Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2025, 33 (2) 144–156
http://journals.rudn.ru/miph

Research article

UDC 519.872
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2025-33-2-144-156

EDN: BMHYOY

Asymptotic analysis of multiserver retrial queueing system with π-defeat of negative arrivals under heavy load

Natalya P. Meloshnikova, Ekaterina A. Fedorova

National Research Tomsk State University, 36 Lenin Ave, Tomsk, 634050, Russian Federation

(received: April 30, 2025; revised: May 10, 2025; accepted: May 14, 2025)

Abstract. The paper studies a multiserver retrial queuing system with π-defeat as a mathematical model of cloud services. The arrival processes of "positive" calls are Poisson. The system has a finite number of servers and the service time for calls at the servers is exponentially distributed. When all servers are busy, calls entering the system transfer to an orbit, where they experience a random delay. After the delay, calls from the orbit attempt to access the service unit according to a multiple access policy. The system also receives a stream of negative calls. Negative calls do not require the service. An negative call "deletes" a random number of calls is the service unit. For the considered model, the Kolmogorov equations are written in the steady state. The method of asymptotic analysis under a heavy load condition is applied for deriving the stationary probability distribution of the number of calls in the orbit. The results of the numerical analysis are presented.

Key words and phrases: mathematical modelling, retrial queueing system, negative calls, asymptotic analysis, heavy load

For citation: Meloshnikova, N. P., Fedorova, E. A. Asymptotic analysis of multiserver retrial queueing system with π-defeat of negative arrivals under heavy load. *Discrete and Continuous Models and Applied Computational Science* **33** (2), 144–156. doi: 10.22363/2658-4670-2025-33-2-144-156. edn: BMHYOY (2025).

1. Introduction

Cloud technologies represent a model for providing computing resources on demand over the Internet, where infrastructure, software, and data are located on remote servers (in the "cloud") rather than on the user's local devices. Users access these resources via the Internet using various devices, such as computers, smartphones, and tablets, and pay only for the resources actually consumed, making cloud technologies a flexible, scalable, and cost-effective solution.

Cloud technologies encompass a wide range of services delivered over the internet, including Infrastructure as a Service (IaaS) from AWS, Azure, and Google Cloud; Platform as a Service (PaaS) such as Heroku and App Engine; Software as a Service (SaaS) like Microsoft 365 and Salesforce; Functions as a Service (FaaS) such as AWS Lambda; as well as Database as a Service (DBaaS) and cloud storage solutions such as Amazon S3, Azure Blob Storage, and Google Cloud Storage, providing users with flexible, scalable, and cost-effective computing resources [1].

Cloud technologies, as and development of new telecommunication networks, remains a priority for science and technology, reflected in diverse approaches ranging from analytical methods [2],the development of architectures for managing network slicing [3] and the performance modeling of mmWave networks [4], requiring the advancement of analytical tools.

Mathematical modeling is critically important for the optimization of costs, the enhancement of performance, and the assurance of reliability by predicting resource consumption, identifying bottlenecks, and planning for scaling. It is particularly important that modeling allows for the consideration of potential negative factors, such as software failures, cyberattacks, and accidents, in order to develop effective protection and redundancy strategies, guaranteeing the uninterrupted operation of the system.

In this paper, we present a mathematical model of a cloud in the form of a multi-server retrial queueing system (RQ system) with negative calls.

Retrial queueing systems are mathematical models of queueing theory widely used to analyze and optimize various telecommunications systems, mobile communication networks, and call centers [5–7]. The main feature of such models is the presence of repeated calls to the server after an unsuccessful attempt to receive the service. There is not a queue in the system, unserved calls go to an orbit (some virtual place), where they perform a random delay. There is a random access protocol for all calls in an orbit.

J. Artalejo and A. Gomez-Corral [5], G. Falin and J. Templeton [6] offered comprehensive treatments in retrial queueing systems, establishing the groundwork for analyzing queues with repeated calls. T. Phung-Duc [7] provides a survey of retrial queueing models, highlighting their theoretical developments and diverse applications.

The concept of negative calls was pioneered by E. Gelenbe [8]. He introduced negative signals that can remove calls from the system, providing a framework for modeling complex interactions in various systems. This concept was further explored in [9, 10], solidifying the mathematical foundation for this area, such models began to be called as G-networks and G-systems. Do [11] offers a valuable bibliography on G-networks, negative calls, and their applications. M. Caglayan [12] highlighted the applications of G-networks to machine learning and energy packet networks.

