Methodology and Computing in Applied Probability (2022) 24:1753–1772
https://doi.org/10.1007/s11009-021-09890-6

Using Infinite-server Resource Queue with Splitting of Requests for Modeling Two-channel Data Transmission

Tatyana Bushkova[1] · Svetlana Moiseeva[1] · Alexander Moiseev[1] · János Sztrik[2] · Ekaterina Lisovskaya[1] · Ekaterina Pankratova[3]

Received: 3 September 2020 / Revised: 1 August 2021 / Accepted: 16 August 2021 / Published online: 9 September 2021

Abstract
New Radio Access Technology 3GPP New Radio has become the fundamental wireless technology in the fifth-generation networks, which allows us to achieve high data rates due to the ability to work in the millimeter-wave band. But the key feature and the main problem of 5G New Radio networks is that people themselves, cars, buildings, etc. are signal blockers, while the base stations of the fourth generation networks have widescreen broadcasting and such small obstacles do not cause loss of connection. Service providers and mobile operators are already testing the proposed technology. In this connection, the scientific community has the task of analyzing the performance of these systems and increasing it in the future. Currently, there are known studies of “basic” mathematical models of such networks. By this term, we mean models built in the simplest possible assumptions. However, due to the justified necessity of introducing new technology into the daily lives of subscribers, service providers pose the scientific community with the task of analyzing the effectiveness of the most appropriate mathematical models. For example, a technology of splitting transmitted data into two streams using as 5G and both 4G transmission technologies is considered now by 3GPP Project Coordination Group. The paper is devoted to such a problem. We consider a mathematical model of the message transmitting with the implementation of the splitting function in the communication networks of New Radio technology in the form of a resource queueing system with a renewal arrival process and non-exponential service. For this problem, an approximation of a stationary two-dimensional probability distribution of the number of occupied resources in parallel service units is obtained. It is shown that this approximation coincides with the Gaussian distribution, and its area of applicability is shown.

Keywords Infinite-server system · Resource queue · Splitting of requests · Parallel servicing · Renewal process · Asymptotic analysis

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

Extended author information available on the last page of the article

1 Introduction

The interest in the study of resource queueing systems is determined by the possibility of their application in the modeling of modern technical devices, next-generation data transmission networks, and computer systems, including cloud computing systems. The theory of stochastic processes, queueing theory, and the theory of teletraffic are served as the basis for the development of both theoretical and applied research in the field of data transmission and processing in information and communication systems. Various mathematical models and methods can be used to analyze the performance indicators of such systems. This includes, for example, queueing systems with repeated calls (retrial queues) Artalejoa and Gomez-Corral (2008); Falin and Templeton (1997), resource queueing systems Naumov et al. (2016); Samouylov et al. (2016); Tikhonenko (2010) and models of stochastic geometry Benkhelifa et al. (2020); Chu et al. (2020).

The growing popularity of wireless network services necessitates the creation of new approaches to assessing the quality of services of communication operators Buturlin et al. (2012); Galinina et al. (2014). In these networks, each active connection requires a certain amount of radio resources (call, message, video) provided to the request at the time of its receipt and should be released at the end of the connection. The amount of required resources is determined by a predefined probability distribution, which can take into account the features of various radio resource allocation schemes when analyzing the performance of wireless networks Galinina et al. (2014); Gudkova et al. (2014). Such systems can be modeled using resource queueing systems Lisovskaya et al. (2019); Naumov et al. (2016); Sopin et al. (2017); Tikhonenko (2010), where each arrival, in addition to one server, takes a certain amount of resource (deterministic or random, discrete or continuous). While the customers are being serviced, it occupies both the server and the resources, at the end of the service, the customers leave the system and frees up the occupied server and resources. Modeling of wireless communication networks using resource queues is represented by a large number of publications. However, most of the results were obtained under simplifying assumptions: deterministic resource requests, exponentially distributed service time, Poisson arrival process, simple configuration of a queueing system, which is associated with the complexity of constructing the corresponding stochastic processes (see references Basharin et al. (2009); Lisovskaya et al. (2019); Naumov and Samouylov (2017) and reviews inside them).

In addition, resource queues allow us to model any features of the distribution of resources in modern wireless networks. The inclusion of signals that trigger the redistribution of resources makes it possible to take into account the mobility of user devices Ageev et al. (2018); Samouylov et al. (2018); Sopin et al. (2017), as well as to model Network Slicing Guan et al. (2018); Song et al. (2019). Multicast technology offers a possible solution to the problem of transferring the same data to a number of devices, which leads to a significant improvement in the spectral efficiency and throughput of a wireless network. Since the same frequency band is used for several devices in the multicast mode, the data transfer speed can reach higher values in comparison to the unicast mode, where for each device only a small separate frequency band is allocated Araniti et al. (2015). Simultaneous servicing at several stations is modeled using resource queues with parallel servicing, and by using multi-stage (tandem) resource queues for sequential servicing at several stations Lisovskaya et al. (2017); Galileyskaya et al. (2019). Also, there is a possibility to take into account the heterogeneity of requests Lisovskaya et al. (2019) and their requirements for various resources Lisovskaya et al. (2019).

According to Cisco (2018) it is expected that by 2022 the number of devices which will connect to the wireless access network will reach 28 billion, this means about 3.5 devices per person as average. According to the same data, it is expected that monthly traffic on all devices will increase 2-5 times depending on the type of device: Smartphone, Tablet, Laptop or PC, Ultra High Definition TV. Moreover, the main part of the traffic will be watching video content, and a 12-fold increase will affect VR/AR traffic. It is worth noting that for such traffic, the amount of allocated radio resource does not affect the duration of the sessions, but affects the quality of the video image (4K, Immersive 360 video, VR, High Dynamic Range, High Frame Rate, 8K).

The 3GPP Project Coordination Group is considering now simultaneously supporting active links between multiple Radio Access Technologies to ensure session continuity in fifth-generation wireless networks. This is due to data transmission over the 5G network is very sensitive to interference on the line of sight, and the movement of the receiver can lead to additional problems in data transmission. One of the solutions may be a splitting of transmitted data into two streams one of which uses 5G New Radio and another uses 4G Long-Term Evolution technology.

