deals with several numerical examples and comparisons to simulation

results showing the advantage of the asymptotic methods, and some comments are made. Finally, the paper ends with a Conclusion.

2 Model description and notations

Let us consider a retrial queuing system of type $M/M/1//N$ with collision of the customers and an unreliable server. The number of sources is N and each of them can generate a primary request with rate λ/N . A source cannot generate a new call until end of the successful service of this customer. If a primary customer finds the server idle, he enters into service immediately, in which the required service time is assumed to be exponentially distributed random variable with parameter μ . Otherwise, if the server is busy, an arriving (primary or repeated) customer involves into collision with customer under service and they both move into the orbit. The retrial time of requests are assumed to be exponentially distributed with rate σ/N . We suppose that the server is unreliable and the lifetime is exponentially distributed random variable with failure rate γ_0 if the server is idle and with rate γ_1 if it is busy. When the server breaks down, it is immediately sent for repair and the recovery time is assumed to be exponentially distributed with rate γ_2 . We deal with the case when the server is down all sources continue generation of customers and send it to the server, similarly customers may retry from the orbit to the server but all arriving customers immediately go into the orbit. Furthermore, in this unreliable model we suppose the interrupted request enters into the orbit immediately and its next service is independent of the interrupted one. All random variables involved in the model construction are assumed to be independent of each other.

Assuming that the systems is operating in steady-state the aim of the present paper is to find the limiting probability distribution of the sojourn time of the tagged customer in the system under conditions of $N \rightarrow \infty$. Furthermore, we can obtain the limiting probability distribution of the number of transitions of the tagged customer into the orbit, and, subsequently, the limiting probability distribution of the sojourn time/waiting time of the customer in the orbit. With the help of a simulation program developed to these purposes we can determine the accuracy and range of applicability of the asymptotic results in prelimit situation.

3 First order asymptotics for the number of customers in the system

Let $i(t)$ be the number of customers in the system at time t , that is, the total number of customers in orbit and in service, and similarly let us denote by $k(t)$ the server state as follows

$$k(t) = \begin{cases} 0, & \text{if the server is free,} \\ 1, & \text{if the server is busy,} \\ 2, & \text{if the server is down (under repair).} \end{cases}$$

We can state the following theorem which has been published and proved in our recent paper (Nazarov et al. 2018). We need this Theorem for the investigation of the waiting time.

Theorem 1 *Let $i(\infty)$ be the number of customers in the system in steady state, then*

$$\lim_{N \rightarrow \infty} E \exp \left\{ jw \frac{i(\infty)}{N} \right\} = \exp \{ jw \kappa_1 \}, \tag{1}$$

where the value of parameter κ_1 is the positive solution of equation

$$(1 - \kappa_1) \lambda - \mu R_1(\kappa_1) = 0, \tag{2}$$

and the stationary distributions of probabilities $R_k(\kappa_1)$ of the service state k are defined as follows

$$R_0(\kappa_1) = \left\{ \frac{\gamma_0 + \gamma_2}{\gamma_2} + \frac{\gamma_1 + \gamma_2}{\gamma_2} \cdot \frac{a(\kappa_1)}{a(\kappa_1) + \gamma_1 + \mu} \right\}^{-1},$$

$$R_1(\kappa_1) = \frac{a(\kappa_1)}{a(\kappa_1) + \gamma_1 + \mu} \cdot R_0(\kappa_1), \quad (3)$$

$$R_2(\kappa_1) = \frac{1}{\gamma_2} [\gamma_0 R_0(\kappa_1) + \gamma_1 R_1(\kappa_1)],$$

here $a(\kappa_1)$ is

$$a(\kappa_1) = (1 - \kappa_1)\lambda + \sigma\kappa_1. \quad (4)$$

The limiting mean value κ_1 allows us to execute further research of the sojourn time of the customer in system, number of transitions of the tagged customer into the orbit, and also the sojourn/waiting time of the customer in the orbit.

4 Sojourn time of the customer in the system

Let us denote by T the steady-state sojourn time of the customer in the system, i.e. T is interval length from the moment of the end of generation of the customer by a source until completion of its successful service.

First of all, we will find limiting mean sojourn time of the customer in the system with the help of Little's formula.

Since

$$(1 - \kappa_1)\lambda \mathbb{E} \left(\frac{1}{N} T \right) = \kappa_1,$$

therefore

$$\mathbb{E} \left(\frac{1}{N} T \right) = \frac{\kappa_1}{(1 - \kappa_1)\lambda}, \quad (5)$$

where κ_1 is asymptotic average value of number of customers in the system, and $(1 - \kappa_1)\lambda$ is the mean arrival intensity of the incoming flow.

However, we should underline that the above equality (5) defines only mean sojourn time of the customer in the system. We would like to carry a more detailed investigation for T of the tagged customer.

The total sojourn time T is simply expressed through residual or virtual sojourn time $T_{res}(t)$. Here $T_{res}(t)$ is interval length from the moment t until the end of the successful service of the tagged customer.

To study the sojourn time we will define the server states as follows:

$$s(t) = \begin{cases} 0, & \text{server is free,} \\ 1, & \text{server is busy, but not by the tagged customer,} \\ 2, & \text{server is down (under repair),} \\ 3, & \text{server is busy by the tagged customer.} \end{cases}$$

Let us introduce the following conditional characteristic functions

$$G_k(u, i) = \mathbb{E} \left\{ e^{juT_{res}(t)} | s(t) = k, i(t) = i \right\}, \quad (6)$$

and assuming that system is operating in steady state, it is not difficult to write the following system of Kolmogorov equations

$$\begin{aligned}
 & \left[ju - \frac{N-i}{N}\lambda - \frac{i}{N}\sigma - \gamma_0 \right] G_0(u, i) + \frac{N-i}{N}\lambda G_1(u, i+1) + \gamma_0 G_2(u, i) \\
 & + \frac{i-1}{N}\sigma G_1(u, i) + \frac{\sigma}{N} G_3(u, i) = 0, \\
 & \left[ju - \frac{N-i}{N}\lambda - \frac{i-1}{N}\sigma - \gamma_1 - \mu \right] G_1(u, i) + \frac{N-i}{N}\lambda G_0(u, i+1) \\
 & + \frac{i-1}{N}\sigma G_0(u, i) + \gamma_1 G_2(u, i) + \mu G_0(u, i-1) = 0, \\
 & \left[ju - \frac{N-i}{N}\lambda - \gamma_2 \right] G_2(u, i) + \frac{N-i}{N}\lambda G_2(u, i+1) + \gamma_2 G_0(u, i) = 0 \\
 & \left[ju - \frac{N-i}{N}\lambda - \frac{i-1}{N}\sigma - \gamma_1 - \mu \right] G_3(u, i) \\
 & + \frac{N-i}{N}\lambda G_0(u, i+1) + \frac{i-1}{N}\sigma G_0(u, i) + \gamma_1 G_2(u, i) + \mu = 0. \tag{7}
 \end{aligned}$$

The solution of system (7) for finite values N causes difficulties therefore we will find their solution under condition of unlimited growing number of sources, that is as $N \rightarrow \infty$.