Later research has expanded upon these foundations, considering various aspects of G-networks and retrial queues with negative calls. P. Bocharov and V. Vishnevsky [13] discussed the development of the theory of multiplicative networks in the context of G-networks. Y. Shin [14] investigated multiserver retrial queues with both negative calls and disasters, while R. Razumchik [15] studied queueing systems with negative arrivals, a “bunker” for displaced calls, and varying service intensities. M. Matalytski and V. Naumenko [16] analyzed queueing networks with bounded waiting times for both positive and negative calls under non-stationary conditions. Further contributions to this field include research on related queuing models. Liu et al. [17] explore a multiserver two-way communication retrial queue subject to disasters and synchronous working vacations, offering insights into the impact of disruptive events and service strategies on system performance. A. Melikov [18] considers inventory queuing systems with negative arrivals. E. Lisovskaya et al. [19] investigate a resource retrial queue with two orbits and negative customers. These works demonstrate a great interest and expansion of the theoretical understanding and practical applications of G-systems with negative calls.

In this paper, we consider a multiserver retrial queuing system with π-defeat as a model for cloud services. In terms of queueing theory, data flow in cloud services is described as an arrival process in multiserver system, which defines a cloud node. The orbit is the storage location for so-called “sleeping” requests, negative arrivals define hacker attacks.

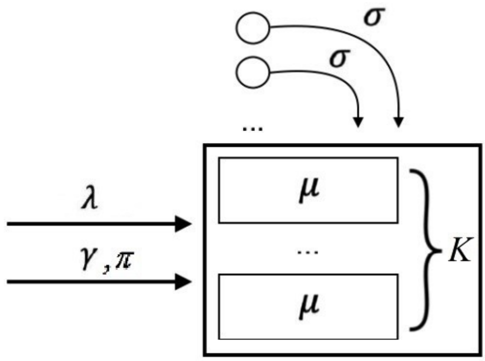

Figure 1. Multiserver retrial queueing systems system with π-defeat

The article is structured as follows. Section 2 describes a multiserver queueing system with retrials and catastrophes in the service unit, formulates the problem statement, and writes down the system of Kolmogorov equations. Section 3 is devoted to the asymptotic analysis method under heavy load conditions, which is applied to solve the system of equations. In Section 4, we demonstrate some numerical experiments that show the accuracy of the asymptotic results. Section 5 is devoted to some concluding remarks.

2. Mathematical model

In this paper, we consider a multiserver retrial queueing systems (Figure 1). The arrival process is Poisson with parameter λ, we will call these calls as "positive." Positive calls arrive into a service unit (which has K servers), until all servers become busy. The service time of each call is distributed exponentially with parameter μ. If all servers are busy, an arrival call goes to an orbit, where it performs a random delay. The delay duration has an exponential distribution with parameter σ. From the orbit, the call again turns to the service unit. If there is a free server, the call begins its service; otherwise, it returns to the orbit to make a next attempt. We suppose that there is a multiple access policy in the orbit.

In addition, negative calls arrive into the system according to Poisson arrival process with parameter γ. A negative call does not need the service; it negatively affects the system. Here, we consider a general model with negative arrivals – π-defeat. It means that a negative call deletes k servicing calls with probability π_k, where $k = \overline{1, K}$. For the probability distribution of the number of deleted calls, the normalization condition holds $\sum_{k=1}^{K} \pi_k = 1$.

Special cases of the presented model are considered in the previous studies, such as RQ with disasters in the service unit ($\pi_k = 0$ for $k = \overline{1, K-1}$ and $\pi_K = 1$) [20], the model with a single destruction ($\pi_1 = 1$ and $\pi_k = 0$ for $k = \overline{2, K}$) [21].

Denote a random process of the number of calls in the orbit by $i(t)$. The problem of obtaining the stationary probability distribution of the number of calls in the orbit is posed.

Due to $i(t)$ is not Markov process, we introduce process $k(t)$ determined states of the service unit as follows:

$$k(t) = \begin{cases} 0, \text{if all servers are free,} \\ 1, \text{if one server is busy,} \\ \dots \\ k, \text{if } k \text{ servers are busy,} \\ \dots \\ K, \text{if all servers are busy.} \end{cases}$$

Obviously, process $\{k(t), i(t)\}$ is a continuous-time Markov chain. Let $P\{k(t) = k, i(t) = i\} = P(k, i)$ be the stationary probabilities that the service unit is in state k and there are i calls in the orbit. The system of Kolmogorov equations in the steady state has the following form

$$\begin{cases} -P(0,i)(\lambda + i\sigma) + P(1,i)\mu + \sum_{k=1}^{K} P(k,i)\gamma \sum_{v=k}^{K} \pi_v = 0, \\ -P(k,i)(\lambda + k\mu + i\sigma + \gamma) + P(k-1,i)\lambda + P(k-1,i+1)\sigma(i+1) + \\ \qquad + P(k+1,i)(k+1)\mu \sum_{v=k+1}^{K} P(v,i)\gamma\pi_{v-k} = 0, \text{ with } 1 \leqslant k \leqslant K-1, \\ -P(K,i)(\lambda + K\mu + \gamma) + P(K-1,i)\lambda + P(K,i-1)\lambda + P(K-1,i+1)\sigma(i+1) = 0. \end{cases} \tag{1}$$

