Probability in the Engineering and Informational Sciences, page 1 of 17, 2018.
doi:10.1017/S0269964818000207

ASYMPTOTIC WAITING TIME ANALYSIS OF A FINITE-SOURCE M/M/1 RETRIAL QUEUEING SYSTEM

E. Sudyko
National Research Tomsk State University, 36 Lenina ave., 634050 Tomsk, Russia
E-mail: esudyko@yandex.ru

A.A. Nazarov
Peoples Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya St, 117198 Moscow, Russia
E-mail: nazarov.tsu@gmail.com

J. Sztrik
University of Debrecen, Debrecen, Hungary
E-mail: sztrik.janos@inf.unideb.hu

The aim of the paper is to derive the distribution of the number of retrial of the tagged request and as a consequence to present the waiting time analysis of a finite-source M/M/1 retrial queueing system by using the method of asymptotic analysis under the condition of the unlimited growing number of sources. As a result of the investigation, it is shown that the asymptotic distribution of the number of retrials of the tagged customer in the orbit is geometric with given parameter, and the waiting time of the tagged customer has a generalized exponential distribution. For the considered retrial queuing system numerical and simulation software packages are also developed. With the help of several sample examples the accuracy and range of applicability of the asymptotic results in prelimit situation are illustrated showing the effectiveness of the proposed approximation.

Keywords: accuracy and area of applicability of approximations, asymptotic analysis, closed queueing system, finite-source queueing system, limiting distribution, number of retrials, retrial queue, waiting time

1. INTRODUCTION

Retrial queues (RQ) have been widely used to model many problems arising in telephone switching systems, telecommunication networks, computer networks and computer systems, call centers, wireless communication systems, and so on. 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 Gomez-Corral [8], Gómez-Corral and Phung-Duc [16], Kim and Kim [18], and references therein.

In many practical situations, it is important to take into account the fact that the rate of generation of new primary calls decreases as the number of customers in the system increases. This can be done with the help of finite-source, or quasi-random input models. RQ with quasi-random input are a recent interest in modeling among others magnetic disk memory systems, cellular mobile networks, computer networks, and local-area networks with non-persistent CSMA/CD protocols, with a star topology, with random access protocols, and with multiple-access protocols, see, for example Alfa and Isotupa [1], Ali and Wei [2], Almási et al. [3], Do et al. [9], Dragieva [12], Ikhlef, Lekadir and Aïssani [17], Lebedev and Ponomarov [21], Wüchner, Sztrik and de Meer [31].

One of the most complicated problems in the RQ-systems is the distribution of the waiting time, the time a customer spends in the orbit. More information concerning the investigation of waiting time can be found, for example in the papers Gomez-Corral and Ramalhoto [15], Neuts [27], Nobel and Tijms [28] for RQ-system with an infinite number of sources, and Artalejo, Chakravarthy and Lopez-Herrero [6], Artalejo and Gomez-Corral [5], Artalejo and Gomez-Corral [7], Falin and Artalejo [13], for finite retrial group or sources.

In this paper, we will investigate the distribution of the number of retrials of a customer together with the distribution of the waiting time of a request in a RQ-system since they are connected to each other. Related results can be found, for example in Alfa and Isotupa [1], Dragieva [10], Dragieva [11], Falin and Artalejo [13], Gharbi and Dutheillet [14], Kvach and Nazarov [20], Nazarov, Kvach and Yampolsky [25], Nazarov and Sudyko [22], Wang, Zhao and Zhang [29], Wang, Zhao and Zhang [30], Zhang and Wang [32]. We will use the method of asymptotic analysis under limiting condition of a growing number of sources as it has been applied, for example in Kvach and Nazarov [20], Nazarov, Kvach and Yampolsky [25], Nazarov and Moiseeva [22].

Although the state space is finite in the case of prelimit situation, that is when N is finite, an exact analysis of the steady-state waiting time distribution, see Falin and Artalejo [13] and the distribution of the number of retrials, see Dragieva [10] is very complicated, for that reason, our main contribution is to propose an asymptotic analysis of these measures by sending the number of sources N to infinity and using of course some proper form of scaling or normalizing. The main advantage of the proposed method is that it directly works on the asymptotic solution to the system of balance equations without deriving explicit expressions of the exact generating and characteristic functions.

The rest of the paper is organized as follows. In Section 2 we describe the mathematical model of the RQ-system and define the main problem of obtaining probability characteristics of the waiting time. In Section 3 the Kolmogorov's equations with respect to the systems state and the residual number of retrials are derived. Section 4 is devoted to the asymptotic analysis of the probability distribution of servers states and we obtain the first order asymptotics (the law of large numbers) for the distribution of the number of customers in the orbit. In Section 5 asymptotic analysis of the generating functions for the residual number and the number of retrials is treated, respectively. Section 6 deals with the approximation of the distribution of the waiting time in the orbit in prelimit situation. In Section 7 with the help of several sample results obtained by numerical and simulation methods the accuracy and range of applicability of the asymptotic results in prelimit situation are illustrated showing the effectiveness of the proposed approximation. Finally, the paper ends with a Conclusion.

2. MATHEMATICAL MODEL

Let us consider a single server RQ-system where primary requests are generated by a finite number of N sources. If the server is idle, then the service of an incoming request starts

 https://doi.org/10.1017/S0269964818000207

immediately; after the service, this request leaves the service facility and the source starts to generate a new primary request for service. If the server is busy, then the incoming request joins to the orbit, and after a random delay, each request generates a retrial attempt in order to capture the server. If the server is idle at the moment of the retrial, then this request occupies the server; otherwise, it returns to the retrial pool thus being delayed by a random amount of time.

We suppose that each source during a random time, which is exponentially distributed with parameter λ/N, generates a request, the service time of which is exponentially distributed with parameter μ. The inter retrial times are supposed to be exponentially distributed with parameter σ/N. Assuming that all durations are independent random variables the aim of our investigation is to get the distribution of the number of retrials and the waiting time in this RQ-system. Supposing that the system is in the stationary regime and denoting the length of the waiting time of the tagged customer by W, and by R the number of retrial attempts to capture the server before beginning the service, it is easy to see that the value R is a non-negative integer; R is equal to zero when the primary request finds the server idle and the service starts immediately.

It is obvious that

$$W = \begin{cases} 0, & \text{if } R = 0 \\ \tau_1 + \tau_2 + \cdots + \tau_R, & \text{if } R > 0 \end{cases} \tag{1}$$

where τ_i denotes the inter retrial times of the tagged customer, which are supposed to be independent and exponentially distributed random variables with parameter σ/N.