For the asymptotic analysis of the sojourn time of the customers in the system we have the following statement

Theorem 2

$$\lim_{N \rightarrow \infty} E \exp \left\{ jw \frac{T}{N} \right\} = q + (1-q) \frac{\sigma q}{\sigma q - jw}, \tag{8}$$

where the value of parameter q is defined by expression

$$q = \frac{(1 - \kappa_1)\lambda}{(1 - \kappa_1)\lambda + \sigma \kappa_1}. \tag{9}$$

Proof Denoting $\frac{1}{N} = \varepsilon$, in the system (7) let us execute the following substitutions

$$u = \varepsilon w, \quad \varepsilon i = x, \quad G_k(u, i) = g_k(w, x, \varepsilon), \tag{10}$$

then (7) can be rewritten as

$$\begin{aligned}
 & [j\varepsilon w - (1-x)\lambda - x\sigma - \gamma_0] g_0(w, x, \varepsilon) + (1-x)\lambda g_1(w, x + \varepsilon, \varepsilon) + \gamma_0 g_2(w, x, \varepsilon) \\
 & + (x - \varepsilon)\sigma g_1(w, x, \varepsilon) + \varepsilon\sigma g_3(w, x, \varepsilon) = 0, \\
 & [j\varepsilon w - (1-x)\lambda - x\sigma - \gamma_1 - \mu + \varepsilon\sigma] g_1(w, x, \varepsilon) + (1-x)\lambda g_0(w, x + \varepsilon, \varepsilon) \\
 & + (x - \varepsilon)\sigma g_0(w, x, \varepsilon) + \gamma_1 g_2(w, x, \varepsilon) + \mu g_0(w, x - \varepsilon, \varepsilon) = 0, \\
 & [j\varepsilon w - (1-x)\lambda - \gamma_2] g_2(w, x, \varepsilon) + (1-x)\lambda g_2(w, x + \varepsilon, \varepsilon) + \gamma_2 g_0(w, x, \varepsilon) = 0 \\
 & [j\varepsilon w - (1-x)\lambda - x\sigma - \gamma_1 - \mu + \varepsilon\sigma] g_3(w, x, \varepsilon) + (1-x)\lambda g_0(w, x + \varepsilon, \varepsilon) \\
 & + (x - \varepsilon)\sigma g_0(w, x, \varepsilon) + \gamma_1 g_2(w, x, \varepsilon) + \mu = 0. \tag{11}
 \end{aligned}$$

Denoting $\lim_{\varepsilon \rightarrow 0} g_k(w, x, \varepsilon) = g_k(w, x)$ and $(1-x)\lambda + x\sigma = a(x)$, carrying out the specified limiting transition in system (11), we obtain

$$\begin{aligned} & -[a(x) + \gamma_0]g_0(w, x) + a(x)g_1(w, x) + \gamma_0g_2(w, x) = 0, \\ & -[a(x) + \gamma_1 + \mu]g_1(w, x) + [a(x) + \mu]g_0(w, x) + \gamma_1g_2(w, x) = 0, \\ & -\gamma_2g_2(w, x) + \gamma_2g_0(w, x) = 0, \\ & -[a(x) + \gamma_1 + \mu]g_3(w, x) + a(x)g_0(w, x) + \gamma_1g_2(w, x) + \mu = 0. \end{aligned}$$

From the first three equations of this system it follows that the functions $g_0(w, x)$, $g_1(w, x)$ and $g_2(w, x)$ coincide, and let us denote by $g(w, x)$ their common value, that is equalities are carried out

$$g_0(w, x) = g_1(w, x) = g_2(w, x) = g(w, x), \quad (12)$$

and from the last equation it is possible to write down

$$g_3(w, x) = \frac{\mu + (a(x) + \gamma_1)g(w, x)}{a(x) + \gamma_1 + \mu}. \quad (13)$$

The solution $g_k(w, x, \varepsilon)$ of the system (11) can be obtained in the form of decomposition

$$\begin{aligned} g_k(w, x, \varepsilon) &= g(w, x) + \varepsilon f_k(w, x) + o(\varepsilon), \quad k = \overline{0, 2}, \\ g_3(w, x, \varepsilon) &= g_3(w, x) + \varepsilon f_3(w, x) + o(\varepsilon), \end{aligned} \quad (14)$$

where $g_3(w, x)$ was expressed above through function $g(w, x)$ by (13). Substituting this decompositions into the first three equations of the system (11), we get the following equalities

$$\begin{aligned} & [j\varepsilon w - a(x) - \gamma_0]g(w, x) - \varepsilon[a(x) + \gamma_0]f_0(w, x) \\ & + (1-x)\lambda \left\{ g(w, x) + \varepsilon \frac{\partial g(w, x)}{\partial x} + \varepsilon f_1(w, x) \right\} \\ & + (x-\varepsilon)\sigma g(w, x) + \gamma_0 \{g(w, x) \\ & + \varepsilon f_2(w, x)\} + \varepsilon x\sigma f_1(w, x) + \varepsilon \sigma g_3(w, x) = o(\varepsilon), \\ & [j\varepsilon w - a(x) - \gamma_1 - \mu]g(w, x) - \varepsilon[a(x) + \gamma_1 + \mu]f_1(w, x) \\ & + (1-x)\lambda \left\{ g(w, x) + \varepsilon \frac{\partial g(w, x)}{\partial x} + \varepsilon f_0(w, x) \right\} + (x-\varepsilon)\sigma g(w, x) \\ & + \gamma_1 \{g(w, x) + \varepsilon f_2(w, x)\} + \varepsilon x\sigma f_0(w, x) \\ & + \mu \left\{ g(w, x) + \varepsilon \frac{\partial g(w, x)}{\partial x} - \varepsilon f_0(w, x) \right\} = o(\varepsilon), \\ & [j\varepsilon w - (1-x)\lambda - \gamma_2]g(w, x) - \varepsilon[(1-x)\lambda + \gamma_2]f_2(w, x) \\ & + (1-x)\lambda \left\{ g(w, x) + \varepsilon \frac{\partial g(w, x)}{\partial x} + \varepsilon f_2(w, x) \right\} \\ & + \gamma_2 \{g(w, x) + \varepsilon f_0(w, x)\} = o(\varepsilon). \end{aligned}$$