Due to the necessity of introducing new technology into the daily life of subscribers, service providers pose the scientific community with the task of analyzing the effectiveness of the most appropriate mathematical models. A detailed review of modern research in the field of communications, modern computer networks, and information systems is reflected in Dudin et al. (2020).

This paper proposes a mathematical model of data transmitting with a splitting transmitted data into two independent streams. We represent the model in the form of a resource queueing system with the splitting of requests for their parallel servicing in two separate units, each of which contains an unlimited number of servers. Arrivals are modeled as a renewal process with a given distribution of interarrival time. The service discipline is as follows: when a request arrives, it is split into two parts and they are processed in parallel units of servers (each part in a single unit). The parts of the request capture (allocate) some amount of resources in their servicing units. This amount we model as random variables given by distribution functions that may be different for different servicing units. Service durations of the parts are stochastically independent of each other and of the amount of allocated resources, and they are determined by given distributions with finite first and second moments. After the service is completed for any part, it leaves the system immediately independently of another part.

The paper studies the two-dimensional stochastic process of the total amount of resources captured in each unit of servers. The complexity of the study of resource systems is due to the fact that at present there is no universal approach to solve the problems of this type. We use asymptotic methods Moiseev and Nazarov (2014, 2016) for the study which give asymptotic expressions acceptable for practical use for the desired characteristics of the system in cases where an exact analysis is impossible. In the paper, we use an asymptotic condition of equivalent growth of service times, the practical meaning of which is that the average service speed is much less than the intensity of requests Bushkova et al. (2019). The paper continues studies started in Bushkova et al. (2019) where a similar model with Poisson arrivals was considered.

Briefly about the content of the paper. In Sect. 2, we describe a mathematical model of the system under study, introduce main notations, and describe the problem under study. Section 3 contains a description of the dynamic screening method adapted to the analysis of resource queues with two units of servers that are used to solve the problem. The balance equation is formulated in this section too. In Sect. 4, we apply an asymptotic analysis procedure to solve the balance equation. As a result, we obtain Gaussian approximation for the stochastic process under study (subsection 4.3). Numerical analysis and example are presented in Sect. 5. Using numerical experiments and simulation, we estimate the applicability area of the obtained approximation. Then, using technique Moiseev and Nazarov (2016), we obtain an estimation of the amount of resources that is enough that the probability of losses is equal to or less than a given value for a similar system with a limited reserve of resources.

2 Mathematical Model

Consider a queueing system with two service units each of which contains an unlimited number of homogeneous service devices (Fig. 1). These units correspond to different transmission environments (channels or technologies).

Arrivals are determined as a renewal process, interarrival periods have cumulative distribution function $A(z)$. Arrived request is split into two parts: the first part goes to the first service unit and the second one goes to the second service unit. Each part occupies any free server in their unit, where they are serviced for a random times with probability distribution functions $B_k(x)$ (we denote the number of parts and respective service unit by $k = 1, 2$). Service times correspond to the part's transmission duration.

When entering into a service unit, a part of the request occupies some resources of the amount $v_k > 0$, $k = 1, 2$ distributed with the probability distribution function $G_k(y)$, $k = 1, 2$. Occupied resources correspond to channel resources captured by the transmission session (e.g., radio frequencies, bandwidths, etc.) After the service of a part of the request is completed, this part leaves the system and free resource occupied by it independently from another part of the request.

Suppose, functions $A(z)$, $B_k(x)$ and $G_k(y), k = 1, 2$ have finite first and second order moments, and the function $A(z)$ satisfies the condition $o(A(\Delta t)) = o(\Delta t)$.

In the paper, we consider the problem of analyzing the total amount of occupied resources in both channels. Problem of parts synchronization for a single request or analyzing and optimizing total transmission time are out of scope. We use an infinite-server model with unlimited resources assuming the channels have great enough bandwidth. Due to the resulting distributions will be Gaussian (see Sect. 4.3), we can use the hyper-ellipsoid method presented in Moiseev and Nazarov (2016) to estimate the number of servers and reserve amount of resources that will be enough for the systems with finite limitations.

Let us denote the total amount of resources occupied in the k -th service unit at time instant t by $V_k(t)$. The goal of the study is to obtain the steady-state probability distribution of

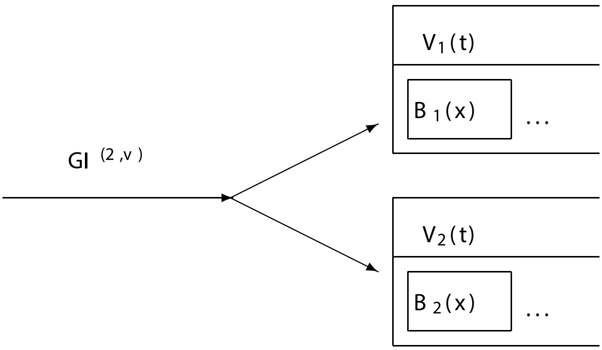

Fig. 1 Resource queueing system with splitting of requests

two-dimensional stochastic process $\{V_1(t), V_2(t)\}$. This process is not Markovian and classic methods of study do not allow to solve the problem. Due to this, we use the dynamic screening method Moiseev and Nazarov (2016) adapted for resource systems Lisovskaya et al. (2016) combined with the splitting of requests.

3 Dynamic Screening Method and Balance Equation

The dynamic screening method allows us to transform the original problem of studying of a queueing system by the analysis of some stochastic (screened) process which is constructed in the following way. Let us draw three parallel time axes. We number them from 0 to 2 (Fig. 2). Axes with numbers 1 and 2 will correspond to the service units with the same numbers, and the axis with number 0 is used to display arrival epochs.

Let some moment of time T in the future be fixed. We denote the probability that a part of a request arrived in the system at the time instant $t \leq T$ will be served at the time instant T by $S_k(t)$, where $k = 1, 2$ is the number of part of the request (or the number of service unit where it is served). These probabilities we call as screening probabilities for the 1st or for the 2nd axis. We will mark epochs on these axes if original arrival was screened on them with probabilities $S_1(t)$ and $S_2(t)$ respectively. A sequence of these epochs on each axis we call as screened flow or process. Obviously, that original arrival do not generate an epoch in k-th screened flow with probability $1 - S_k(t)$, $k = 1, 2$. This case corresponds to the situation when a part leaves the k-th unit of servers before the time instant T. Generally, the $S_k(t)$ functions also depend on the fixed moment T, i.e. $S_k(T, t)$, but in the future we will omit T to simplify the notation.