In addition, the normalization condition must be taken into account:

$$\sum_{k=0}^{K} \sum_{i=0}^{\infty} P(k,i) = 1.$$

Let us introduce the partial characteristic functions:

$$H_k(u) = \sum_{i=0}^{\infty} e^{jui} P(k,i).$$

Then we rewrite System (1) as follows

$$\begin{cases} -H_0(u)\lambda + j\sigma\frac{\partial H_0(u)}{\partial u} + H_1(u)\mu + \gamma \sum_{k=1}^{K} H_k(u) \sum_{v=k}^{K} \pi_v = 0, \\ -H_k(u)(\lambda + k\mu + \gamma) + j\sigma\frac{\partial H_k(u)}{\partial u} - j\sigma e^{-ju}\frac{\partial H_{k-1}(u)}{\partial u} + +H_{k-1}(u)\lambda + \\ \qquad + H_{k+1}(u)(k+1)\mu + \gamma \sum_{v=k+1}^{K} H_v(u)\pi_{v-k} = 0, \text{ with } 1 \leqslant k \leqslant K-1, \\ -H_K(u)(\lambda(1-e^{ju}) + K\mu + \gamma) + H_{K-1}(u)\lambda - j\sigma e^{-ju}\frac{\partial H_{K-1}(u)}{\partial u} = 0. \end{cases} \tag{2}$$

Summing up all equations of System (2), we obtain an additional equation

$$j\sigma \sum_{k=0}^{K-1} \frac{\partial H_k(u)}{\partial u} + H_K(u)\lambda e^{ju} = 0. \tag{3}$$

3. Asymptotic analysis

System (2)-(3) is a differential equations system of K functions, thus it is necessary to apply specific methods for its solving. We propose the method of the asymptotic analysis under a heavy load condition in this paper.

Let us denote the load parameter

$$\rho = \frac{\lambda}{K\mu}.$$

The steady state condition is written $\rho < S$, where S is an upper limit value of the load parameter, such as the system being unstable for $\rho \geq S$. Sometimes, S is called a throughput parameter. Note that an expression for S is unknown for the considered model, but further we will derive it.

The asymptotic condition of a heavy load is described as $\rho \to S$ or $\varepsilon \to 0$, where $\varepsilon = S - \rho$.

The method of the asymptotic analysis consists of several steps. First of all, we introduce the asymptotic notations:

$$\lambda = (S - \varepsilon)K\mu,\ u = \varepsilon w,\ H_k(u) = \varepsilon F_k(w, \varepsilon),\ H_K(u) = F_K(w, \varepsilon).$$

From system (2)–(3), the following system of the asymptotic equation is obtained

$$\begin{cases} -F_0(w,\varepsilon)\varepsilon(S-\varepsilon)K\mu + j\sigma\frac{\partial F_0(w,\varepsilon)}{\partial w} + F_1(w,\varepsilon)\varepsilon\mu + F_K(w,\varepsilon)\gamma\pi_K + \gamma\sum\limits_{k=1}^{K-1} F_k(w,\varepsilon)\varepsilon\sum\limits_{v=k}^{K}\pi_v = 0, \\ -F_k(w,\varepsilon)\varepsilon((S-\varepsilon)K\mu + k\mu + \gamma) + j\sigma\dfrac{\partial F_k(w,\varepsilon)}{\partial w} + F_K(w,\varepsilon)\gamma\pi_{K-k} - j\sigma e^{-j\varepsilon w}\dfrac{\partial F_{k-1}(w,\varepsilon)}{\partial w} + \\ \qquad + F_{k-1}(w,\varepsilon)\varepsilon(S-\varepsilon)K\mu + F_{k+1}(w,\varepsilon)\varepsilon(k+1)\mu + \gamma\sum\limits_{v=k+1}^{K-1} F_v(w,\varepsilon)\varepsilon\pi_{v-k} = 0,\ 1 \leqslant k \leqslant K-2, \\ -F_{K-1}(w,\varepsilon)\varepsilon((S-\varepsilon)K\mu + (K-1)\mu + \gamma) + j\sigma\dfrac{\partial F_{K-1}(w,\varepsilon)}{\partial w} - j\sigma e^{-j\varepsilon w}\dfrac{\partial F_{K-2}(w,\varepsilon)}{\partial w} + \\ \qquad + F_{K-2}(w,\varepsilon)\varepsilon(S-\varepsilon)K\mu + F_K(w,\varepsilon)K\mu + F_K(w,\varepsilon)\gamma\pi_1 = 0,\ k = K-1, \\ -F_K(w,\varepsilon)((S-\varepsilon)K\mu(1-e^{j\varepsilon w}) + K\mu + \gamma) + F_{K-1}(w,\varepsilon)(S-\varepsilon)K\mu\varepsilon - j\sigma e^{-j\varepsilon w}\frac{\partial F_{K-1}(w,\varepsilon)}{\partial w} = 0, \end{cases} \tag{4}$$

and

$$j\sigma\sum_{k=0}^{K-1}\frac{\partial F_k(w,\varepsilon)}{\partial w} + F_K(w,\varepsilon)(S-\varepsilon)K\mu e^{j\varepsilon w} = 0. \tag{5}$$