Let us define two random variables for the tagged request: the residual waiting time $W_{\text{res}}(t)$ as the length of interval from the moment t until the server start servicing the tagged request and the residual number of retrials $R_{\text{res}}(t)$ as the number of attempts until the tagged request finds the server idle. Note that $W_{\text{res}}(t)$ and $R_{\text{res}}(t)$ have positive values and they have the following relation

$$W_{\text{res}}(t) = \tau_1 + \tau_2 + \cdots + \tau_{R_{\text{res}}(t)}. \tag{2}$$

It is clear that W, $W_{\text{res}}(t)$, R, $R_{\text{res}}(t)$ and the number of requests in the orbit depend on N and they are prelimit random variables. But for the simpler notation, this dependence is not shown directly, but we should remember to this fact.

In this paper, we will first deal with the asymptotic (for $N \longrightarrow \infty$) distribution of $R_{\text{res}}(t)$, R and then the approximation of the prelimit (for $N < \infty$) distribution of $W_{\text{res}}(t)$ and W. The asymptotic distributions will be obtained by the asymptotic method, whereas the prelimit distributions by numerical and simulation approaches. Then our aim is to compare these distributions and to show the effectiveness of the asymptotic method.

3. KOLMOGOROV'S EQUATIONS

This section is important for the investigations carried out in the following sections. The used asymptotic methods are similar but they are different and our aim is that the readers understand this effective approach. That is the reason why we show the detailed analysis in the given sections.

Let us derive first the steady-state Kolmogorov's equations. Let $Q(t)$ be the number of requests in the orbit at the time t, and let $C(t)$ define the state of the server at the time t

as follows:

$$C(t) = \begin{cases} 0, & \text{if the server is idle,} \\ 1, & \text{if the server is busy.} \end{cases}$$

Due to the exponential distribution of the times involved in the model construction, it is easy to see that $(C(t), Q(t))$ is a two-dimensional Markov process. Let us denote its distribution by

$$P_k(j,t) = P(C(t) = k, Q(t) = j).$$

Since the described Markov chain is homogeneous and irreducible, having a finite number of states, its stationary distribution $P_k(j) = P(C = k, Q = j)$ exists and satisfy the following balance equations

$$\begin{aligned} \lambda P_0(0) &= \mu P_1(0), \\ \left(\lambda \frac{N-1}{N} + \mu\right) P_1(0) &= \lambda P_0(0) + \sigma \frac{1}{N} P_0(1), \quad j = 0, \\ \left(\lambda \frac{N-j}{N} + \sigma \frac{j}{N}\right) P_0(j) &= \mu P_1(j), \\ \left(\lambda \frac{N-j-1}{N} + \mu\right) P_1(j) &= \lambda \frac{N-j}{N} P_0(j) \\ &\quad + \sigma \frac{j+1}{N} P_0(j+1) + \lambda \frac{N-j}{N} P_1(j-1), \quad 1 \le j \le N-2, \\ \left(\lambda \frac{1}{N} + \sigma \frac{N-1}{N}\right) P_0(N-1) &= P_1(N-1)\mu, \\ \mu P_1(N-1) &= \lambda \frac{1}{N} P_0(N-1) + \lambda \frac{1}{N} P_1(N-2), \quad j = N-1. \end{aligned} \tag{3}$$

This can be written in a shorter form as

$$\begin{aligned} \left(\lambda \frac{N-j}{N} + \sigma \frac{j}{N}\right) P_0(j) &= \mu P_1(j), \\ \left(\lambda \frac{N-j-1}{N} + \mu\right) P_1(j) &= \lambda \frac{N-j}{N} P_0(j) + \sigma \frac{j+1}{N} P_0(j+1) + \lambda \frac{N-j}{N} P_1(j-1), \end{aligned}$$

with the convention that $P_0(N) = P_1(-1) = 0$.

Let us denote the conditional generating functions for the residual number of retrials $R_{\text{res}}(t)$ of the tagged request by

$$G_k(z, j, t) = E\{z^{R_{\text{res}}(t)} \mid C(t) = k, Q(t) = j\}. \tag{4}$$

Since $(C(t), Q(t))$ is in stationary regime these generating functions do not depend on t. Thus using standard method it is easy to see that the steady-state generating functions

$G_k(z,j) = E\{z^{R_{\text{res}}} \mid C = k, Q = j\}$ satisfy the following system of equations

$$\begin{aligned}
&\lambda G_0(z,0) = \lambda G_1(z,0),\\
&\left(\lambda\frac{N-1}{N}+\mu\right)G_1(z,0) = \mu G_0(z,0) + \lambda\frac{N-1}{N}G_1(z,1), \quad j=0,\\
&\left(\lambda\frac{N-j}{N}+\sigma\frac{j}{N}\right)G_0(z,j) = \lambda\frac{N-j}{N}G_1(z,j) + \sigma\frac{j-1}{N}G_1(z,j-1) + \sigma\frac{1}{N}z,\\
&\left(\lambda\frac{N-j-1}{N}+\sigma\frac{1}{N}+\mu\right)G_1(z,j) = \mu G_0(z,j) + \lambda\frac{N-j-1}{N}G_1(z,j+1)\\
&\qquad + \sigma\frac{1}{N}zG_1(z,j), \quad 1 \le j \le N-2,\\
&\left(\lambda\frac{1}{N}+\sigma\frac{N-1}{N}\right)G_0(z,N-1) = \lambda\frac{1}{N}G_1(z,N-1) + \sigma\frac{N-2}{N}G_1(z,N-2) + \sigma\frac{1}{N}z,\\
&\left(\sigma\frac{1}{N}+\mu\right)G_1(z,N-1) = \mu G_0(z,N-1) + \sigma\frac{1}{N}zG_1(z,N-1), \quad j=N-1.
\end{aligned} \tag{5}$$

Eqs. (3) and (5) will be used for the investigation of the asymptotic distribution of the residual number retrials and then for the approximation of the distribution of the waiting time of the tagged request.

First, we consider the method of asymptotic analysis under the limiting condition $N \to \infty$, then in Section 7, we perform a numerical and simulation analysis for the prelimit situation $N < \infty$ to show the efficiency of the asymptotic method.