Let us denote the number of requests serviced in k-th unit at time instant t by $i_k(t)$, and the total number of epochs screened on k-th axis before time instant t by $n_k(t)$. Suppose that at some initial instant in time $t_0 < T$ the system is empty. Then, according to the method of constructing the screened processes, the following equality holds Moiseev and Nazarov (2016):

$$\mathrm{P}\{i_k(T) = m\} = \mathrm{P}\{n_k(T) = m\},$$

for $k = 1, 2$ and $m = 0, 1, \ldots$ This is the main idea of the dynamic screening method: we analyze two-dimensional process $\{n_1(t), n_2(t)\}$ and, substituting $t = T$, we obtain a result for the process $\{i_1(t), i_2(t)\}$ at time instant T.

For resource queueing systems, we denote the amount of resources allocated by the original request arrived at time instant t and screened on the k-th axis by $W_k(t)$, $k = 1, 2$. Then the probability distribution of the amount of resources allocated in the system at time instant T coincides with the probability distribution in the screened process Lisovskaya et al. (2016):

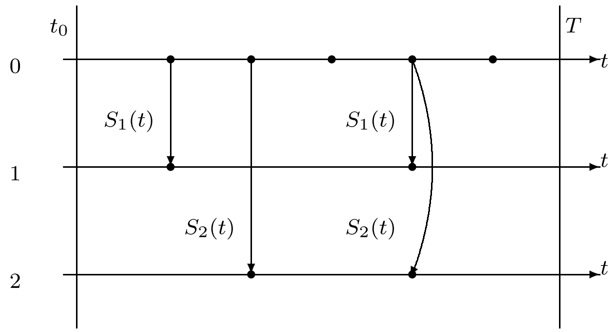

Fig. 2 Constructing of the screened processes

$$\mathrm{P}\left\{V_1(T) < z_1, V_2(T) < z_2\right\} = \mathrm{P}\left\{W_1(T) < z_1, W_2(T) < z_2\right\}. \tag{1}$$

So, we can analyze process $\{W_1(t), W_2(t)\}$ instead of process $\{V_1(t), V_2(t)\}$ and obtain a result for the last one by substituting $t = T$ into the result for process $\{W_1(t), W_2(t)\}$.

Process $\{W_1(t), W_2(t)\}$ is not Markovian, therefore we consider a three-dimensional Markov process $\{z(t), W_1(t), W_2(t)\}$, where $z(t)$ is the length of the interval from time instant t until the next request arrives in the system (residual time for the arrival process). We denote its probability distribution as follows:

$$P(z, w_1, w_2, t) = \mathrm{P}\{z(t) < z, W_1(t) < w_1, W_2(t) < w_2\}.$$

The following events may happen in interval $(t, t + \Delta t]$:

- no new customers enter into the system with probability $P(z + \Delta t, w_1, w_2, t) - P(\Delta t, w_1, w_2, t)$;
- exactly one customer enters into the system and will be in both service units in moment T, and the length of the interarrival period until the next customer's arrival is less than z – with probability $A(z)S_1(t)S_2(t) \int_0^{w_1}\int_0^{w_2} P(\Delta t, w_1 - y_1, w_2 - y_2, t)dG_1(y_1)dG_2(y_2)$;
- exactly one customer enters into the system and will be in only first service unit in moment T, and the length of the interarrival period until the next customer's arrival is less than z – with probability $A(z)S_1(t)(1 - S_2(t)) \int_0^{w_1} P(\Delta t, w_1 - y_1, w_2, t)dG_1(y_1)$;
- exactly one customer enters into the system and will be in only second service unit in moment T, and the length of the interarrival period until the next customer's arrival is less than z – with probability $A(z)(1 - S_1(t))S_2(t) \int_0^{w_2} P(\Delta t, w_1, w_2 - y_2, t)dG_2(y_2)$;
- exactly one customer enters into the system and will finish service on both service units before moment T, and the length of the interarrival period until the next customer's arrival is less than z – with probability $A(z)P(\Delta t, w_1, w_2, t)(1 - S_1(t))(1 - S_2(t))$;
- more than one customer enters into the system with probability $o(\Delta t)$.

For the three-dimensional Markov process, we apply the formula of total probability and write down the equality

$$\begin{aligned} P(z, w_1, w_2, t + \Delta t) = {} & \left[P(z + \Delta t, w_1, w_2, t) - P(\Delta t, w_1, w_2, t)\right] \\ & + A(z)\Bigg[P(\Delta t, w_1, w_2, t)(1 - S_1(t))(1 - S_2(t)) \\ & + S_1(t)(1 - S_2(t)) \int_0^{w_1} P(\Delta t, w_1 - y_1, w_2, t)dG_1(y_1) \\ & + (1 - S_1(t))S_2(t) \int_0^{w_2} P(\Delta t, w_1, w_2 - y_2, t)dG_2(y_2) \\ & + S_1(t)S_2(t) \int_0^{w_1}\int_0^{w_2} P(\Delta t, w_1 - y_1, w_2 - y_2, t)dG_1(y_1)dG_2(y_2)\Bigg] + o(\Delta t). \end{aligned} \tag{2}$$

Assuming that the partial derivatives $\frac{\partial P(z, w_1, w_2, t)}{\partial t}, \frac{\partial P(z, w_1, w_2, t)}{\partial z}$ exist, we obtain the following integral differential equation:

$$\begin{aligned} \frac{\partial P(z, w_1, w_2, t)}{\partial t} = & \frac{\partial P(z, w_1, w_2, t)}{\partial z} \\ & + \left[A(z)(1 - S_1(t))(1 - S_2(t)) - 1\right] \frac{\partial P(0, w_1, w_2, t)}{\partial z} \\ & + A(z) \Bigg[S_1(t)(1 - S_2(t)) \int_0^{w_1} \frac{\partial P(0, w_1 - y_1, w_2, t)}{\partial z} dG_1(y_1) \\ & + (1 - S_1(t))S_2(t) \int_0^{w_2} \frac{\partial P(0, w_1, w_2 - y_2, t)}{\partial z} dG_2(y_2) \\ & + S_1(t)S_2(t) \int_0^{w_1} \int_0^{w_2} \frac{\partial P(0, w_1 - y_1, w_2 - y_2, t)}{\partial z} dG_1(y_1)dG_2(y_2) \Bigg], \end{aligned} \tag{3}$$

with initial conditions

$$P(z, dw_1, dw_2, t_0) = r(z)\delta_{(0,0)}(dw_1 \times w_2).$$

Here $r(z)$ is the stationary distribution of the process $z(t)$ satisfying the differential equation Moiseeva and Nazarov (2012)

$$r'(z) + r'(0)[A(z) - 1] = 0, \tag{4}$$

whose solution is defined as follows: $r(z) = \kappa_1 \int_0^z (1 - A(x))dx$, where $\kappa_1 = \left[\int_0^\infty (1 - A(u))du \right]^{-1}$.