Under $\varepsilon \to 0$ in system (4)–(5), we have

$$\begin{cases} j\sigma F_0'(w) + F_K(w)\gamma\pi_K = 0, \\ j\sigma F_k'(w) - j\sigma F_{k-1}'(w) + F_K(w)\gamma\pi_{K-k} = 0,\ 1 \leqslant k \leqslant K-2, \\ j\sigma F_{K-1}'(w) - j\sigma F_{K-2}'(w) + F_K(w)K\mu + F_K(w)\gamma\pi_1 = 0, \\ -j\sigma F_{K-1}'(w) - F_K(w)(K\mu + \gamma) = 0, \end{cases} \tag{6}$$

and

$$j\sigma\sum_{k=0}^{K-1}\frac{\partial F_k(w)}{\partial w} + F_K(w)SK\mu = 0. \tag{7}$$

From system (6), it can be obtained the following expressions:

$$\begin{cases} j\sigma F_0'(w) = -F_K(w)\gamma\pi_K, \\ j\sigma F_k'(w) = -F_K(w)\gamma(\pi_K + \pi_{K-1} + \ldots + \pi_{K-k}) = 0,\ 1 \leqslant k \leqslant K-2, \\ j\sigma F_{K-1}'(w) = -F_K(w)(K\mu + \gamma) = 0. \end{cases} \tag{8}$$

From (8), let us express:

$$j\sigma \sum_{k=0}^{K-1} \frac{\partial F_k(w)}{\partial w} = -F_K(w)K\mu - F_K(w)\gamma(K\pi_K + (K-1)\pi_{K-1} + \ldots + 2\pi_2 + \pi_1). \tag{9}$$

Substituting (9) into (7), we obtain:

$$S = \frac{b\gamma}{K\mu} + 1,$$

where $b = \sum_{k=1}^{K} k\pi_k$ is a mean of the number of servicing calls deleted by negative arrivals.

In the next step of the asymptotic analysis, the following expansions for functions $F_k(w,\varepsilon)$ will be used:

$$F_k(w,\varepsilon) = F_k(w) + \varepsilon f_k(w) + O(\varepsilon^2). \tag{10}$$

Substituting (10) into system (4)–(5) and taking into account (6)–(7), we derive

$$\begin{cases} -SK\mu\varepsilon F_0(w) + j\sigma F_0'(w) + j\sigma\varepsilon f_0'(w) + \mu\varepsilon F_1(w) + \\ \quad + \gamma \sum_{k=1}^{K-1} \varepsilon F_k(w) \sum_{v=k}^{K-1} \pi_v + \gamma\pi_K F_K(w) + \gamma\varepsilon\pi_K F_K(w) = 0, \\ -(k\mu + \gamma)\varepsilon F_k(w) - SK\mu\varepsilon F_k(w) + j\sigma F_k'(w) + j\sigma\varepsilon f_k'(w) - \\ \qquad - j\sigma F_{k-1}'(w) + j\sigma(j\varepsilon w)F_{k-1}'(w) - j\sigma\varepsilon f_{k-1}' + SK\mu\varepsilon F_{k-1}(w) + \\ \quad + (k+1)\mu\varepsilon F_{k+1}(w) + \gamma \sum_{v=k+1}^{K-1} \varepsilon\pi_{v-k} F_v(w) + \gamma\varepsilon\pi_{K-k} f_K(w) + \gamma\pi_{K-k} F_K(w) = 0, k = 1 \ldots K-2, \\ -SK\mu\varepsilon F_{K-1}(w) - ((K-1)\mu + \gamma)\varepsilon F_{K-1}(w) + j\sigma F_{K-1}'(w) + \\ \qquad + j\sigma\varepsilon f_{K-1}'(w) - j\sigma F_{K-2}'(w) + j\sigma(j\varepsilon w)F_{K-2}'(w) - j\sigma\varepsilon f_{K-2}' + \\ \quad + SK\mu\varepsilon F_{K-2}(w) + K\mu F_K(w) + K\mu\varepsilon f_K(w) + \gamma\pi_1 F_K(w) + \gamma\varepsilon\pi_1 f_K(w) = 0, k = K-1, \\ -SK\mu(j\varepsilon w)F_K(w) - (K\mu + \gamma)F_K(w) - (K\mu + \gamma)\varepsilon f_K(w) + \\ \quad + SK\mu\varepsilon F_{K-1}(w) - j\sigma(1 - j\varepsilon w)F_{K-1}'(w) + j\sigma(j\varepsilon w)F_{K-1}'(w) - j\sigma\varepsilon f_{K-1}'(w) = 0, \end{cases} \tag{11}$$

and

$$j\sigma \sum_{k=0}^{K-1} \frac{\partial f_k(w)}{\partial w} + SK\mu(jw)F_K(w) - K\mu F_K(w) + SK\mu f_K(w) = 0. \tag{12}$$