4. ASYMPTOTIC ANALYSIS OF THE DISTRIBUTION OF THE SERVER'S STATES AND THE NUMBER OF CUSTOMERS IN THE ORBIT

Let us introduce the following steady-state partial characteristic functions

$$H_k(u) = \sum_{j=0}^{N-1} e^{iuj} P_k(j) \tag{6}$$

where $i = \sqrt{-1}$ is the imaginary unit. It can easily be seen that

$$\sum_{j=0}^{N-1} e^{iuj} j P_k(j) = -i\frac{\partial H_k(u)}{\partial u}. \tag{7}$$

For functions $H_k(u)$ according to Eqs. (7) and (3) we can obtain the following equations

$$\begin{cases}
\dfrac{i}{N}(\sigma-\lambda)\dfrac{\partial H_0(u)}{\partial u} - \lambda H_0(u) + \mu H_1(u) = 0,\\
-\dfrac{i}{N}\left(e^{-iu}\sigma-\lambda\right)\dfrac{\partial H_0(u)}{\partial u} - \dfrac{i}{N}\lambda\left(1-e^{iu}\right)\dfrac{\partial H_1(u)}{\partial u}\\
\quad +\lambda H_0(u) - \left\{\mu + \lambda\dfrac{N-1}{N}\left(1-e^{iu}\right)\right\}H_1(u) = 0.
\end{cases}$$

Denoting $\varepsilon = 1/N$, and making the following substitutions

$$u = \varepsilon w, \quad H_k(u) = F_k(w,\varepsilon), \tag{8}$$

on the one hand, we can simplify the equations and the other hand, we can use some asymptotics for certain functions. Of course, since it is a substitution u remains fix. In probabilistic term, this substitution results in the scaled number of requests in the orbit, or the normed number of requests as we will refer to it later on. The interpretation is clear since if the number of sources increases the request in the orbit also increases.

Thus for functions $F_k(w,\varepsilon)$ we can get the following system of equations

$$\begin{cases} i(\sigma-\lambda)\dfrac{\partial F_0(w,\varepsilon)}{\partial w} - \lambda F_0(w,\varepsilon) + \mu F_1(w,\varepsilon) = 0, \\ -i\left(e^{-i\varepsilon w}\sigma - \lambda\right)\dfrac{\partial F_0(w,\varepsilon)}{\partial w} - i\lambda\left(1-e^{i\varepsilon w}\right)\dfrac{\partial F_1(w,\varepsilon)}{\partial w} \\ +\lambda F_0(w,\varepsilon) - \{\mu + \lambda(1-\varepsilon)\left(1-e^{i\varepsilon w}\right)\}F_1(w,\varepsilon) = 0. \end{cases} \tag{9}$$

Since it is very complicated to get the exact solution of this system, we will use asymptotics methods. Let us investigate this system in two stages as the authors carried in Kvach and Nazarov [20], Nazarov and Sudyko [23], Nazarov and Sudyko [24], Nazarov and Moiseeva [22].

Stage 1. Passing to the limit as $\varepsilon \to 0$ in system Eq. (9), after denoting

$$\lim_{\varepsilon\to 0} F_k(w,\varepsilon) = F_k(w),$$

we obtain the following system

$$\begin{cases} i(\sigma-\lambda)\dfrac{\partial F_0(w)}{\partial w} - \lambda F_0(w) + \mu F_1(w) = 0, \\ -i(\sigma-\lambda)\dfrac{\partial F_0(w)}{\partial w} + \lambda F_0(w) - \mu F_1(w) = 0. \end{cases} \tag{10}$$

Let us notice that this system consists of two equivalent equations.

Stage 2. Adding together equations of system Eq. (9) we get

$$i\sigma\left(1-e^{-i\varepsilon w}\right)\frac{\partial F_0(w,\varepsilon)}{\partial w} + i\lambda\left(e^{i\varepsilon w}-1\right)\frac{\partial F_1(w,\varepsilon)}{\partial w} + \lambda(1-\varepsilon)\left(e^{i\varepsilon w}-1\right)F_1(w,\varepsilon) = 0.$$

Using first order asymptotics to the exponential function this equation can be rewritten as follows

$$i\sigma\left(i\varepsilon w + o(\varepsilon)\right)\frac{\partial F_0(w,\varepsilon)}{\partial w} + i\lambda\left(i\varepsilon w + o(\varepsilon)\right)\frac{\partial F_1(w,\varepsilon)}{\partial w} + \lambda(1-\varepsilon)\left(i\varepsilon w + o(\varepsilon)\right)F_1(w,\varepsilon) = 0.$$

Passing to the limit $\varepsilon \to 0$ here, we have

$$i\sigma\frac{\partial F_0(w)}{\partial w} + i\lambda\frac{\partial F_1(w)}{\partial w} + \lambda F_1(w) = 0$$

which together with Eq. (10) we obtain the following system of equations

$$\begin{cases} i(\sigma-\lambda)\dfrac{\partial F_0(w)}{\partial w} - \lambda F_0(w) + \mu F_1(w) = 0, \\ i\sigma\dfrac{\partial F_0(w)}{\partial w} + i\lambda\dfrac{\partial F_1(w)}{\partial w} + \lambda F_1(w) = 0. \end{cases} \tag{11}$$

We give a constructive solution to this system, namely, we find the solution in the following form

$$F_k(w) = R(k)\exp\{iw\kappa\}, \tag{12}$$

where $R(k) = P(C = k), \quad k = 0, 1.$
Substituting Eq. (12) into system Eq. (11) for the probability distribution $R(k)$ we get a homogeneous system of two algebraic equations, namely

$$\begin{cases} i(\sigma-\lambda)i\kappa R(0) - \lambda R(0) + \mu R(1) = 0, \\ i\sigma i\kappa R(0) + i\lambda i\kappa R(1) + \lambda R(1) = 0. \end{cases}$$

This can be rewritten as

$$\begin{cases} -\{\lambda(1-\kappa) + \sigma\kappa\}R(0) + \mu R(1) = 0, \\ -\sigma\kappa R(0) + \lambda(1-\kappa)R(1) = 0. \end{cases} \tag{13}$$

The homogeneous system Eq. (13) has a non-trivial solution if and only if its determinant vanishes, therefore we have

$$\lambda(1-\kappa)\big[\lambda(1-\kappa)+\sigma\kappa\big] - \mu\sigma\kappa = 0, \tag{14}$$

resulting the following quadratic equation

$$q(\kappa) = \lambda(\lambda-\sigma)\kappa^2 + \{\sigma(\lambda-\mu) - 2\lambda^2\}\kappa + \lambda^2 = 0, \tag{15}$$

which determines the value of κ for Eq. (12).

Note that $q(0) = \lambda^2$ and $q(1) = -\sigma\mu$, therefore there exists at least one root in $(0, 1)$. Since $\kappa_1\kappa_2 = \lambda/(\lambda - \sigma)$ the other root cannot be in $(0, 1)$ thus $0 < \kappa < 1$ is unique for all values of parameters λ, μ and σ.