Let us introduce partial characteristic functions

$$h(z, u_1, u_2, t) = \int_0^\infty e^{ju_1 w_1} \int_0^\infty e^{ju_2 w_2} P(z, dw_1, dw_2, t),$$

where $j = \sqrt{-1}$.

$$\begin{aligned} \frac{\partial h(z, u_1, u_2, t)}{\partial t} = & \frac{\partial h(z, u_1, u_2, t)}{\partial z} + \frac{\partial h(0, u_1, u_2, t)}{\partial z} \Bigg\{ A(z) - 1 \\ & + A(z) \Big[S_1(t)(G_1^*(u_1) - 1) + S_2(t)(G_2^*(u_2) - 1) \\ & + S_1(t)S_2(t)(G_1^*(u_1) - 1)(G_2^*(u_2) - 1) \Big] \Bigg\}. \end{aligned} \tag{5}$$

with initial condition

$$h(z, u_1, u_2, t_0) = r(z). \tag{6}$$

4 Asymptotic Analysis

Since it is not possible to find an explicit form of a solution to the problem (5)–(6), we will seek the solution under the asymptotic condition of equivalent growing of service times in the units of servers. This asymptotic condition means proportional growth of the average service times in both service units and it is taken from practice. For example, during transmitting high-quality streaming video, resources are used to provide quality of the content, and service time corresponds to the duration of transmitting process. So, in such a case, average service times in both transmission channels will be proportional each other.

4.1 Asymptotic Analysis of the First Order

Denote average service time in k-th unit of servers by $b_k = \int_0^\infty \left[1 - B_k(x)\right]dx$, $k = 1, 2$. Then we can write the asymptotic condition of equivalent growing of service times in the units of servers in the form $b_k \to \infty$, $k = 1, 2$ while $\lim_{b_1\to\infty} \frac{b_1}{b_2} = \text{const}$.

Theorem 1 *The first-order asymptotic characteristic function $h^{(1)}(z, u_1, u_2, t) \approx h(z, u_1, u_2, t)$ of the probability distribution of the process $\{z(t), W_1(t), W_2(t)\}$ under the condition of equivalently growing service times has the form*

$$h^{(1)}(z, u_1, u_2, t) = r(z) \exp\left\{ j\kappa_1 \sum_{k=1}^{2} u_k a_k \int_{t_0}^{t} S_k(\tau) d\tau \right\}, \tag{7}$$

where $a_k = \int_0^\infty y_k dG_k(y_k)$ is the average amount of resources allocated by one request in the k-th service unit, $k = 1, 2$.

Proof Let us perform the following substitutions in Eqs. (5)–(6):

$$\begin{aligned} b_1 &= \frac{1}{q_1\varepsilon},\ b_2 = \frac{1}{q_2\varepsilon},\ t\varepsilon = \tau,\ t_0\varepsilon = \tau_0, \\ S_1(t) &= \bar{S}_1(\tau),\ S_2(t) = \bar{S}_2(\tau),\ u_1 = \varepsilon x_1,\ u_2 = \varepsilon x_2, \\ h(z, u_1, u_2, t) &= F_1(z, x_1, x_2, \tau, \varepsilon), \end{aligned} \tag{8}$$

we derive the equation

$$\begin{aligned} \varepsilon \frac{\partial F_1(z, x_1, x_2, \tau, \varepsilon)}{\partial \tau} &= \frac{\partial F_1(z, x_1, x_2, \tau, \varepsilon)}{\partial z} + \frac{\partial F_1(0, x_1, x_2, \tau, \varepsilon)}{\partial z} \Big\{ A(z) - 1 \\ &+ A(z)\big[S_1(\tau)(G_1^*(\varepsilon x_1) - 1) + S_2(\tau)(G_2^*(\varepsilon x_2) - 1) \\ &+ S_1(\tau)S_2(\tau)(G_1^*(\varepsilon x_1) - 1)(G_2^*(\varepsilon x_2) - 1)\big] \Big\}, \end{aligned} \tag{9}$$

with initial condition

$$F_1(z, x_1, x_2, \tau_0, \varepsilon) = r(z). \tag{10}$$

Performing the limit transition $\varepsilon \to 0$ in Eq. (9) and using the expansions

$$e^{j\varepsilon x_k y_k} = 1 + j\varepsilon x_k y_k + O(\varepsilon^2),$$
$$G_k^*(\varepsilon x_k) = \int_0^\infty e^{j\varepsilon x_k y_k} dG_k(y_k) = 1 + j\varepsilon x_k \int_0^\infty y_k dG_k(y_k) + O(\varepsilon^2)$$
$$= 1 + j\varepsilon x_k a_k + O(\varepsilon^2),$$

and

$$\lim_{\varepsilon\to 0} G^*(\varepsilon x_k) = \lim_{\varepsilon\to 0}(1 + j\varepsilon x_k a_k + O(\varepsilon^2)) = 1,$$

we obtain

$$\frac{\partial F_1(z, x_1, x_2, \tau)}{\partial z} + \frac{\partial F_1(0, x_1, x_2, \tau)}{\partial z}[A(z) - 1] = 0.$$