Under limit $\varepsilon \to 0$, system (11)–(12) is written as

$$\begin{cases} -SK\mu F_0(w) + j\sigma f_0'(w) + \mu F_1(w) + \gamma \sum_{k=1}^{K-1} F_k(w) \sum_{v=k}^{K} \pi_v + \gamma\pi_K f_K(w) = 0, \\ -(SK\mu + k\mu + \gamma)F_k(w) + j\sigma f_k'(w) + j\sigma jwF_{k-1}'(w) - j\sigma f_{k-1}'(w) + SK\mu F_{k-1}(w) + \\ \qquad + (k+1)\mu F_{k+1}(w) + \gamma \sum_{v=k+1}^{K-1} \pi_{v-k} F_v(w) + \gamma\pi_{K-k} f_K(w) = 0, \ k = 1...K-2, \\ -(SK\mu + (K-1)\mu + \gamma)F_{K-1}(w) + j\sigma f_{K-1}'(w) + j\sigma jwF_{K-2}'(w) - \\ - j\sigma f_{K-2}'(w) + SK\mu F_{K-2}(w) + K\mu f_K(w) + \gamma\pi_1 f_K(w) = 0, \\ -jwSK\mu F_K(w) - (K\mu + \gamma)f_K(w) + SK\mu F_{K-1}(w) + j\sigma jwF_{K-1}'(w) - j\sigma f_{K-1}'(w) = 0, \end{cases} \tag{13}$$

and

$$j\sigma \sum_{k=0}^{K-1} \frac{\partial f_k(w)}{\partial w} + K\mu(Sjw - 1)F_K(w) + SK\mu f_K(w) = 0. \tag{14}$$

Let us express $j\sigma f_0'(w)$ from the first equation of System (13):

$$j\sigma f_0'(w) = SK\mu F_0(w) - \mu F_1(w) - \gamma \sum_{k=1}^{K-1} F_k(w) \sum_{v=k}^{K} \pi_v - \gamma\pi_K f_K(w).$$

From system (13), the following expressions can be derived

$$j\sigma f_k'(w) = SK\mu F_k(w) - j\sigma jw \sum_{v=0}^{k-1} F_v'(w) - ((k+1)\mu + \gamma)F_{k+1}(w) - \\ - \gamma \sum_{m=k+2}^{K-1} F_m(w) \sum_{v=m-k}^{K} \pi_v - \gamma f_K(w) \sum_{v=K-k}^{K} \pi_v \ \text{ for } k = \overline{1, K-2}, \tag{15}$$

and

$$j\sigma f_{K-1}'(w) = SK\mu F_{K-1}(w) - j\sigma jw \sum_{v=0}^{K-2} F_v'(w) - (K\mu + \gamma) f_K(w). \tag{16}$$

Let us denote $g_k = \sum_{v=k}^{K} \pi_v$. Taking into account (15)–(16), we derive

$$j\sigma \sum_{k=0}^{K-1} f_k'(w) = -\gamma \sum_{k=2}^{K-1} F_k(w)(g_k + 1) + \sum_{k=0}^{K-1} SK\mu F_k(w) - \sum_{k=1}^{K-1} k\mu F_k(w) - \\ - \gamma F_1(w) - \gamma \sum_{k=1}^{K-3} \sum_{m=k+2}^{K-1} F_m(w) g_{m-k} - j\sigma jw \sum_{k=1}^{K-1} \sum_{v=0}^{k-1} F_v'(w) - f_K(w) SK\mu.$$

The last step is substituting all derived expression into Equation (14):

$$-\gamma \sum_{k=2}^{K-1} F_k(w)(g_k + 1) + SK\mu \sum_{k=0}^{K-1} F_k(w) - \mu \sum_{k=1}^{K-1} kF_k(w) - \gamma F_1(w) - \\ - \gamma \sum_{k=1}^{K-3} \sum_{m=k+2}^{K-1} F_m(w) g_{m-k} - jw\gamma F_K(w) \sum_{v=2}^{K} (v-1) g_v + K\mu(Sjw - 1)F_K(w) = 0.$$