Equating here coefficients at identical degrees of ε , we receive

$$- [a(x) + \gamma_0]f_0(w, x) + (1-x)\lambda f_1(w, x) + \gamma_0 f_2(w, x)$$

$$\begin{aligned}
 &= (\sigma - jw)g(w, x) - (1 - x)\lambda \frac{\partial g(w, x)}{\partial x} - \sigma g_3(w, x), \\
 &\quad - [a(x) + \gamma_1 + \mu] f_1(w, x) + [a(x) + \mu] f_0(w, x) + \gamma_1 f_2(w, x) \\
 &= [\mu - (1 - x)\lambda] \frac{\partial g(w, x)}{\partial x} - jwg(w, x), \\
 &\quad - \gamma_2 f_2(w, x) + \gamma_2 f_0(w, x) = -jwg(w, x) - (1 - x)\lambda \frac{\partial g(w, x)}{\partial x}.
 \end{aligned}$$

Multiplying the first equality by $R_0(x)$, second by $R_1(x)$, third by $R_2(x)$ and adding them, we obtain equation

$$[\mu R_1(x) - (1 - x)\lambda] \frac{\partial g(w, x)}{\partial x} + [\sigma R_0(x) - jw]g(w, x) - \sigma R_0(x)g_3(w, x) = 0.$$

In this equation let us set $x = \kappa_1$ and denote

$$g(w, \kappa_1) = g(w), \quad g_3(w, \kappa_1) = g_3(w), \quad a(\kappa_1) = a, \quad R_k(\kappa_1) = R_k,$$

then we can write it in the form

$$(\sigma R_0 - jw)g(w) = \sigma R_0 g_3(w).$$

Then using (13) we get

$$\left(\sigma R_0 \frac{\mu}{a + \gamma_1 + \mu} - jw \right) g(w) = \sigma R_0 \frac{\mu}{a + \gamma_1 + \mu}. \tag{15}$$

Here carry out the equality

$$R_0 \frac{\mu}{a + \gamma_1 + \mu} = \frac{\mu}{a} R_0 \frac{a}{a + \gamma_1 + \mu} = \frac{\mu R_1}{a} = \frac{\lambda(1 - \kappa_1)}{\lambda(1 - \kappa_1) + \sigma \kappa_1} = q,$$

which coincides with (9), therefore from equalities (15) and (13) we obtain

$$g(w) = \frac{\sigma q}{\sigma q - jw}, \quad g_3(w) = \frac{\mu}{a + \gamma_1 + \mu} + \frac{\gamma_1 + \mu}{a + \gamma_1 + \mu} g(w). \tag{16}$$

Since R_0 is the probability that an arriving customer finds the server idle, hence $1 - R_0$ is the probability that it goes into the orbit. For the limiting characteristic function of the total sojourn time T by using the law of total probability we have

$$\lim_{N \rightarrow \infty} E \exp \left\{ jw \frac{T}{N} \right\} = R_0 g_3(w) + (1 - R_0)g(w) = q + (1 - q) \frac{\sigma q}{\sigma q - jw},$$

which coincides with (8).

The theorem is proved. □

Definition 1 The probability distribution, whose characteristic function has the form

$$h(w) = q + (1 - q) \frac{\alpha}{\alpha - jw},$$

is called a Generalized Exponential (GE) distribution with parameters q and α . Here α is the parameter of an exponential distribution. It easy to see that it is a mixed-type distribution, the random variable takes a zero value with probability q and exponentially distributed with parameter α with probability $1 - q$. More information about GE distribution can be found in, for example Kouvatsos (1994) and Wüchner et al. (2010).

Corollary 1 *Characteristic function of the sojourn time of the customer in the system in a preliminary situation of finite N can be approximated by a function of the form*

$$E e^{juT} \approx q + (1 - q) \frac{\sigma q}{\sigma q - juN}, \quad (17)$$

which follows obviously from Eq. (8).

Let us note that the form of the function (17) is determined by a single parameter (9)

$$q = \frac{(1 - \kappa_1)\lambda}{(1 - \kappa_1)\lambda + \sigma\kappa_1}, \quad (18)$$

which depends only on the asymptotic average κ_1 , defined in Theorem 1. That is for finding the function (17) it is enough to execute a first-order asymptotics, find value κ_1 and applying equality (18) we obtain value of the parameter q determining the form of the function (17).

In the following let us find the average value of the normalized sojourn time of the tagged customer in the system which is

$$v = E \left(\frac{1}{N} T \right) = (1 - q) \frac{1}{\sigma q} = \frac{1}{\sigma} \cdot \frac{1 - q}{q} = \frac{1}{\sigma} \frac{\sigma\kappa_1}{(1 - \kappa_1)\lambda} = \frac{\kappa_1}{(1 - \kappa_1)\lambda},$$

that coincides with the result (5), obtained by the Little's formula.

5 Distribution of the number of transitions/retrials of the tagged customer into the orbit

Sojourn time T of the customer in the system consists of total sojourn/waiting time W of the customer in the orbit and the total sojourn time of the customer under service. For systems with collisions and an unreliable server the sojourn time of the customer at the server has a rather complex structure, since it contains terms of zero duration when the customer from the orbit finds the server broken or busy by another customer. It has a non-zero terms of the services interrupted by collisions and failures of the server and finally a single term of successfully completed service, after which the customer leaves the system.

The structure of the sojourn/waiting time W of the customer in the orbit is simpler, since it consists of a random number of transitions to the orbit of the random durations of a single stay of the customer in the orbit. The probability distribution of this duration is given and it is exponential with parameter $\frac{\sigma}{N}$. If the discrete probability distribution of number of transitions of the customer into the orbit is found, then it is not difficult to find distribution of the total sojourn time of the tagged customer in the orbit.

Let us denote by ν the number of transitions of the tagged customer into the orbit in steady-state. We should note that obviously the random variables T , W , ν depend on N but for the simplification of notations is not shown explicitly. Applying method of the asymptotic analysis under condition of unlimited growing number of sources, we can find the probability distribution of ν .

We have the following statement

Theorem 3 *Let ν be the number of transitions of the tagged customer into the orbit, then*

$$\lim_{N \rightarrow \infty} E z^\nu = \frac{q}{1 - (1 - q)z}, \quad (19)$$

where parameter q is defined in Theorem 2 and has the form

$$q = \frac{(1 - \kappa_1)\lambda}{a}, \tag{20}$$

here

$$a = (1 - \kappa_1)\lambda + \sigma\kappa_1.$$

Proof Denote $v_{res}(t)$ the residual number of transitions of the tagged customer into the orbit, that is number of transitions into the orbit of the tagged customer from moment t till the end of its successful service.