According to system Eq. (13) and by considering the normalization condition $R(0) + R(1) = 1$, the probability distribution $R(k)$ can be calculated as

$$R(0) = \frac{\mu}{\lambda(1-\kappa)+\sigma\kappa+\mu}, \quad R(1) = \frac{\lambda(1-\kappa)+\sigma\kappa}{\lambda(1-\kappa)+\sigma\kappa+\mu}. \tag{16}$$

Consequently, since the asymptotic partial characteristic functions $F_k(w)$ has the form Eq. (12) and the characteristic function is the sum of the partial characteristic functions it is not difficult to see that we get the law of large numbers in probability theory, namely the normalized number of requests in the orbit under limiting condition of growing number of sources $N \to \infty$ converges weakly to the deterministic value κ, which has been determined from Eq. (15). Actually, it is the mean of the normalized number of requests in the orbit. With the help of formulas, this statement can be stated in the following way

$$\lim_{N\to\infty} \mathsf{E}\exp\left\{iw\frac{Q}{N}\right\} = \exp\{iw\kappa\}. \tag{17}$$

Hence the mean number of requests in the orbit $E(Q)$ can be approximated by $N\kappa$.
Using characteristic functions instead of generating functions is preferable since it helps us to get the law of large numbers and in case we are interested in the distribution we can apply the inversion formula.

5. ASYMPTOTIC ANALYSIS OF THE GENERATING FUNCTIONS FOR THE RESIDUAL NUMBER OF RETRIALS

Similarly, let us solve the system of equations Eq. (5) for the conditional generating functions $G_k(z,i)$ by the method of asymptotic analysis under growing number of sources $N \to \infty$.

Denoting $\varepsilon = 1/N$ and making the following substitutions

$$\frac{j}{N} = j\varepsilon = x, \quad G_k(z,j) = S_k(z,x,\varepsilon), \tag{18}$$

system Eq. (5) can be rewritten by the help of functions $S_k(z,x,\varepsilon)$ and we get a system of two equations, namely

$$\begin{aligned} &-[\lambda(1-x)+\sigma x]S_0(z,x,\varepsilon)+\lambda(1-x)S_1(z,x,\varepsilon)\\ &\quad+\sigma(x-\varepsilon)S_1(z,x-\varepsilon,\varepsilon)+\varepsilon\sigma z=0,\\ &-[\lambda(1-x-\varepsilon)+\sigma\varepsilon+\mu]S_1(z,x,\varepsilon)+\mu S_0(z,x,\varepsilon)\\ &\quad+\lambda(1-x-\varepsilon)S_1(z,x+\varepsilon,\varepsilon)+\varepsilon\sigma z S_1(z,x,\varepsilon)=0. \end{aligned} \tag{19}$$

As before let us investigate this system in two stages.

Stage 1. Passing to the limit as $\varepsilon \to 0$ in system Eq. (19), denoting

$$\lim_{\varepsilon\to 0} S_k(z,x,\varepsilon) = S_k(z,x)$$

for functions $S_k(z,x)$ we obtain the system of equations

$$\begin{aligned} &-[\lambda(1-x)+\sigma x]S_0(z,x)+\lambda(1-x)S_1(z,x)+\sigma x S_1(z,x)=0,\\ &-[\lambda(1-x)+\mu]S_1(z,x)+\mu S_0(z,x)+\lambda(1-x)S_1(z,x)=0. \end{aligned}$$

This system consists of two equivalent equations for which we may write

$$S_0(z,x) = S_1(z,x) = S(z,x). \tag{20}$$

Stage 2. Let us write the solution $S_k(z,x,\varepsilon)$ of the system Eq. (19) in the form of a decomposition, which is first order asymptotic

$$S_k(z,x,\varepsilon) = S(z,x) + \varepsilon S_k(z,x) + o(\varepsilon) \tag{21}$$

and rewrite system Eq. (19) as follows

$$\begin{aligned} &-[\lambda(1-x)+\sigma x]S_0(z,x,\varepsilon)+[\lambda(1-x)+\sigma x]S_1(z,x,\varepsilon)\\ &\quad+\varepsilon\frac{\partial(\sigma x S_1(z,x,\varepsilon))}{\partial x}+\varepsilon\sigma z=o(\varepsilon),\\ &-[\lambda(1-x)+\mu+\varepsilon(\sigma-\lambda)]S_1(z,x,\varepsilon)+\mu S_0(z,x,\varepsilon)+\lambda(1-x)S_1(z,x,\varepsilon)\\ &\quad+\varepsilon\frac{\partial(\lambda(1-x)S_1(z,x,\varepsilon))}{\partial x}+\varepsilon\sigma z S_1(z,x,\varepsilon)=o(\varepsilon). \end{aligned}$$

Substituting decomposition Eq. (21) into this system after simple algebraic calculations we have

$$\begin{aligned} &-[\lambda(1-x)+\sigma x]S_0(z,x)+[\lambda(1-x)+\sigma x]S_1(z,x)=\frac{\partial(\sigma x S(z,x))}{\partial x}-\sigma z,\\ &-\mu S_1(x,z)-\mu S_0(x,z)=\frac{\partial(\lambda(1-x)S(z,x))}{\partial}x-(\sigma z-\sigma+\lambda)S(z,x). \end{aligned}$$

As a result for function $S(z,x)$ we obtain the following equation

$$\{[\lambda(1-x)+\sigma x]\lambda(1-x)-\mu\sigma x\}\frac{\partial S(z,x)}{\partial}x \\ +\{[\lambda(1-x)+\sigma x]\sigma(z-1)-\sigma\mu\}S(z,x)+\mu\sigma z=0. \quad \textbf{(22)}$$

Let us remember that in the previous section we have obtained the normalized number of requests in the orbit Q/N under a growing number of sources $N\to\infty \quad (\varepsilon\to 0)$ converges to a deterministic value κ, therefore substituting $j\varepsilon = x$ in Eq. (18), value x satisfies $x=\kappa$.
Since κ is the solution to equation Eq. (14) which is the coefficient of the derivative $\partial S(z,x)/\partial x$ then the coefficient is zero in Eq. (22) and $x=\kappa$.
Denoting $S(z)=S(z,\kappa)$ we can rewrite equation Eq. (22) as follows

$$\{[\lambda(1-\kappa)+\sigma\kappa](z-1)-\mu\}S(z)+\mu z=0. \quad \textbf{(23)}$$