Suppose function $F_1(z, x_1, x_2, \tau, \varepsilon)$ is differentiable in $\varepsilon = 0$. Then, taking into account (4), we will find it in the expansion form (the Maclaurin series on powers of ε)

$$F_1(z, x_1, x_2, \tau, \varepsilon) = r(z)\Phi_1(x_1, x_2, \tau) + O(\varepsilon), \tag{11}$$

where $\Phi_1(x_1, x_2, \tau)$ is some function which satisfies initial condition

$$\Phi_1(x_1, x_2, \tau_0) = 1.$$

Then let us substitute (11) into (9):

$$\varepsilon r(z)\frac{\partial \Phi_1(x_1, x_2, \tau)}{\partial \tau} + O(\varepsilon^2) = r'(z)\Phi_1(x_1, x_2, \tau) + \lambda\Phi_1(x_1, x_2, \tau)$$
$$\cdot \Big\{A(z) - 1 + A(z)\big[S_1(\tau)(G_1^*(\varepsilon x_1) - 1) + S_2(\tau)(G_2^*(\varepsilon x_2) - 1)$$
$$+ S_1(\tau)S_2(\tau)(G_1^*(\varepsilon x_1) - 1)(G_2^*(\varepsilon x_2) - 1)\big]\Big\} + O(\varepsilon^2), \tag{12}$$

where denoting $r'(0) = \lambda$. Let us, make in (12) limit transition $z \to \infty$:

$$\varepsilon\frac{\partial \Phi_1(x_1, x_2, \tau)}{\partial \tau} + O(\varepsilon^2) = \lambda\Phi_1(x_1, x_2, \tau)$$
$$\cdot \Big\{A(z) - 1 + A(z)\big[S_1(\tau)(G_1^*(\varepsilon x_1) - 1) + S_2(\tau)(G_2^*(\varepsilon x_2) - 1)$$
$$+ S_1(\tau)S_2(\tau)(G_1^*(\varepsilon x_1) - 1)(G_2^*(\varepsilon x_2) - 1)\big]\Big\} + O(\varepsilon^2),$$

Dividing the obtained equation by ε and tend $\varepsilon \to 0$, we derive the following differential equation for function $\Phi_1(x_1, x_2, \tau)$:

$$\frac{\partial \Phi_1(x_1, x_2, \tau)}{\partial \tau} = j\kappa_1\Phi_1(x_1, x_2, \tau)[\bar{S}_1(\tau)x_1 a_1 + \bar{S}_2(\tau)x_2 a_2],$$

which solution is as follows:

$$\Phi_1(x_1, x_2, \tau) = \exp\left\{ j\kappa_1 \sum_{k=1}^{2} x_k a_k \int_{\tau_0}^{\tau} \bar{S}_k(\xi) d\xi \right\}. \tag{13}$$

Substituting (13) into (11) and performing substitutions inverse to (8), we obtain

$$\begin{aligned} h(z, u_1, u_2, t) = F_1(z, x_1, x_2, \tau, \varepsilon) &\approx r(z) \exp\left\{ j\kappa_1 \sum_{k=1}^{2} x_k a_k \int_{\tau_0}^{\tau} \bar{S}_k(\xi) d\xi \right\} \\ &= r(z) \exp\left\{ j\kappa_1 \sum_{k=1}^{2} u_k a_k \int_{t_0}^{t} S_k(\tau) d\tau \right\} = h^{(1)}(z, u_1, u_2, t). \end{aligned} \tag{14}$$

The theorem is proved.

Remark Function $F_1(z, x_1, x_2, \tau, \varepsilon)$ must satisfy the condition of existing finite derivative in $\varepsilon = 0$ to we be able to use expansion (11). But the problem of checking such property or trying to establish its meaningful for specific distribution/characteristic functions seem very hard to be solved and lay out of the topic of the paper. Due to this, we rely on the fact that we obtain an expression for the characteristic function at the end of the derivations, and on the comparison of this analytical result with a simulation one presented in Sect. 5. The same applies to the differentiability of function $F_2(z, x_1, x_2, \tau, \varepsilon)$ in the next section.

Letting $z \to \infty$ in (14), we obtain the complete characteristic function of the stochastic process $\{V_1(t), V_2(t)\}$, and setting $t = T$ and letting $t_0 \to -\infty$, we obtain the characteristic function of the stationary probability distribution of the two-dimensional process under study:

$$h(u_1, u_2) = \exp\left\{ j\kappa_1 \sum_{k=1}^{2} u_k a_k b_k \right\}.$$

The obtained approximation determines only the means of resource amounts. To construct a qualitatively more accurate approximation, we perform a second-order asymptotic analysis.

4.2 Asymptotic Analysis of the Second Order

Let us prove the following statement.

Theorem 2 *The second-order asymptotic characteristic function* $h^{(2)}(z, u_1, u_2, t) \approx h(z, u_1, u_2, t)$ *of the probability distribution of the process* $\{z(t), W_1(t), W_2(t)\}$ *under the condition of equivalently growing service times has the form*

$$h^{(2)}(z,u_1,u_2,t) = r(z)\exp\left\{ j\kappa_1 \sum_{k=1}^{2} u_k a_k \int_{t_0}^{t} S_k(\tau)d\tau \right.$$
$$+ \sum_{k=1}^{2} \frac{(ju_k)^2}{2}\left[\kappa_1 \alpha_k \int_{t_0}^{t} S_k(\tau)d\tau + \kappa_2 a_k^2 \int_{t_0}^{t} S_k^2(\tau)d\tau\right] \tag{15}$$
$$\left. + ju_1 ju_2(\kappa_1+\kappa_2)a_1 a_2 \int_{t_0}^{t} S_1(\tau)S_2(\tau)d\tau \right\},$$

where $\kappa_2 = \kappa_1^3(\sigma^2 - a^2)$, $\alpha_k = \int_0^\infty y_k^2 dG_k(y_k)$, $k = 1, 2$, *a and* σ^2 *are the mean and variance of random variable given by distribution function* $A(x)$.