Let us introduce the conditional generating functions

$$G_k(z, i) = \mathbb{E} \left\{ z^{v_{res}(t)} | s(t) = k, i(t) = i \right\}, \tag{21}$$

where $k = \overline{0, 3}$ is the server states, defined in the previous section. Assuming that system is operating in steady state for the conditional generating functions $G_k(z, i)$ we obtain the following system of Kolmogorov equations

$$\begin{aligned} & \left[\frac{N-i}{N}\lambda + \frac{i}{N}\sigma + \gamma_0 \right] G_0(z, i) \\ &= \frac{N-i}{N}\lambda G_1(z, i+1) + \gamma_0 G_2(z, i) + \frac{i-1}{N}\sigma G_1(z, i) + \frac{\sigma}{N} G_3(z, i), \\ & \left[\frac{N-i}{N}\lambda + \frac{i-1}{N}\sigma + \gamma_1 + \mu \right] G_1(z, i) \\ &= \frac{N-i}{N}\lambda G_0(z, i+1) + \frac{i-2}{N}\sigma G_0(z, i) + \frac{\sigma}{N} z G_0(z, i) \\ & \quad + \gamma_1 G_2(z, i) + \mu G_0(z, i-1), \\ & \left[\frac{N-i}{N}\lambda + \frac{\sigma}{N} + \gamma_2 \right] G_2(z, i) = \frac{N-i}{N}\lambda G_2(z, i+1) \\ & \quad + \frac{\sigma}{N} z G_2(z, i) + \gamma_2 G_0(z, i), \\ & \left[\frac{N-i}{N}\lambda + \frac{i-1}{N}\sigma + \gamma_1 + \mu \right] G_3(z, i) \\ &= \frac{N-i}{N}\lambda z G_0(z, i+1) + \frac{i-1}{N}\sigma z G_0(z, i) + \gamma_1 z G_2(z, i) + \mu. \end{aligned} \tag{22}$$

Similarly to the proof of the Theorem 2, denoting $\frac{1}{N} = \varepsilon$, introducing the substitutions

$$i\varepsilon = x, \quad G_k(z, i) = F_k(z, x, \varepsilon), \quad a(x) = (1-x)\lambda + \sigma x, \tag{23}$$

we obtain the following system of equations

$$\begin{aligned} & -[a(x) + \gamma_0] F_0(z, x, \varepsilon) + (1-x)\lambda F_1(z, x + \varepsilon, \varepsilon) + \gamma_0 F_2(z, x, \varepsilon) \\ & \quad + (x - \varepsilon)\sigma F_1(z, x, \varepsilon) + \varepsilon\sigma F_3(z, x, \varepsilon) = 0, \\ & -[a(x) + \gamma_1 + \mu - \varepsilon\sigma] F_1(z, x, \varepsilon) \\ & \quad + (1-x)\lambda F_0(z, x + \varepsilon, \varepsilon) + \gamma_1 F_2(z, x, \varepsilon) \\ & \quad + (x - 2\varepsilon)\sigma F_0(z, x, \varepsilon) + \varepsilon\sigma z F_0(z, x, \varepsilon) + \mu F_0(z, x - \varepsilon, \varepsilon) = 0, \\ & -[(1-x)\lambda + \gamma_2 + \varepsilon\sigma] F_2(z, x, \varepsilon) + (1-x)\lambda F_2(z, x + \varepsilon, \varepsilon) \end{aligned}$$

$$\begin{aligned}
& + \gamma_2 F_0(z, x, \varepsilon) + \varepsilon \sigma z F_2(z, x, \varepsilon) = 0, \\
& - [a(x) + \gamma_1 + \mu - \varepsilon \sigma] F_3(z, x, \varepsilon) + (1-x)\lambda z F_0(z, x, \varepsilon) \\
& + (x-\varepsilon)\sigma z F_0(z, x, \varepsilon) + \gamma_1 z F_2(z, x, \varepsilon) + \mu = 0.
\end{aligned} \tag{24}$$

Taking the limiting transition under condition $\varepsilon \rightarrow 0$ and denoting $\lim_{\varepsilon \rightarrow 0} F_k(z, x, \varepsilon) = F_k(z, x)$, system (24) can be rewritten as

$$\begin{aligned}
& - [a(x) + \gamma_0] F_0(z, x) + a(x) F_1(z, x) + \gamma_0 F_2(z, x) = 0, \\
& - [a(x) + \gamma_1 + \mu] F_1(z, x) + [a(x) + \mu] F_0(z, x) + \gamma_1 F_2(z, x) = 0, \\
& - \gamma_2 F_2(z, x) + \gamma_2 F_0(z, x) = 0, \\
& - [a(x) + \gamma_1 + \mu] F_3(z, x) + a(x) z F_0(z, x) + \gamma_1 z F_2(z, x) + \mu = 0.
\end{aligned} \tag{25}$$

From the first three equations of this system we can observe that the functions

$$F_0(z, x) = F_1(z, x) = F_2(z, x)$$

are equal. Designating by $F(z, x)$ their common value, from the fourth equation of the system (25) let us derive expression for function $F_3(z, x)$

$$F_3(z, x) = \frac{[a(x) + \gamma_1] z F(z, x) + \mu}{a(x) + \gamma_1 + \mu}. \tag{26}$$

To find the function $F(z, x)$, let us introduce the solution $F_k(z, x, \varepsilon)$ of system (24) in the form of the following decomposition

$$\begin{aligned}
F_k(z, x, \varepsilon) &= F(z, x) + \varepsilon f_k(z, x) + o(\varepsilon), \quad k = \overline{0, 2}, \\
F_3(z, x, \varepsilon) &= F_3(z, x) + \varepsilon f_3(z, x) + o(\varepsilon).
\end{aligned}$$

Substituting these decompositions into the first three equations of system (24), we obtain equalities

$$\begin{aligned}
& - [a(x) + \gamma_0] \{F(z, x) + \varepsilon f_0(z, x)\} - \varepsilon \sigma F(z, x) \\
& + (1-x)\lambda \left\{ F(z, x) + \varepsilon \frac{\partial F(z, x)}{\partial x} + \varepsilon f_1(z, x) \right\} + \varepsilon \sigma F_3(z, x) \\
& + x\sigma \{F(z, x) + \varepsilon f_1(z, x)\} + \gamma_0 \{F(z, x) + \varepsilon f_2(z, x)\} = o(\varepsilon), \\
& - [a(x) + \gamma_1 + \mu] \{F(z, x) + \varepsilon f_1(z, x)\} + \varepsilon \sigma F(z, x) \\
& + (1-x)\lambda \left\{ F(z, x) + \varepsilon \frac{\partial F(z, x)}{\partial x} + \varepsilon f_0(z, x) \right\} + \varepsilon \sigma z F(z, x) \\
& + x\sigma \{F(z, x) + \varepsilon f_0(z, x)\} + \gamma_1 \{F(z, x) + \varepsilon f_2(z, x)\} - 2\varepsilon \sigma F(z, x) \\
& + \mu \left\{ F(z, x) - \varepsilon \frac{\partial F(z, x)}{\partial x} + \varepsilon f_0(z, x) \right\} = o(\varepsilon), \\
& - [(1-x)\lambda + \gamma_2] \{F(z, x) + \varepsilon f_2(z, x)\} - \varepsilon \sigma F(z, x) \\
& + (1-x)\lambda \left\{ F(z, x) + \varepsilon \frac{\partial F(z, x)}{\partial x} + \varepsilon f_2(z, x) \right\} + \varepsilon \sigma z F(z, x) \\
& + \gamma_2 \{F(z, x) + \varepsilon f_0(z, x)\} = o(\varepsilon).
\end{aligned}$$