Obviously the solution $S(z)$ of Eq. (23) is

$$S(z)=\frac{\mu z}{\mu-[\lambda(1-\kappa)+\sigma\kappa](z-1)}=z\frac{1-R(1)}{1-zR(1)}, \quad \textbf{(24)}$$

where $R(1)$ was given in Eq. (16).
Denoting $R(1)=\rho$, we may write the generating function $S(z)$ in Eq. (24) in the form

$$S(z)=\lim_{N\to\infty} Ez^{R_{\text{res}}}=z\frac{1-\rho}{1-z\rho}. \quad \textbf{(25)}$$

Then according to Eq. (25) it is easy to see that the asymptotic probability distribution of the residual number of retrials R_{res} for the tagged request is geometric

$$P(r)=\lim_{N\to\infty} P(R_{\text{res}}=r)=(1-\rho)\rho^{r-1}, \quad r=1,2,3,\ldots \quad \textbf{(26)}$$

where

$$\rho=R(1)=\frac{\lambda(1-\kappa)+\sigma\kappa}{\lambda(1-\kappa)+\sigma\kappa+\mu}. \quad \textbf{(27)}$$

Thus

$$P(R_{\text{res}}=r)\approx(1-\rho)\rho^{r-1}, \quad r=1,2,3,\ldots$$

6. DISTRIBUTION OF THE WAITING TIME IN THE ORBIT

The probability distribution $P(r)$ given in Eq. (26) allows us to obtain other steady-state performance measures of the systems such as the approximation of the residual waiting time W_{res}, asymptotic distribution of the number of retrials R and the approximation of the distribution of the waiting time W for the tagged request.

6.1. Residual waiting time

Using equality Eq. (2), let us find the Laplace transform of the residual waiting time in steady state

$$g(\alpha) = E\{e^{-\alpha W_{\text{res}}}\}. \tag{28}$$

Let us notice that even the number of retrials depends on the inter retrial times occurred so far, due to the memoryless property of the exponentially distributed inter retrial times and the geometrically distributed number of retrials which also holds the memoryless property, the conditional waiting time is Erlang distributed. Thus by the help of the law of total probability we get

$$\begin{aligned} E\{e^{-\alpha W_{\text{res}}}\} &= \sum_{r=1}^{\infty} E\{e^{-\alpha W_{\text{res}}}|R_{\text{res}} = r\}P(R_{\text{res}} = r) \\ &= \sum_{r=1}^{\infty} E\{e^{-\alpha(\tau_1+\tau_2+\cdots+\tau_r)}\}P(R_{\text{res}} = r). \end{aligned}$$

Since the inter retrial times τ_k are independent and identically distributed, denoting by $Ee^{-\alpha\tau_k} = \varphi(\alpha)$ their common Laplace transform, we may approximate function $g(\alpha)$ in the form

$$g(\alpha) \approx \sum_{r=1}^{\infty} \varphi(\alpha)^r (1-\rho)\rho^{r-1} = \varphi(\alpha)\frac{(1-\rho)}{1-\rho\varphi(\alpha)}.$$

Since τ_k are exponentially distributed with parameter σ/N thus their common Laplace transform is

$$\varphi(\alpha) = \frac{\sigma/N}{\alpha + \sigma/N},$$

therefore

$$g(\alpha) \approx \frac{\sigma/N}{\alpha+\sigma/N}\frac{(1-\rho)}{1-\rho\frac{\sigma/N}{\alpha+\sigma/N}} = \frac{(1-\rho)\sigma/N}{\alpha+\sigma/N-\rho\sigma/N} = \frac{(1-\rho)\sigma/N}{\alpha+(1-\rho)\sigma/N}.$$

It means that the distribution of the residual waiting time W_{res} can be approximated by an exponential distribution with parameter $(1-\rho)\sigma/N$.

We are confident that our method is correct. In our recently submitted paper Nazarov et al. [26], we have used the same asymptotic method directly to the asymptotic distribution of the waiting time in prelimit situation. Then the limiting distribution of the number of retrials has been obtained and consequently by using the same approach presented here the asymptotic distribution of the waiting time has been determined resulting in the same distribution.

6.2. Number of retrials

If the primary request finds the server idle, which has the probability $R(0) = 1 - R(1) = 1 - \rho$ its service starts immediately. In this case, the number of retrials R for such request is equal to zero. Otherwise, if the primary request finds the server busy with probability $R(1) = \rho$, then the tagged request joins to the orbit and in this case the number of retrials R

is equal to the residual number of retrials R_{res}, therefore by using the law of total probability we get

$$\lim_{N\to\infty} P(R=0) = 1-\rho,$$
$$\lim_{N\to\infty} P(R=r) = R(1)(1-\rho)\rho^{r-1} = (1-\rho)\rho^r, \quad r=1,2,3,\ldots$$

Consequently, the asymptotic distribution of the number of retrials can be written as

$$\lim_{N\to\infty} P\{R=r\} = (1-\rho)\rho^r, \quad r=0,1,2,3,\ldots \tag{29}$$

6.3. Waiting time

According to Eqs. (1) and (26), using the law of total probability the Laplace transform of the waiting time can be approximated by

$$g_W(\alpha) = E\{e^{-\alpha W}\} \approx P(R=0) + \sum_{r=1}^{\infty} \varphi(\alpha)^r P(R=r) = 1-\rho + \sum_{r=1}^{\infty} \varphi(\alpha)^r (1-\rho)\rho^{r-1}$$
$$= 1-\rho + \frac{\varphi(\alpha)(1-\rho)\rho}{1-\rho\varphi(\alpha)} = 1-\rho+\rho\frac{(1-\rho)\sigma/N}{\alpha+(1-\rho)\sigma/N}.$$

Thus the waiting time W is equal to zero with probability $1-\rho$, and it is exponentially distributed with parameter $(1-\rho)\sigma/N$ with probability ρ. This mixed-type distribution is called generalized exponential (GE) distribution with parameters $1-\rho$ and $(1-\rho)\sigma/N$, or other words the conditional waiting time is exponentially distributed, see Kouvatsos [19], Wüchner, Sztrik and de Meer [31]. Consequently, the mean waiting time can be approximated as

$$E(W) \approx \rho\frac{1}{(1-\rho)\sigma/N} = \frac{\rho}{1-\rho}\frac{1}{\sigma/N}. \tag{30}$$

This formula gives us an approximation for the mean waiting time in the prelimit situation. However, this expectation could be calculated with the help of the Little-formula, too. It is not difficult to see that $\mu R(1) = \lambda(1-\kappa)$ which is the average arrival rate of the requests, thus $E(W) \approx N\kappa/(\mu R(1)$. We can use these two different formulas to verify our calculations since the results should be identical.

7. NUMERICAL AND SIMULATION RESULTS AND COMPARATIVE ANALYSIS

The probability distribution $P(r)$ given in Eq. (26) has been obtained by the method of asymptotic analysis under the limiting condition $N\to\infty$.
However, it is worth investigating the range of applicability of these asymptotic results for prelimit situations for a finite value of N.