Substituting (8), we obtain

$$\frac{\gamma}{j\sigma}(K\mu+\gamma)F_K(w)(g_{K-1}+1)+\frac{\gamma}{j\sigma}\sum_{k=2}^{K-2}F_K(w)g_{K-k}(g_k+1)-\frac{SK\mu}{j\sigma}(K\mu+\gamma)F_K(w)-\frac{SK\mu}{j\sigma}\sum_{k=0}^{K-2}\gamma F_K(w)g_{K-k}+$$
$$+\frac{\mu(K-1)}{j\sigma}(K\mu+\gamma)F_K(w)+\frac{\mu}{j\sigma}\sum_{k=1}^{K-2}k\gamma F_K(w)g_{K-k}-\frac{\gamma}{j\sigma}\gamma F_K(w)g_{K-1}-$$
$$-\gamma\sum_{k=1}^{K-3}\Big(-\frac{(K\mu+\gamma)}{j\sigma}F_K(w)g_{K-1-k}-\frac{\gamma}{j\sigma}\sum_{m=k+2}^{K-2}F_K(w)g_{K-m}g_{m-k}\Big)+$$
$$+j\gamma F_K(w)\sum_{v=2}^{K}(v-1)g_v+jw\gamma F_K'(w)\sum_{v=2}^{K}(v-1)g_v+KS\mu jF_K(w)+K\mu(Sjw-1)F_K'(w)=0.$$

After some mathematical transformations, we obtain a differential equation for function $F_K(w)$ of the following form:

$$F_K(w)\alpha+j\sigma(1-jw\beta)F'(w)=0, \tag{17}$$

where

$$\alpha=\sigma\beta+\mu+\gamma\Big(b+\frac{(K-1)^2}{K}+2\pi_{K-1}+3\pi_K+\frac{d}{K}\Big)-\frac{\gamma}{K\mu}(f_2+b+$$
$$+\pi_{K-1}+2\pi_K+1)+\frac{\gamma^2}{K\mu}\Big(K(b-1)+3\pi_{K-1}+4\pi_K+2-C\Big),$$
$$\beta=S+\frac{\gamma}{K\mu}\sum_{k=2}^{K}(k-1)g_k,\ f_n=\sum_{k=n}^{K-2}g_{K-k}g_k, g_k=\sum_{v=k}^{K}\pi_v, \tag{18}$$
$$d=\sum_{k=0}^{K-2}kg_{K-k}, c_{k+2}=\sum_{m=k+2}^{K-2}g_{K-m}\sum_{n=m-k}^{m-1}\pi_n,\ C=\sum_{k=1}^{K-3}(f_{k+2}+c_{k+2}).$$

The solution to the equation (17) is:

$$F_K(w)=\Big(1-jw\beta\Big)^{-\frac{\alpha}{\sigma}}.$$

Turning up the asymptotic notation, the characteristic function of the number of calls in the orbit

$$H(u)=F_K\Big(\frac{u}{\varepsilon}\Big)+O(\varepsilon).$$

In this way, we finally conclude that the asymptotic characteristic function of the number of calls in the orbit of the considered model under a heavy load condition

$$h(u)=F_K\Big(\frac{u}{S-\rho}\Big)=\Big(1-\frac{ju\beta}{S-\rho}\Big)^{-\frac{\alpha}{\sigma}}, \tag{19}$$

has the form of the characteristic function of the gamma distribution with parameters α and β defined by (18).

3.1. Special cases

As mentioned above, the studied retrial queueing system with π-defeat generalizes models considered previously in [20, 21]. So we can compare the asymptotic results in two special cases:

1. Model with single destruction: $\pi_1=1$ and $\pi_k=0$ for $k=\overline{2,K}$.
2. Model with disasters in the service unit: $\pi_K=1$ and $\pi_k=0$ for $k=\overline{1,K-1}$.

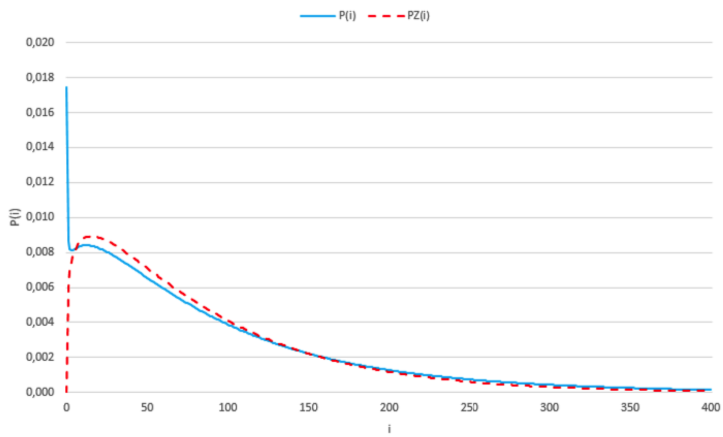

Figure 2. Exact and asymptotic probability distributions for $\rho = 0.99$

By substituting the corresponding values of π_k into (19) the following corollaries can be formulated (which coincide with [20] and [21]).

Corollary 1. The asymptotic characteristic function of the probability distribution of the number of calls in orbit in RQ M/M/K with single destruction of negative calls has the form of the gamma distribution function (19) with parameters $\alpha = \mu + \gamma + \sigma,\ \beta = 1,\ S = 1 + \dfrac{\gamma}{K\mu}$.