Equating here coefficients at identical degrees ε , for functions $f_k(z, x)$ we receive a heterogeneous system of linear algebraic equations

$$\begin{aligned}
 & - [a(x) + \gamma_0] f_0(z, x) + a(x) f_1(z, x) + \gamma_0 f_2(z, x) \\
 = & \sigma \{F(z, x) - F_3(z, x)\} - (1 - x)\lambda \frac{\partial F(z, x)}{\partial x}, \\
 & - [a(x) + \gamma_1 + \mu] f_1(z, x) + [a(x) + \mu] f_0(z, x) + \gamma_1 f_2(z, x) \\
 = & (1 - z)\sigma F(z, x) + [\mu - (1 - x)\lambda] \frac{\partial F(z, x)}{\partial x}, \\
 & - \gamma_2 f_2(z, x) + \gamma_2 f_0(z, x) = (1 - z)\sigma F(z, x) - (1 - x)\lambda \frac{\partial F(z, x)}{\partial x}. \tag{27}
 \end{aligned}$$

Multiplying the first equality by $R_0(x)$, second by $R_1(x)$, third by $R_2(x)$ and adding this products, we obtain equation

$$\begin{aligned}
 & \sigma R_0(x) \{F(z, x) - F_3(z, x)\} + (1 - z)\sigma F(z, x) \{R_1(x) + R_2(x)\} \\
 & + [\mu R_1(x) - (1 - x)\lambda] \frac{\partial F(z, x)}{\partial x} = 0, \tag{28}
 \end{aligned}$$

since the sum of the left parts of system (27) multiplied respectively by $R_k(x)$ is equal to zero. Substituting into (28) $x = \kappa_1$, denoting $R_k(x) = R_k$, $F_k(z, x) = F_k(z)$ and taking into account equality $\mu R_1 - (1 - x)\lambda = 0$ from Theorem 1, we have

$$R_0 \{F(z) - F_3(z)\} + (1 - z)F(z) \{R_1 + R_2\} = 0,$$

which, owing to equality (26), can be rewritten as

$$\{R_0 + (1 - z)(1 - R_0)\} F(z) = R_0 F_3(z) = R_0 \frac{(a + \gamma_1) z F(z)}{a + \gamma_1 + \mu} + R_0 \frac{\mu}{a + \gamma_1 + \mu}. \tag{29}$$

After performing simple transformations, we obtain

$$\left\{ 1 - z \left[1 - R_0 \frac{\mu}{a + \gamma_1 + \mu} \right] \right\} F(z) = R_0 \frac{\mu}{a + \gamma_1 + \mu}. \tag{30}$$

In the proof of Theorem 2 it was shown that

$$q = R_0 \frac{\mu}{a + \gamma_1 + \mu} = \frac{\mu R_1}{a} = \frac{(1 - \kappa_1)\lambda}{a},$$

coinciding with (20), and from (30) it is possible to obtain

$$F(z) = \frac{q}{1 - (1 - q)z}. \tag{31}$$

Applying the law of total probability to the generating functions of ν we obtain

$$E z^\nu = R_0 F_3(z) + (1 - R_0) z F(z),$$

which, taking into account (29) can be rewritten as

$$E z^\nu = \{R_0 + (1 - z)(1 - R_0)\} F(z) + (1 - R_0) z F(z) = F(z).$$

Therefore, owing to (31) we get

$$E z^\nu = F(z) = \frac{q}{1 - (1 - q)z},$$

coinciding with (19).

The theorem is proved. □

Corollary 2 The probability distribution $P\{v = n\}$, $n = \overline{0, \infty}$ of the number of transitions of the tagged customer into the orbit is geometric and has the form

$$P\{v = n\} = q(1 - q)^n, \quad n = \overline{0, \infty}. \quad (32)$$

The proof obviously follows from the form (19) of the generating function of v .

6 Sojourn/waiting time of the tagged customer in the orbit

Theorem 4 Characteristic function of the sojourn/waiting time W of the tagged customer in an orbit has the form

$$Ee^{juW} = q + (1 - q) \frac{\sigma q}{\sigma q - juN}. \quad (33)$$

Proof Since the characteristic function of the time τ , the single retrial time of the tagged customer in the orbit has the form

$$Ee^{ju\tau} = \frac{\sigma}{\sigma - juN}, \quad (34)$$

then owing to (32) and (34) by using the law of total expectations we have

$$\begin{aligned} Ee^{juW} &= \sum_{n=0}^{\infty} \left(\frac{\sigma}{\sigma - juN} \right)^n q(1 - q)^n = \frac{q}{1 - (1 - q) \frac{\sigma}{\sigma - juN}} \\ &= q \frac{\sigma - juN}{\sigma q - juN} = q + (1 - q) \frac{\sigma q}{\sigma q - juN}. \end{aligned} \quad (35)$$

The theorem is proved. \square

Corollary 3 In the case of $N \rightarrow \infty$ the limiting probability distributions of the normalized sojourn time of the customer in the system T and the normalized sojourn time of the customer in the orbit W coincide.

Proof In the characteristic function (35) let us execute replacement (10) from Theorem 2

$$u = \varepsilon i = \frac{i}{N},$$

then function (35) can be rewritten as

$$Ee^{jt \frac{W}{N}} = q + (1 - q) \frac{\sigma q}{\sigma q - jt},$$

coinciding with (8).

The corollary is proved. \square

It is quite counter intuitive since the sojourn time in the system is the sum of the waiting time in the orbit and the service time. We should mention that in prelimit situation it is the case. But in our model we deal with normalized waiting and response times, that is these times are divided by the number of sources. So the response time is the sum of the waiting time and the total service time, not the required service time since the service is many times interrupted due to collisions and failures of the server. If we divide both sides by N and N tends to infinity their limit is the same as the ratio of total service time and N tends to zero.