Let us denote the prelimit probability distribution by $\pi(r) = P\{R_{\mathrm{res}} = r\}$ and since it is an essential contribution to the comparison, we show how it can be obtained by numerical methods. In Dragieva [10], a similar problem was treated by a rather complicated way that is, we did not want to use that method. Instead, we propose our own way which is also standard but has not been used frequently since the investigation of a number of retrials is not so popular due to its complexity.

According to Eqs. (4) and (5), the conditional generating functions $G_k(z,j)$ can be written as

$$G_k(z,j) = E\{z^{R_{\text{res}}}|C=k,Q=j\} = \sum_{r=1}^{\infty} z^r P\{R_{\text{res}}=r|C=k,Q=j\}. \quad (31)$$

Denoting the conditional probability by

$$\Pi_k(r,j) = P\{R_{\text{res}}=r|C=k,Q=j\}$$

Equation (31) can be rewritten as follows

$$G_k(z,j) = \sum_{r=1}^{\infty} z^r \Pi_k(r,j). \quad (32)$$

It should be noted that these conditional probabilities are well defined for each $j=0,\ldots,N-1$ even for $j=0$ because it gives the conditional probability of the residual number of retrials given the orbit is empty. Since $R_{\text{res}}(t)$ denotes the residual number of retrial from time t which are positive it means that the request does not enter to the server immediately thus the orbit can be empty.

Using the law of total probability we can get

$$\pi(r) = \sum_{j=0}^{N-1} \{\Pi_0(r,j)P_0(j) + \Pi_1(r,j)P_1(j)\}, \quad (33)$$

where the unconditional probabilities $P_k(j)$ are solutions to the Kolmogorov's system of Eq. (3) and normalization condition. To obtain the conditional probabilities $\Pi_k(r,j)$, let us substitute power series Eq. (32) into system Eq. (5) and equate coefficients of the corresponding powers of z. Then for the probabilities $\Pi_k(r,j)$ we can obtain the following systems of equations

For $r=1$

$$\begin{aligned}
&-\lambda\Pi_0(1,0) + \lambda\Pi_1(1,0) = 0,\\
&-\left(\lambda\frac{N-1}{N}+\mu\right)\Pi_1(1,0) + \mu\Pi_0(1,0) + \lambda\frac{N-1}{N}\Pi_1(1,1) = 0, \quad j=0,\\
&-\left(\lambda\frac{N-j}{N}+\sigma\frac{j}{N}\right)\Pi_0(1,j) + \lambda\frac{N-j}{N}\Pi_1(1,j) + \sigma\frac{j-1}{N}\Pi_1(1,j-1) = 0,\\
&-\left(\lambda\frac{N-j-1}{N}+\sigma\frac{1}{N}+\mu\right)\Pi_1(1,j) + \mu\Pi_0(1,j)\\
&\quad+\lambda\frac{N-j-1}{N}\Pi_1(1,j+1) = 0, \quad 1\le j\le N-2,\\
&-\left(\lambda\frac{1}{N}+\sigma\frac{N-1}{N}\right)\Pi_0(1,N-1) + \lambda\frac{1}{N}\Pi_1(1,N-1) + \sigma\frac{N-2}{N}\Pi_1(1,N-2) = 0,\\
&-\left(\sigma\frac{1}{N}+\mu\right)\Pi_1(1,N-1) + \mu\Pi_0(1,N-1) = 0, \quad j=N-1.
\end{aligned} \quad (34)$$

Table 1. Comparisons for exact and asymptotic results

N	10	50	100	200	300	$N \to \infty$
$\pi(1)$	0.4228	0.3933	0.3896	0.3878	0.3872	0.3860
$\pi(1) - P(1)$	0.0368	0.0073	0.0036	0.0018	0.0012	

For $r \geq 2$

$$\begin{aligned}
&-\lambda\Pi_0(r,0) + \lambda\Pi_1(r,0) = 0,\\
&-\left(\lambda\frac{N-1}{N} + \mu\right)\Pi_1(r,0) + \mu\Pi_0(r,0) + \lambda\frac{N-1}{N}\Pi_1(r,1) = 0, \quad j = 0,\\
&-\left(\lambda\frac{N-j}{N} + \sigma\frac{j}{N}\right)\Pi_0(r,j) + \lambda\frac{N-j}{N}\Pi_1(r,j) + \sigma\frac{j-1}{N}\Pi_1(r,j-1) = 0,\\
&-\left(\lambda\frac{N-j-1}{N} + \sigma\frac{1}{N} + \mu\right)\Pi_1(r,j) + \mu\Pi_0(r,j)\\
&\quad + \lambda\frac{N-j-1}{N}\Pi_1(r,j+1) + \sigma\frac{1}{N}\Pi_1(r-1,j) = 0, \quad 1 \leq j \leq N-2,\\
&-\left(\lambda\frac{1}{N} + \sigma\frac{N-1}{N}\right)\Pi_0(r,N-1) + \lambda\frac{1}{N}\Pi_1(r,N-1)\\
&\quad + \sigma\frac{N-2}{N}\Pi_1(r,N-2) = 0,\\
&-\left(\sigma\frac{1}{N} + \mu\right)\Pi_1(r,N-1) + \mu\Pi_0(r,N-1) + \sigma\frac{1}{N}\Pi_1(r-1,N-1) = 0, \quad j = N-1.
\end{aligned} \tag{35}$$

Solving Eqs. (3) and (34) by numerical methods for given values of parameters λ, μ, σ and N, we find values of unconditional probabilities $P_k(j)$ and conditional probabilities $\Pi_k(1,j)$ for all $0 \leq j \leq N-1$. After that, substituting these values into Eq. (33), we can get the probability $\pi(1)$ as the first step for our comparisons

$$\pi(1) = \sum_{j=0}^{N-1}\{\Pi_0(1,j)P_0(j) + \Pi_1(1,j)P_1(j)\}.$$

It should be noted since the distribution $P(r)$ is geometric it is enough to calculate the first term from which the consecutive terms can be obtained multiplying them by ρ. Thus the first terms play an important role. As the main contribution of our paper is the analysis of the number of retrials we would like to show the effectiveness of the asymptotic approach in details. As an example in Table 1 the values of $\pi(1)$ are listed for different N while the parameters $\lambda = 1.2$, $\mu = 1$, $\sigma = 2$ are fixed.