Corollary 2.It can be derived from (19) that the asymptotic characteristic function of the probability distribution of the number of calls in orbit in RQ M/M/K with disasters in the service unit has the form of the gamma distribution function (19) with parameters $\beta = 1 + \frac{\gamma(K+1)}{2\mu},\ S = 1 + \frac{\gamma}{\mu}$,

$$\alpha = \sigma\beta + \mu + \gamma(2K-1) - \frac{\gamma(K-1)(2+\gamma(K-2))}{2K} + \frac{\gamma^2(K^2+K-4)}{2K\mu}.$$

Note that we have proved that the asymptotic distribution under a heavy load has a gamma form for all cases.

4. Numerical analysis

To analyze the range of applicability of the proposed asymptotic method, we numerically compare asymptotic distribution $PZ(i)$ and exact distribution $P(i)$ obtained using a numerical algorithm for various values of the system parameters.

As a measure of the asymptotic method accuracy, we use the Kolmogorov distance:

$$\triangle = \left| \sum_{n=0}^{i} PZ(n) - P(n)) \right|.$$

As an example, we present the calculation for the following values of system parameters:

$$\rho = 0.99 \cdot S,\ \mu = 1,\ \gamma = 0.01,\ \sigma = 5,\ K = 5,\ \pi = [0.247\ 0.603\ 0.101\ 0.045\ 0.004]$$

From Figure 4, we can see that the exact distribution has the value of the zero state probability ($P\{i(t) = 0\}$) much greater than others $P(i)$. This is the main feature of the model for quite large values of γ (negative calls arrive more frequently and more calls are deleted).

Table 1

Kolmogorov distances for various values of the parameter ρ

	$\gamma = 0.01$	$\gamma = 0.001$	$\gamma = 0.0001$
$\rho = 0.99 \cdot S$	0.120	0.026	0.016
$\rho = 0.98 \cdot S$	0.120	0.041	0.034
$\rho = 0.97 \cdot S$	0.06	0.053	0.052

In this way the approximation has a quite big error in the point $i(t) = 0$ ($\Delta = 0.120$, Table 1), while if we analyze the entire remaining range of $i(t) > 0$ (excluding the zero point), the Kolmogorov distance does not $\sum_{i \neq 0} \Delta < 0.01$. Unfortunately, it is not possible to analytically estimate the probability P_0 from System 1 or by the asymptotic analysis.

The results of the comparison of distributions for different values of are presented in Table 1 and Figure 4.

From Table 1, it can be concluded that the accuracy of the approximation increases as the system load increases and a negative arrival rate decreases.

5. Conclusion

In this study, a multiserver RQ-system with π-defeat is considered as a mathematical model of cloud services. The asymptotic analysis method under a heavy load condition is applied. It is proven that the asymptotic characteristic function of the distribution of the number of calls in the orbit has the form of the gamma distribution function with the obtained parameters. A formula for the system throughput is obtained. Numerical analysis is presented, demonstrating the accuracy of the approximation.

Author Contributions: Conceptualization, N. Meloshnikova and E. Fedorova; methodology E. Fedorova; software, N. Meloshnikova; validation, N. Meloshnikova; formal analysis, N. Meloshnikova and E. Fedorova; investigation, N. Meloshnikova and E. Fedorova; writing—original draft preparation, N. Meloshnikova; writing—review and editing, E. Fedorova. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kavis, M. J. *Architecting the Cloud: Design Decisions for Cloud Computing Service Models* 224 pp. (Wiley; 1st edition, 2014).
2. Boccardi, F., Heath, R. W., Lozano, A., Marzetta, T. L. & Popovski, P. Five disruptive technology directions for 5G. *IEEE Communications Magazine* **52,** 74–80 (May 2014).
3. Ferrus, R., Sallent, O., Pérez-Romero, J. & Agustí, R. Management of Network Slicing in 5G Radio Access Networks. *Functional Framework and Information Models* (2018).