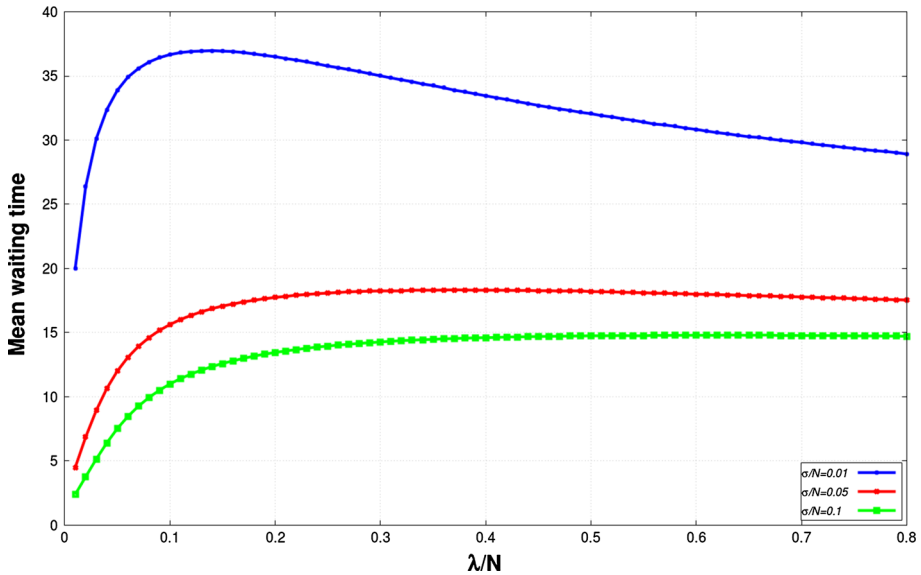


Fig. 1 Mean waiting time in the orbit without collisions by numerical and simulation methods

7 Numerical results and comparative analysis

After the theoretical results we would like to illustrate how close the asymptotic results are to the simulation results. Since for a finite N the system is rather complicated it is almost hopeless to find the exact distributions analytically or even numerically. That is why we have developed a special simulation package written in C++ to get the estimations to the involved probabilities, means and variances. It is not the aim of the present paper to introduce this package, the interested reader is referred to Tóth et al. (2017) for details.

In the following comparisons of prelimit distributions obtained with the help of stochastic simulation and by asymptotic ones are considered and several illustrative examples show the accuracy and range of applicability of the proposed asymptotic method.

Let us start our analysis with the mean waiting time of a customer in the system without collision treated in Zhang and Wang (2013) by means of a numerical method. Figure 1 shows the mean waiting time as the function the arrival rate for different values of retrial rate σ . Of course it is a prelimit situation with corresponding parameters, respectively. Using the notation of the present paper, which is different from the cited one, we have the following values

$$N = 10, \quad \mu = 1, \quad \gamma_0 = 0, \quad \gamma_1 = 0.1, \quad \gamma_2 = 1$$

Using simulation we have reproduced the results for systems without collisions obtained by a numerical procedure in Zhang and Wang (2013). In the next step by means of simulation we treat systems with collisions under the same parameter setup in the case of systems without collisions. Figure 2 illustrates how the mean waiting time changes. Of course, due to the collisions the values should be higher as it was expected. In both cases there is a maximum which a special phenomenon in finite-source retrial queues noticed by several authors, see for example Almási et al. (2005), Artalejo and Gómez-Corral (2008), Falin and Artalejo (1998), Sztrik (2005), Wang et al. (2011) and Zhang and Wang (2013).

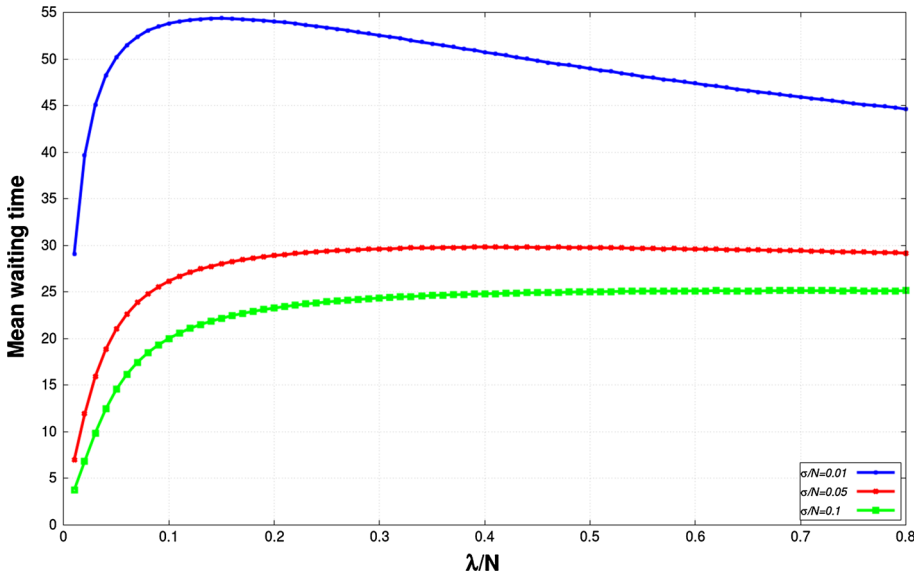


Fig. 2 Mean waiting time in the orbit with collisions by simulation

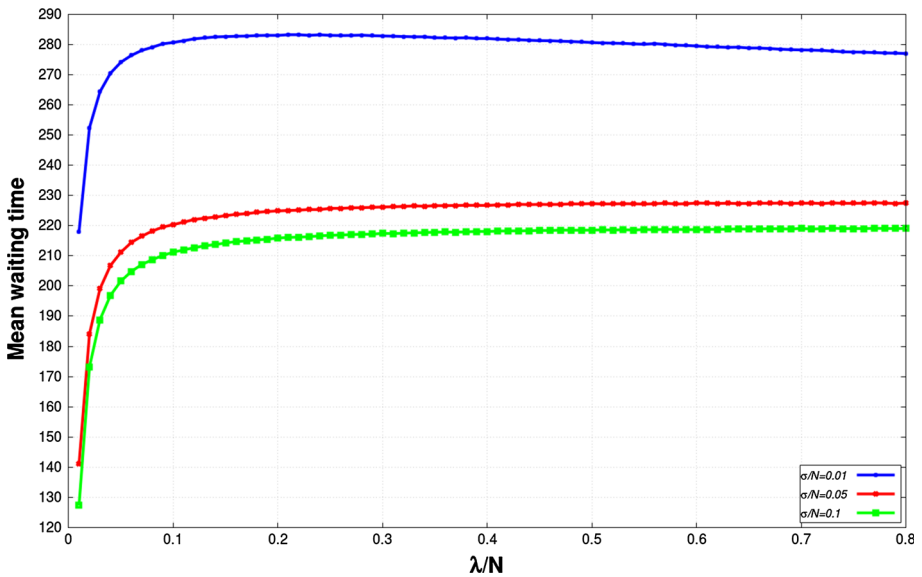


Fig. 3 Asymptotic mean waiting time in the orbit

Under the same parameter setup, except for $N = 100$ in the case of collisions the asymptotic and simulation results were coincide. This situation is illustrated in Fig. 3. We can observe that after a certain value of the arrival rate the the mean waiting time remains almost unchanged.