Proof Taking into account expression (7) for function $h^{(1)}(z,u_1,u_2,t)$, we make the following substitution in (3):

$$h(z,u_1,u_2,t) = h_2(z,u_1,u_2,t)\exp\left\{ j\kappa_1 \sum_{k=1}^{2} u_k a_k \int_{t_0}^{t} S_k(\tau)d\tau \right\}. \tag{16}$$

We obtain

$$\frac{\partial h_2(z,u_1,u_2,t)}{\partial t} + h_2(z,u_1,u_2,t) j\kappa_1 \sum_{k=1}^{2} a_k u_k S_k(t) = \frac{\partial h_2(z,u_1,u_2,t)}{\partial z}$$
$$+ \frac{\partial h_2(0,u_1,u_2,t)}{\partial z}\Big\{ A(z) - 1 + A(z)\big[S_1(t)(G_1^*(u_1)-1) + S_2(t)(G_2^*(u_2)-1)$$
$$+ S_1(t)S_2(t)(G_1^*(u_1)-1)(G_2^*(u_2)-1)\big]\Big\}.$$

Let us make the following substitutions in this equation:

$$b_1 = \frac{1}{q_1\varepsilon^2},\ b_2 = \frac{1}{q_2\varepsilon^2},\ t\varepsilon^2 = \tau,\ t_0\varepsilon^2 = \tau_0,$$
$$S_1(t) = \bar{S}_1(\tau),\ S_2(t) = \bar{S}_2(\tau),\ u_1 = \varepsilon x_1,\ u_2 = \varepsilon x_2, \tag{17}$$
$$h_2(z,u_1,u_2,t) = F_2(z,x_1,x_2,\tau,\varepsilon).$$

We obtain equation

$$
\begin{aligned}
\varepsilon^2 \frac{\partial F_2(z,x_1,x_2,\tau,\varepsilon)}{\partial \tau} &+ F_2(z,x_1,x_2,\tau,\varepsilon) j\kappa_1 \sum_{i=1}^{2} a_k \varepsilon x_k \bar{S}_k(\tau) \\
&= \frac{\partial F_2(z,x_1,x_2,\tau,\varepsilon)}{\partial z} + \frac{\partial F_2(0,x_1,x_2,\tau,\varepsilon)}{\partial z} \Big\{ A(z) - 1 \\
&+ A(z)\big[\bar{S}_1(\tau)(G_1^*(\varepsilon x_1) - 1) + \bar{S}_2(\tau)(G_2^*(\varepsilon x_2) - 1) \\
&+ \bar{S}_1(\tau)\bar{S}_2(\tau)(G_1^*(\varepsilon x_1) - 1)(G_2^*(\varepsilon x_2) - 1)\big] \Big\}
\end{aligned}
\tag{18}
$$

with initial condition

$$
F_2(z,x_1,x_2,\tau_0,\varepsilon) = r(z). \tag{19}
$$

Let us make here asymptotic transition $\varepsilon \to 0$, denoting $\lim_{\varepsilon\to 0} F_2(z,x_1,x_2,\tau,\varepsilon) = F_2(z,x_1,x_2,\tau)$, we derive

$$
\frac{\partial F_2(z,x_1,x_2,\tau)}{\partial z} + \frac{\partial F_2(0,x_1,x_2,\tau)}{\partial z}(A(z) - 1) = 0.
$$

Suppose function $F_2(z,x_1,x_2,\tau,\varepsilon)$ has finite derivatives of the second and greater orders in $\varepsilon = 0$. Then we can represent this function in the expansion form

$$
\begin{aligned}
F_2(z,x_1,x_2,\tau,\varepsilon) = \Phi_2(x_1,x_2,\tau)\big[r(z) + j\varepsilon(x_1 a_1 \bar{S}_1(\tau) \\
+ x_2 a_2 \bar{S}_2(\tau)) f_2(z)\big] + O(\varepsilon^2),
\end{aligned}
\tag{20}
$$

where $\Phi_2(x_1,x_2,\tau)$ and $f_2(z)$ are some functions. The first one satisfies the initial condition $\Phi_2(x_1,x_2,\tau_0) = 1$.

Expansion (20) is the Maclaurin series on powers of ε where the second term is chosen in the specific form to avoid cumbersome derivations that the authors have made before their presentation in the paper, including the property $f_2(\infty) = 0$.

Substituting (20) into (18), dividing on ε and setting $\varepsilon \to 0$, we derive the following equation for function $f_2(z)$:

$$
\kappa_1(r(z) - A(z)) - f_2'(z) - f_2'(0)(A(z) - 1) = 0,
$$

where $r(z) = \kappa_1 \int_0^z (1 - A(u))du$.

Let us substitute expansion (20) into (18) and make asymptotic transition $z \to \infty$, we derive the following equation for function $\Phi_2(x_1,x_2,\tau)$:

$$
\begin{aligned}
\frac{\partial \Phi_2(x_1,x_2,\tau)}{\partial \tau} = \Phi_2(x_1,x_2,\tau)\Big\{ \sum_{k=1}^{2} \frac{(jx_k)^2}{2}\big[\kappa_1 \alpha_k \bar{S}_k(\tau) + \kappa_2 a_k^2 \bar{S}_k^2(\tau)\big] \\
+ jx_1 jx_2(\kappa_1 + \kappa_2) a_1 a_2 \bar{S}_1(\tau)\bar{S}_2(\tau) \Big\}.
\end{aligned}
$$

Its solution is as follows:

$$\Phi_2(x_1,x_2,\tau) = \exp\left\{\sum_{k=1}^{2}\frac{(jx_k)^2}{2}\left[\kappa_1\alpha_k\int_{\tau_0}^{\tau}\bar{S}_k(\xi)d\xi + \kappa_2 a_k^2\int_{\tau_0}^{\tau}\bar{S}_k^2(\xi)d\xi\right]\right.$$
$$\left. + jx_1jx_2(\kappa_1+\kappa_2)a_1a_2\int_{\tau_0}^{\tau}\bar{S}_1(\xi)\bar{S}_2(\xi)d\xi\right\}.$$

Substituting this expression into (20) and making substitutions inverse to (17), we obtain:

$$\begin{aligned} h_2(z,u_1,u_2,t) &= F_2(z,x_1,x_2,\tau,\varepsilon) \\ &\approx r(z)\exp\left\{\sum_{k=1}^{2}\frac{(ju_k)^2}{2}\left[\kappa_1\alpha_k\int_{t_0}^{t}S_k(\tau)d\tau + \kappa_2 a_k^2\int_{t_0}^{t}S_k^2(\tau)d\tau\right]\right. \\ &\left. + ju_1ju_2(\kappa_1+\kappa_2)a_1a_2\int_{t_0}^{t}S_1(\tau)S_2(\tau)d\tau\right\} = h^{(2)}(z,u_1,u_2,t). \end{aligned} \tag{21}$$