4. Wu, J., Wang, M., Chan, Y. C., Wong, E. W. M. & Kim, T. Performance Evaluation of 5G mmWave Networks with Physical-Layer and Capacity-Limited Blockingm. *IEEE 21st International Conference on High Performance Switching and Routing (HPSR),* 1–6 (May 2020).
5. Artalejo, J. R. & Gomez-Corral, A. *Retrial Queueing Systems* 318 pp. (Springer Berlin, 2008).
6. Falin, G. & Templeton, J. *Retrial Queues* 320 pp. (Taylor & Francis, 1997).
7. Phung-Duc, T. Retrial Queueing Models: A Survey on Theory and Applications. *Stochastic Operations Research in Business and Industry,* 1–26 (May 2017).
8. Gelenbe, E. Product-form queueing networks with negative and positive customers. *Journal of Applied Probability* **7,** 656–663 (1991).
9. Gelenbe, E., Glynn, P. & Sigman, K. Queues with Negative Arrivals. *J. Appl. Probab* **28,** 245–250 (1991).
10. Gelenbe, E. G-networks with signals and batch removal. *Probability in the Engineering and Informational Sciences Vol. 7,* 335–343 (1993).
11. Do, T. Bibliography on G-networks, negative customers and applications. *Mathematical and Computer Modelling,* 205–212 (2011).
12. Caglayan, M. G-networks and their applications to machine learning, energy packet networks and routing: introduction to the special issue. *Probability in the Engineering and Informational Sciences* **31,** 381–395 (2017).
13. Bocharov, P. & Vishnevsky, V. G-networks: development of the theory of multiplicative networks. Russian. *Automation and Telemechanics Vol. 5,* 70–74 (2003).
14. Shin, Y. Multi-server retrial queue with negative customers and disasters. *Queueing Syst,* 223–337 (2007).
15. Pechinkin, A. & Razumchik, R. A Method for Calculating Stationary Queue Distribution in a Queuing System with Flows of Ordinary and Negative Claims and a Bunker for Superseded Claims. *J. Commun. Technol. Electron.* **57,** 882–891 (2012).
16. Matalytski, M. & Naumenko, V. Analysis of the queueing network with a random bounded waiting time of positive and negative customers at a non-stationary regime. *Journal of Applied Mathematics and Computational Mechanics* **16,** 97–108 (2017).
17. Liu, T., Hsu, H. & Chang, F. Multi-Server Two-Way Communication Retrial Queue Subject to Disaster and Synchronous Working Vacation. *Algorithms* **18,** 24 (May 2025).
18. Melikov, A., Poladova, L., Edayapurath, S. & Sztrik, J. Single-Server Queuing-Inventory Systems with Negative Customers and Catastrophes in the Warehouse. *Mathematics* **11** (2023).
19. Lisovskaya, E., Fedorova, E., Salimzyanov, R. & Moiseeva, S. Resource Retrial Queue with Two Orbits and Negative Customers. *Mathematics* **10,** 321 (May 2022).
20. Meloshnikova, N. & Fedorova, E. *Asymptotic Analysis of a multiserver RQ-system with disasters in a service block under heavy load* Russian. in *Information and telecommunication technologies and mathematical modeling of high-tech systems: Proceedings of the All-Russian conference with international participation* (RUDN, Moscow, Russian Federation, April 7–11, 2025), (in print).
21. Meloshnikova, N. *Study Retrial Queue M|M|N with negative calls under heavy load* Russian. in *Proceedings of Tomsk State University. Physics and Mathematics Series: Mathematical and Software Support of Information, Technical, and Economic Systems: Materials of the IX International Scientific Conference.* **307** (Tomsk State University Publishing House, Tomsk, Russian Federation, May 26–28, 2021), 165–168.

Information about the authors

Meloshnikova, Natalya P.—PhD-student, Junior researcher of Laboratory of queueing theory and teletraffic theory of National Research Tomsk State University (e-mail: meloshnikovana@gmail.com, phone: +7 (913) 213-83-71, ORCID: 0000-0002-8708-124X, ResearcherID: MTF-1866-2025, Scopus Author ID: 58304893200)

Fedorova, Ekaterina A.—PhD in Physical and Mathematical Sciences, Associate Professor of Department of Probability Theory and Mathematical Statistic of National Research Tomsk State University (e-mail: ekat_fedorova@mail.ru, phone: +7 (952) 801-93-94, ORCID: 0000-0001-8933-5322, ResearcherID: E-3161-2017, Scopus Author ID: 56439120600)

УДК 519.872
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2025-33-2-144-156 EDN: BMHYOY

Асимптотический анализ многолинейной RQ-системы с π-поражением в условии большой загрузки

Н. П. Мелошникова, Е. А. Фёдорова

Национальный исследовательский Томский государственный университет, пр. Ленина, д. 36, Томск, 634050, Российская Федерация

Аннотация. В работе исследуется многолинейная RQ-система с π-поражением как математическая модель облачных сервисов. На вход системы поступает простейший поток «положительных» заявок. В системе конечное число обслуживающих приборов, время обслуживания заявок на приборах распределено по экспоненциальному закону. Когда все приборы заняты, заявки поступающие в систему переходят на орбиту, где осуществляют случайную задержку. После осуществления задержки, заявки с орбиты обращаются к блоку обслуживания согласно политике множественного доступа. Также в систему поступает поток так называемых «отрицательных» заявок. Отрицательная заявка не нуждается в обслуживании: при поступлении она удаляет случайное число обслуживаемых заявок. Для рассматриваемой модели записаны уравнения Колмогорова в стационарном режиме. Предлагается метод асимптотического анализа в условии большой загрузки для нахождения стационарного распределения вероятностей числа заявок на орбите. Представлены результаты численного анализа.

Ключевые слова: математическое моделирование, система массового обслуживания с повторными вызовами, отрицательные заявки, асимптотический анализ, большая загрузка