Table 1 The main system performance measures at various values of σ

σ	0.1	1	10	100	1000
μR_1	0.113	0.267	0.399	0.430	0.433
q	0.713	0.455	0.131	0.017	0.002
v	4.53	1.66	1.10	1.01	1.00

In the following example we consider the limiting distributions showing the effect of the retrial rates on some performance measures. If for illustrations we take

$$\lambda = 1, \quad \mu = 1, \quad \gamma_0 = 0.1, \quad \gamma_1 = 0.2, \quad \gamma_2 = 1$$

then in Table 1 for some σ the values of the following characteristics are given: μR_1 is the system throughput, q is the probability of successful service of the customer for the first attempt, v is the average value of the normalized sojourn time of the customer in the system.

Based on the results given in Table 1, it is possible to draw some conclusions, namely, with increasing intensity σ of repeated calls to the server from orbit we have noticed that

1. Throughput increases which is not natural since an increase of number of the collisions should decrease of throughput of the system. However, higher retrial rate involves shorter waiting time, so the customer could leave the system earlier.
2. The probability of successful service of the customer for the first attempt decreases as it is expected.
3. The average value of the normalized sojourn time of the customer in the system is reduced, which is associated with a significant decrease of the single sojourn time of the customer in orbit.

Previously we have obtained that the probability distribution of the number of transitions of the tagged customer into the orbit is geometric with parameter q , defined in the (20). Let us find out how close the limiting results are to the simulation results and at what values N this approximation is admissible.

Let us denote by $P_{as}(v = n)$ the asymptotic geometric distribution of probabilities with parameter q and by $P_s(v = n)$ the probability distribution of the number of transitions of the tagged customer into the orbit, obtained with the help of our simulation program. Furthermore, let us determine the accuracy (error) of approximation of distribution by mean of Kolmogorov distance Δ which for probability distributions $P_{as}(v = n)$ and $P_s(v = n)$ is defined as

$$\Delta = \max_{0 \leq i < \infty} \left| \sum_{n=0}^i (P_{as}(v = n) - P_s(v = n)) \right|.$$

Realizing the simulation program for

$$\lambda = 1, \quad \mu = 1, \quad \sigma = 4, \quad \gamma_2 = 1$$

and applying the approximation (32), we provide the Kolmogorov distance Δ for various values N and $\gamma = \gamma_0 = \gamma_1$ in Table 2.

We can see, what is expected that by increasing N the Kolmogorov distance should decrease, but with this parameter setup there is no essential reduction if $N > 50$.

Let us see the mean number of retrials under the same condition as before (Table 3).

Again we can observe what was expected, as N increases the mean number of retrials increases since there are more and more customers in the system resulting more and more

Table 2 Kolmogorov distance between distribution $P_s(i)$ and approximation of the geometric distribution $P_{as}(i)$ for various values of the parameters N and γ

	$N = 20$	$N = 30$	$N = 50$	$N = 100$	$N = 200$
$\gamma = 0.05$	0.026	0.016	0.009	0.005	0.003
$\gamma = 0.1$	0.024	0.015	0.009	0.004	0.002
$\gamma = 0.5$	0.017	0.011	0.006	0.004	0.001

Table 3 Mean number of retrials in prelimit and limiting situations for various values of parameters N and γ

	$N = 20$	$N = 30$	$N = 50$	$N = 100$	$N = 200$	limiting
$\gamma = 0.05$	5.512	5.727	5.842	5.900	5.930	5.977
$\gamma = 0.1$	6.090	6.233	6.334	6.415	6.442	6.494
$\gamma = 0.5$	10.336	10.501	10.640	10.715	10.777	10.821

collisions. At the same time it can also be seen how the mean number of retrials increases as the failure rate of the server increases. It should be underlined that each time the limiting values give very good approximations showing the effectiveness of the asymptotic method.

8 Conclusion

In this paper, the sojourn time analysis of a finite-source retrial queuing system with collision of customers where the server is subjects to random breakdowns and repairs depending on whether it is idle or busy was presented. The research has been conducted by method of asymptotic analysis under condition of unlimited growing number of sources. As a result of our investigation it was shown that the probability distribution of the number of transitions of the customer into the orbit is geometric, and the normalized sojourn time of the customer in the system has Generalized Exponential distribution. Parameters of these distributions were found. In addition, it was proved that the limiting distributions of the normalized sojourn time of the customer in the system and the normalized sojourn time of the customer in the orbit at $N \rightarrow \infty$ coincide. For this system a simulation program was developed, by the help of which the distribution of the number of transitions of the customer into the orbit was obtained. Several illustrative numerical results were discussed in details which showed the accuracy and range of applicability of the asymptotic results at finite values of N .