4.3 Main result

Let us set in (15) $z \to \infty$, $t = T$ to satisfy (1), and $t_0 \to -\infty$ to reach steady-state regime, then we obtain the following expression for characteristic function of joint distribution of resources allocated in both units of servers in steady-state regime:

$$\begin{aligned} h(u_1,u_2) = \exp&\left\{j\kappa_1(u_1a_1b_1 + u_2a_2b_2)\right. \\ &+ \frac{(ju_1)^2}{2}(\kappa_1\alpha_1b_1 + \kappa_2a_1^2\beta_1) + \frac{(ju_2)^2}{2}(\kappa_1\alpha_2b_2 + \kappa_2a_2^2\beta_2) \\ &\left. + ju_1ju_2(\kappa_1+\kappa_2)a_1a_2\beta_{12}\right\}, \end{aligned} \tag{22}$$

where

$$\lim_{t_0\to-\infty}\int_{t_0}^{T}S_k(x)dx\bigg|_{t=T} = \int_{-\infty}^{0}\left[1-B_k(T-x)\right]dx = \int_{0}^{\infty}\left[1-B_k(x)\right]dx \doteq b_k,$$

$$\lim_{t_0\to-\infty}\int_{t_0}^{T}S_k^2(x)dx\bigg|_{t=T} = \int_{-\infty}^{0}\left[1-B_k(T-x)\right]^2dx = \int_{0}^{\infty}\left[1-B_k(x)\right]^2dx \doteq \beta_k,\ k=1,2,$$

$$\lim_{t_0\to-\infty}\int_{t_0}^{T}S_1(x)S_2(x)dx\bigg|_{t=T} = \int_{-\infty}^{0}\left[1-B_1(T-x)\right]\left[1-B_2(T-x)\right]dx = \int_{0}^{\infty}\left[1-B_1(x)\right]\left[1-B_2(x)\right]dx \doteq \beta_{12}.$$

So, we can see that the probability distribution of the two-dimensional process $\{V_1(t), V_1(t)\}$ is asymptotically Gaussian with mean vector

$$\mathbf{a} = \kappa_1 [a_1 b_1 \;\; a_2 b_2] \tag{23}$$

and covariance matrix

$$\mathbf{K} = \begin{bmatrix} \kappa_1 \alpha_1 b_1 + \kappa_2 a_1^2 \beta_1 & (\kappa_1 + \kappa_2) a_1 a_2 \beta_{12} \\ (\kappa_1 + \kappa_2) a_1 a_2 \beta_{12} & \kappa_1 \alpha_2 b_2 + \kappa_2 a_2^2 \beta_2 \end{bmatrix}. \tag{24}$$

We may notice that in the case of the Poisson arrival process we obtain exact matching with a result obtained in Bushkova et al. (2019).

5 Numerical Analysis and Example

Since we have no exact results to compare obtained approximation with, we use simulation and perform analysis of the approximation accuracy and applicability area in a numerical form.

Consider the queueing system described in Sect. 2 with the following parameters. Inter-arrival time has a gamma distribution with shape and rate parameters both equal to 0.5. So, the intensity of the arrival process κ_1 is equal to 1, and we obtain $\kappa_2 = 1$ too.

Let us represent asymptotic parameter ε in the form $\varepsilon = 1/N$, where N be some scale parameter. This is made just for convenience of presentation.

Service times have gamma distributions with shape and rate parameters equal to 0.5 and 0.5ε for the first unit of servers, and 1.5 and 0.75ε for the second one. So, the average service time b_1 is equal to N for the first unit of servers and $b_2 = 2N$ for the second one.

We choose gamma distributions for the probability distribution of resources allocated by a single request. Their shape and rate parameters are equal both to 2.5 for the first service unit, and 0.4, 0.8 for the second one. So, the average amount of resources allocated by a request in the first unit of servers (a_1) is equal to 1 and it is equal to 0.5 for the second unit (a_2).

So, according to (23), we obtain the average amount of resources allocated in each unit of servers in a steady-state regime equal to N.

Varying scale (asymptotic) parameter N, we can obtain an applicability area for obtained results. To estimate the accuracy of results, we perform simulations and calculate error in the form of Kolmogorov distance

$$d = \max_{y_1, y_2} \left| F(y_1, y_2) - G(y_1, y_2) \right|,$$

where $F(y_1, y_2)$ is empiric cumulative distribution function constructed on the base of simulations results and $G(y_1, y_2)$ is two-dimensional Gaussian cumulative distribution function with means vector (23) and covariance matrix (24).

Obtained results are presented in the first row of Table 1. The same analysis also is done for various types of probability distributions for allocated resources with parameters that give $a_1 = 1$ and $a_2 = 0.5$ as for the case of a gamma distribution. In Table 1, you can find the results for Uniform, Weibull, Log-normal, and Pareto distributions. As we can notice, for all cases, while N is growing, Kolmogorov distance d is decreasing, therefore, results (22)–(24) become more precise. For example, if we choose $d \leq 0.05$ as enough error for the results accuracy, then we can conclude that results (22)–(24) are applicable for $N \geq 40$ ($\varepsilon \leq 0.025$). Figures 3 and 4 show the probability density functions of the Gaussian approximation (22)–(24) in comparison with distribution built on the base of simulations

Table 1 Kolmogorov distance for various values of scale parameter *N* and various distributions of allocated resources

N	10	20	40	100	200
Gamma	0.091	0.067	0.046	0.031	0.019
Uniform	0.086	0.066	0.043	0.032	0.025
Weibull	0.091	0.067	0.048	0.033	0.019
Log-normal	0.091	0.063	0.049	0.031	0.025
Pareto	0.092	0.060	0.044	0.028	0.024

result for the first and the second service units and various values of scale parameter N. As we can see, while N is increasing, the approximation and actual distribution get closer to each other.