The value in the last column of Table 1 has been determined from Eq. (26) for probability $P(1)$, which has been obtained by the method of asymptotic analysis, while the second row of Table 1 shows us the difference $\pi(1) - P(1)$ for several values of N. The results show what we have expected, namely by increasing number of sources the corresponding prelimit probabilities decrease and tend to the limiting value.

Assuming that the acceptable error to obtain the prelimit probability $\pi(1)$ is not more than 0.01, it may be approximated by the asymptotic probability $P(1)$ for $N \geq 50$.

TABLE 2. Comparisons for exact and asymptotic probabilities

r	1	2	3	4	5	6	7	8	9
$\pi(r)$	0.393	0.235	0.144	0.088	0.054	0.033	0.020	0.013	0.008
$P(r)$	0.386	0.237	0.146	0.089	0.055	0.034	0.021	0.013	0.008

TABLE 3. Kolmogorov distance between exact and asymptotic distribution

N	10	30	50	100	200	300
$\Delta(N)$	0.0368	0.0192	0.0073	0.0024	0.0018	0.0012

Similarly, solving systems of equations Eq. (35) by numerical methods for given values of parameters λ, μ, σ and N, we can find the values of conditional probabilities $\Pi_k(r, j)$ for all $0 \leq j \leq N - 1$ and any range of $r = 1, 2, 3, \ldots$..

Substituting these values into Eq. (33), we can calculate the prelimit probability distribution $\pi(r)$ of the number of retrials for the tagged request.

Particularly, in Table 2 for the parameters $\lambda = 1.2$, $\mu = 1$, $\sigma = 2$ and $N = 50$ values of the prelimit probabilities $\pi(r)$, and the asymptotic probabilities $P(r)$ are listed and could be compared.

The results confirm us again that under these parameters setup the obtained distributions are very close to each other and the asymptotic method is very effective.
Another method to measure the proximity of $\pi(r)$ and $P(r)$ distributions is, for example the Kolmogorov's distance defined as follows

$$\Delta = \max_{1 \leq r < \infty} \left| \sum_{j=0}^{r} P(j) - \sum_{j=0}^{r} \pi(j) \right|,$$

which means the maximum absolute difference between the distribution functions.

In Table 3 values of $\Delta(N)$ of the Kolmogorov's distance between the asymptotic distribution $P(r)$ and the prelimit distribution $\pi(r)$ are listed for various N and under parameters $\lambda = 1.2$, $\mu = 1$, $\sigma = 2$.

We can observe again what we expected, namely by increasing N the error should decrease. It is acceptable to approximate the prelimit distribution $\pi(r)$ by the asymptotic distribution $P(r)$ when the values of the Kolmogorov's distance $\Delta(N)$, which characterizes the error of approximation, are small enough. Assuming that the acceptable error is not more than 0.01, we can conclude that the approximation is reasonable and acceptable for $N \geq 50$.

To see the effectiveness of the approximation, let us consider further examples collected in Table 4 in which we can see the impact of λ when $\mu = 1$, $\sigma = 2$ are fixed. Additional performance measures, such as the mean number of retrials in prelimit and limiting case and their differences are added. Tested by the numerical results a simulation software package has been developed with the aim to investigate finite-source RQ systems with non-exponentially distributed random variables. We could estimate the distribution of the number of customers in the system, distribution of a number of retrials, mean and variance of the waiting and response times. Our results for the non-exponential case will be the topic of a future paper. Since the approximation of the mean waiting time basically depends on the approximation of the mean number of retrials these measures are not listed.

TABLE 4. Comprehensive comparisons of exact and asymptotic metrics

λ	N	10	50	100	200	300	$N \to \infty$
1,2	$\pi(1)$	0,4228	0,3933	0,3896	0,3878	0,3872	0,3861
	Difference	0,0367	0,0072	0,0035	0,0017	0,0011	
	E(Number of retrials)	1,4318	1,5588	1,5736	1,5817	1,5847	1,5906
	Difference	0,1588	0,0318	0,017	0,0089	0,0059	
	$\Delta(N)$	0,0368	0,0073	0,0036	0,0018	0,0012	
5	$\pi(1)$	0,3163	0,2966	0,2943	0,2932	0,2928	0,2918
	Difference	0,0245	0,0048	0,0025	0,0014	0,001	
	E(Number of retrials)	2,1594	2,3685	2,3949	2,4073	2,4129	2,4271
	Difference	0,2677	0,0586	0,0322	0,0198	0,0142	
	$\Delta(N)$	0,0833	0,0176	0,0087	0,0041	0,0028	
10	$\pi(1)$	0,3092	0,2853	0,2824	0,2811	0,2806	0,2796
	Difference	0,0296	0,0057	0,0028	0,0015	0,001	
	E(Number of retrials)	2,5003	2,3684	2,5368	2,5534	2,5599	2,5765
	Difference	0,0762	0,2081	0,0397	0,0231	0,0166	
	$\Delta(N)$	0,1545	0,0341	0,01687	0,0084	0,0056	
20	$\pi(1)$	0,3133	0,2817	0,2777	0,2756	0,2751	0,2736
	Difference	0,0397	0,0081	0,0041	0,002	0,0015	
	E(Number of retrials)	2,1568	2,5421	2,5953	2,6231	2,6313	2,6549
	Difference	0,4981	0,1128	0,0596	0,0318	0,0236	
	$\Delta(N)$	0,2706	0,0663	0,0337	0,0166	0,0111	
50	$\pi(1)$	0,3281	0,2851	0,2778	0,2741	0,2726	0,2702
	Difference	0,0579	0,0149	0,0076	0,0039	0,0024	
	E(Number of retrials)	1,8131	2,4898	2,5905	2,6427	2,6621	2,7009
	Difference	0,8878	0,2111	0,1104	0,0582	0,0388	
	$\Delta(N)$	0,4495	0,1568	0,0827	0,0421	0,0283	

We can observe that λ/N has the highest impact on the measures. The smaller its value the better the approximation is. The Difference stands for the difference between the prelimit and asymptotic values. These examples illustrated us the effectiveness of the proposed asymptotic method by the help of which it is much easier to approximate the prelimit distributions than by applying numerical procedures (if possible) or simulation methods. It can easily be seen because the distribution of the number of retrials in the asymptotic and prelimit case are close to each other, measured by the Kolmogorov's distance therefore the asymptotic and prelimit distribution of the waiting time are also close to each due to the law of total probability.

Finally, it should be underlined again that the exact determination of the distribution of the waiting time and the number of retrials is a much more complicated task as we could see, for example in Artalejo [4], Artalejo and Gomez-Corral [8], Dragieva [10], and Falin and Artalejo [13]. We believe that the asymptotic method will play even more important role in investigating more complicated systems when the exact distributions are almost impossible to obtain. We have some yet unpublished results concerning systems when only simulation and asymptotic results are available.