Acknowledgements Open access funding provided by University of Debrecen (DE). The authors are very grateful to the reviewers for their valuable comments and suggestions which improved the quality and the presentation of the paper. The work/publication of J. Sztrik is supported by the EFOP-3.6.1-16-2016-00022 project. The project is co-financed by the European Union and the European Social Fund.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alfa, A. S., & Isotupa, K. (2004). An $M/PH/k$ retrial queue with finite number of sources. *Computers & Operations Research*, 31(9), 1455–1464.
- Ali, A. A., & Wei, S. (2015). Modeling of coupled collision and congestion in finite source wireless access systems. In *2015 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 1113–1118). IEEE.
- Almási, B., Bérczes, T., Kuki, A., Sztrik, J., & Wang, J. (2016). Performance modeling of finite-source cognitive radio networks. *Acta Cybernetica*, 22(3), 617–631.
- Almási, B., Roszik, J., & Sztrik, J. (2005). Homogeneous finite-source retrial queues with server subject to breakdowns and repairs. *Mathematical and Computer Modelling*, 42(5–6), 673–682.
- Amador, J. (2010). On the distribution of the successful and blocked events in retrial queues with finite number of sources. In *Proceedings of the 5th international conference on Queueing theory and network applications* (pp. 15–22). ACM.
- Artalejo, J. (1998). Retrial queues with a finite number of sources. *Journal of the Korean Mathematical Society*, 35(3), 503–525.
- Artalejo, J. R., & Gómez-Corral, A. (2008). *Retrial queueing systems. A computational approach*. Berlin: Springer.
- Balsamo, S., Dei Rossi, G. L., & Marin, A. (2013). Modelling retrial-upon-conflict systems with product-form stochastic petri nets. In *International Conference on Analytical and Stochastic Modeling Techniques and Applications* (pp. 52–66). Springer.
- Choi, B. D., Shin, Y. W., & Ahn, W. C. (1992). Retrial queues with collision arising from unslotted CSMA/CD protocol. *Queueing Systems*, 11(4), 335–356.
- Do, T. V., Wüchner, P., Bérczes, T., Sztrik, J., & De Meer, H. (2014). A new finite-source queueing model for mobile cellular networks applying spectrum renting. *Asia-Pacific Journal of Operational Research*, 31(2), 19.
- Dragieva, V. I. (2013). A finite source retrial queue: Number of retrials. *Communications in Statistics-Theory and Methods*, 42(5), 812–829.
- Dragieva, V. I. (2014). Number of retrials in a finite source retrial queue with unreliable server. *Asia-Pacific Journal of Operational Research*, 31(2), 23.
- Dragieva, V. I. (2016). Steady state analysis of the $M/G/1/N$ queue with orbit of blocked customers. *Annals of Operations Research*, 247(1), 121–140.
- Falin, G., & Artalejo, J. (1998). A finite source retrial queue. *European Journal of Operational Research*, 108(2), 409–424.
- Falin, G., & Templeton, J. (1997). *Retrial queues*. London: Chapman & Hall.
- Falin, G. I. (1977). Waiting time in a single-channel queueing system with repeated calls. *Moscow University Computational Mathematics and Cybernetics*, 4, 66–69.
- Falin, G. I. (1984). Asymptotic investigation of fully available switching systems with high repetition intensity of blocked calls. *Moscow University Mathematics Bulletin*, 39, 72–77.
- Falin, G. I. (1986). On the waiting-time process in a single-line queue with repeated calls. *Journal of Applied Probability*, 23, 185–192.
- Falin, G. I. (1988). Virtual waiting time in systems with repeated calls. *Moscow University Mathematics Bulletin*, 43(6), 6–10.
- Gharbi, N., & Dutheillet, C. (2011). An algorithmic approach for analysis of finite-source retrial systems with unreliable servers. *Computers & Mathematics with Applications*, 62(6), 2535–2546.
- Gharbi, N., & Ioualalen, M. (2006). GSPN analysis of retrial systems with servers breakdowns and repairs. *Applied Mathematics and Computation*, 174(2), 1151–1168.
- Gómez-Corral, A. (2010). On the applicability of the number of collisions in p-persistent CSMA/CD protocols. *Computers & Operations Research*, 37(7), 1199–1211.
- Gómez-Corral, A., & Phung-Duc, T. (2016). Retrial queues and related models. *Annals of Operations Research*, 247(1), 1–2. <https://doi.org/10.1007/s10479-016-2305-2>.
- Ikhlef, L., Lekadir, O., & Aïssani, D. (2016). MRSPN analysis of Semi-Markovian finite source retrial queues. *Annals of Operations Research*, 247(1), 141–167.
- Kim, J. S. (2010). Retrial queueing system with collision and impatience. *Communications of the Korean Mathematical Society*, 25(4), 647–653.
- Kim, J., & Kim, B. (2016). A survey of retrial queueing systems. *Annals of Operations Research*, 247(1), 3–36.
- Kouvatsos, D. D. (1994). Entropy maximisation and queueing network models. *Annals of Operations Research*, 48(1), 63–126.

- Kumar, B. K., Rukmani, R., Thangaraj, V., & Krieger, U. R. (2010a). A single server retrieval queue with Bernoulli feedback and collisions. *Journal of Statistical Theory and Practice*, 4(2), 243–260.
- Kumar, B. K., Vijayalakshmi, G., Krishnamoorthy, A., & Basha, S. S. (2010b). A single server feedback retrieval queue with collisions. *Computers & Operations Research*, 37(7), 1247–1255.
- Kvach, A., & Nazarov, A. (2015). Sojourn time analysis of finite source Markov retrieval queuing system with collision. In *International Conference on Information Technologies and Mathematical Modelling* (pp. 64–72). Springer.
- Nazarov, A., & Moiseeva S. P. (2006). *Methods of asymptotic analysis in queueing theory* (in Russian). NTL Publishing House of Tomsk University.
- Nazarov, A., Sztrik, J., & Kvach, A. (2018). A survey of recent results in finite-source retrieval queues with collisions. In *Information Technologies and Mathematical Modelling. Queueing Theory and Applications* (pp. 1–15). Springer.
- Nazarov, A., Kvach, A., & Yampolsky, V. (2014). *Asymptotic analysis of closed Markov retrieval queuing system with collision* (pp. 334–341). Cham: Springer International Publishing.
- Nazarov, A., & Sudyko, E. (2010). Method of asymptotic semi-invariants for studying a mathematical model of a random access network. *Problems of Information Transmission*, 46(1), 86–102.
- Nazarov, A., Sztrik, J., Kvach, A., & Bérczes, T. (2018). Asymptotic analysis of finite-source M/M/1 retrieval queuing system with collisions and server subject to breakdowns and repairs. *Annals of Operations Research*, 277, 213–229. <https://doi.org/10.1007/s10479-018-2894-z>.
- Peng, Y., Liu, Z., & Wu, J. (2014). An M/G/1 retrieval G-queue with preemptive resume priority and collisions subject to the server breakdowns and delayed repairs. *Journal of Applied Mathematics and Computing*, 44(1–2), 187–213.
- Sudyko, E., Nazarov, A., & Sztrik, J. (2018). Asymptotic waiting time analysis of a finite-source M/M/1 retrieval queuing system. *Probability in the Engineering and Informational Sciences*, 33, 387–403.
- Sztrik, J. (2005). Tool supported performance modelling of finite-source retrieval queues with breakdowns. *Publicationes Mathematicae*, 66(1–2), 197–211.
- Sztrik, J., Almasi, B., & Roszik, J. (2006). Heterogeneous finite-source retrieval queues with server subject to breakdowns and repairs. *Journal of Mathematical Sciences New York*, 132(5), 677–685.
- Tóth, Á., Bérczes, T., Sztrik, J., & Kvach, A. (2017). Simulation of finite-source retrieval queueing systems with collisions and a non-reliable server. In *International Conference on Distributed Computer and Communication Networks* (pp. 146–158). Springer.
- Wang, J., Zhao, L., & Zhang, F. (2010). Performance analysis of the finite source retrieval queue with server breakdowns and repairs. In *Proceedings of the 5th international conference on queueing theory and network applications* (pp. 169–176). ACM.
- Wang, J., Zhao, L., & Zhang, F. (2011). Analysis of the finite source retrieval queues with server breakdowns and repairs. *Journal of Industrial and Management Optimization*, 7(3), 655–676.
- Wüchner, P., Sztrik, J., & de Meer, H. (2010). Finite-source retrieval queues with applications. In *Proceedings of 8th international conference on applied informatics*, Eger, Hungary, Vol. 2, pp. 275–285.
- Zhang, F., & Wang, J. (2013). Performance analysis of the retrieval queues with finite number of sources and service interruptions. *Journal of the Korean Statistical Society*, 42(1), 117–131.