8. CONCLUSION

In this paper, an analysis of the number of retrials and the waiting time in a finite-source retrial queuing system was presented. The research has been conducted by the method of asymptotic analysis under the condition of an unlimited growing number of sources. As a result of the investigation, it was shown that the asymptotic distribution of the number of retrials of the customers in the orbit is geometric, and the distribution of the waiting time of the customers can be approximated by a GE distribution with given parameters. For the considered retrial queueing system a numerical and simulation software package has been developed with the help of which using several sample examples the accuracy and range of applicability of the asymptotic results in prelimit situation were illustrated and thus demonstrated the efficiency of the proposed approximations. In the near future, we would like to investigate these type of systems with non-exponentially distributed random variables by using asymptotic and simulation methods.

Acknowledgments

The authors are very grateful to the reviewers for their valuable comments and suggestions which improved the quality and the presentation of the paper. They especially thank the constructive remarks of Professor Ivo Adan.

Funding

The publication was financially supported by the Ministry of Education and Science of the Russian Federation (Agreement number 02.a03.21.0008) and Peoples Friendship University of Russia (RUDN University).

References

1. Alfa, A.S. & Isotupa, K. (2004). An $M/PH/k$ retrial queue with finite number of sources. *Computers & Operations Research* 31(9): 1455–1464.
2. Ali, A.-A. & Wei, S. (2015). Modeling of coupled collision and congestion in finite source wireless access systems. In: *Wireless Communications and Networking Conference (WCNC), 2015 IEEE, New Orleans, LA, USA*, pp. 1113–1118.
3. Almási, B., Bérczes, T., Kuki, A., Sztrik, J., & Wang, J. (2016). Performance modeling of finite-source cognitive radio networks. *Acta Cybernetics* 22(3): 617–631.
4. Artalejo, J. (1998). Retrial queues with a finite number of sources. *Journal of the Korean Mathematical Society* 35: 503–525.
5. Artalejo, J. & Gomez-Corral, A. (2005). Waiting time in the M/M/c queue with finite retrial group. *Bulletin of Kerala Mathematics Association* 2: 1–17.
6. Artalejo, J., Chakravarthy, S., & Lopez-Herrero, M. (2007). The busy period and the waiting time analysis of a MAP/Mic queue with finite retrial group. *Stochastic Analysis and Applications* 25: 445–469.
7. Artalejo, J.R. & Gomez-Corral, A. (2007). Waiting time analysis of the M/G/1 queue with finite retrial group. *Naval Research Logistics* 54: 524–529.
8. Artalejo, J. & Gomez-Corral, A. (2008). *Retrial queueing systems: a computational approach*. Berlin, Heidelberg, Germany: Springer.
9. 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.
10. Dragieva, V.I. (2013). A finite source retrial queue: number of retrials. *Communications in Statistics-Theory and Methods* 42(5): 812–829.
11. 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.
12. 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.
13. Falin, G. & Artalejo, J. (1998). A finite source retrial queue. *European Journal of Operational Research* 108: 409–424.
14. 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.
15. Gomez-Corral, A. & Ramalhoto, M. (2000). On the waiting time distribution and the busy period of a retrial queue with constant retrial rate. *Stochastic Modelling and Applications* 3: 37–47.

16. Gómez-Corral, A. & Phung-Duc, T. (2016). Retrial queues and related models. *Annals of Operations Research* 247(1): 1–2.
17. Ikhlef, L., Lekadir, O., & Aïssani, D. (2016). MRSPN analysis of Semi-Markovian finite source retrial queues. *Annals of Operations Research* 247(1): 141–167.
18. Kim, J. & Kim, B. (2016). A survey of retrial queueing systems. *Annals of Operations Research* 247(1): 3–36.
19. Kouvatsos, D.D. (1994). Entropy maximisation and queueing network models. *Annals of Operations Research* 48(1): 63–126.
20. Kvach, A. & Nazarov, A. (2015). *Sojourn time analysis of finite source markov retrial queuing system with collision*. Cham, Switzerland: Springer International Publishing, Cham, Ch. 1, pp. 64–72.
21. Lebedev, E. & Ponomar'ov, V. (2008). Optimization of retrial queue with finite source of calls. *Visnyk. Seriya: Fizyko-Matematychni Nauky*. Tarasa Shevchenka: Kyïvs'kyj Universytet Imeni 2008(2): 91–97.
22. Nazarov, A. & Moiseeva, S.P. (2006). *Methods of asymptotic analysis in queueing theory*. (In Russian). Tomsk, Russia: NTL Publishing House of Tomsk University.
23. 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.
24. Nazarov, A. & Sudyko, E.A. (2010). Method of asymptotic semiinvariants for studying a mathematical model of a random access network. *Problemy Peredachi Informatsii* 46(1): 94–111.
25. Nazarov, A., Kvach, A., & Yampolsky, V. (2014). *Asymptotic analysis of closed markov retrial queuing system with collision*. Cham, Switzerland: Springer International Publishing, Cham, Ch. 1, pp. 334–341.
26. Nazarov, A., Sztrik, J., Kvach, A., & Tóth, A. (2017). Asymptotic sojourn time analysis of Markov finite-source M/M/1 retrial queueing system with collisions and server subject to breakdowns and repairs. Markov Processes and Related Fields Submitted.
27. Neuts, M. (1968). The joint distribution of the virtual waiting time and the residual busy period for the M/G/1 queue. *Journal of Applied Probability* 5, 224–229.
28. Nobel, R. & Tijms, H. (2006). Waiting-time probabilities in the M/G/1 retrial queue. *Statistica Neerlandica* 60, 73–78.
29. Wang, J., Zhao, L., & Zhang, F. (2010). Performance analysis of the finite source retrial queue with server breakdowns and repairs. In: *Proceedings of the 5th International Conference on Queueing Theory and Network Applications*. Beijing, China: *ACM*, pp. 169–176.
30. Wang, J.,, Zhao, L., & Zhang, F. (2011). Analysis of the finite source retrial queues with server breakdowns and repairs. *Journal of Industrial and Management Optimization* 7(3): 655–676.
31. Wüchner, P., Sztrik, J., & de Meer, H. (2010). Finite-source retrial queues with applications. In: *Proceedings of 8th International Conference on Applied Informatics, Eger, Hungary. Vol. 2*, pp. 275–285.
32. Zhang, F. & Wang, J. (2013). Performance analysis of the retrial queues with finite number of sources and service interruptions. *Journal of the Korean Statistical Society* 42(1): 117–131.