In practice, we usually have no infinite reserve of resources to be used, but we can use the results from Sect. 4.3 in the case of the system with limited resources. Consider the problem of design such transmission system and the task of choosing maximum values (limits) of resources provided in each channel. On one hand, these limit values should be minimal as possible to make a cost of the system very low. On the other

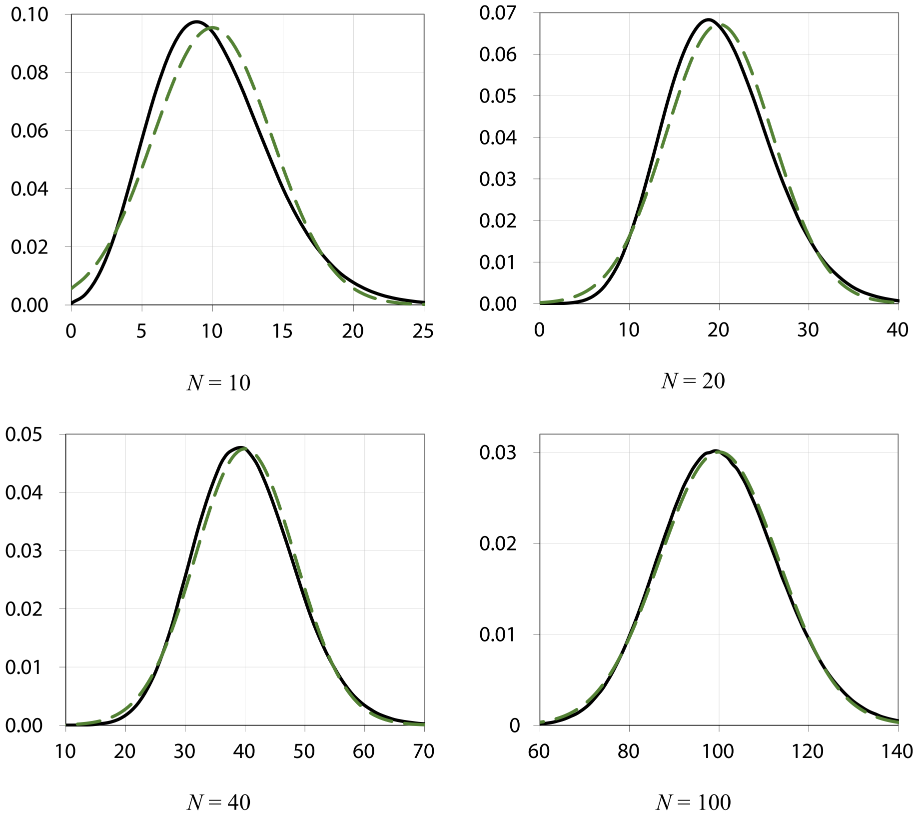

Fig. 3 Probability density functions of Gaussian approximation (dashed line) and distribution of simulations result (solid line) for the first service unit for various values of scale parameter N

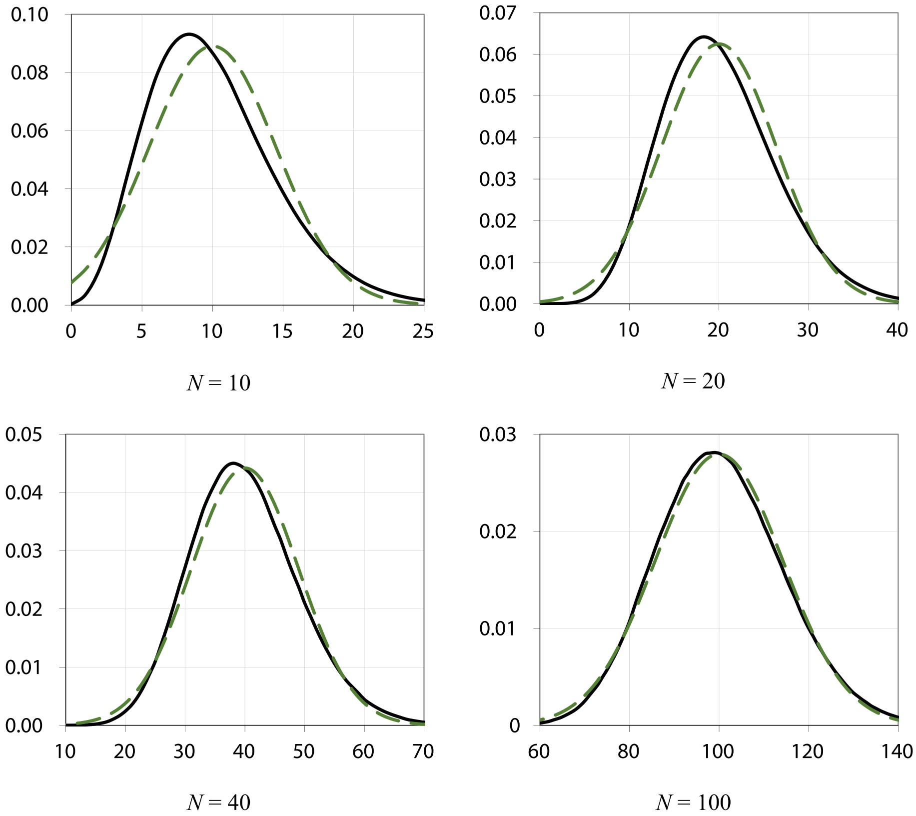

Fig. 4 Probability density functions of Gaussian approximation (dashed line) and distribution of simulations result (solid line) for the second service unit for various values of scale parameter N

hand, these values should be enough to avoid losing of requests due to the absence of resources for their servicing. We call values of the total amount in resources that satisfy both criteria as *optimal* and denote them by a_k^{opt} where k is the number of the channel. Evaluation of these values is based on the Gaussian distribution (22)–(24) and technique explained in Moiseev and Nazarov (2016). Consider the probability that a request arrived in the system will be rejected due to there is no enough amount of resources for its transmission in any of the channels because the most resources are occupied by other requests. Let us estimate the total amount of resources in each channel that will be enough to make this probability be not more than some given value p_{loss}. According to Moiseev and Nazarov (2016), due to the joint probability distribution of occupied resources is two-dimensional Gaussian, we can evaluate the optimal values of the total resources in each channel using the following expressions:

$$a_k^{opt} = a_k + r\sqrt{K_{kk}}, \quad k = 1, 2. \tag{25}$$

Here r is the so-called hyper-ellipsoid radius that depends on p_{loss} and can be evaluated using the method described in Moiseev and Nazarov (2016).

For example, consider the system described above with gamma distribution of resource allocation. We choose $N = 40$ to be sure that approximations (22)–(24) are accurate

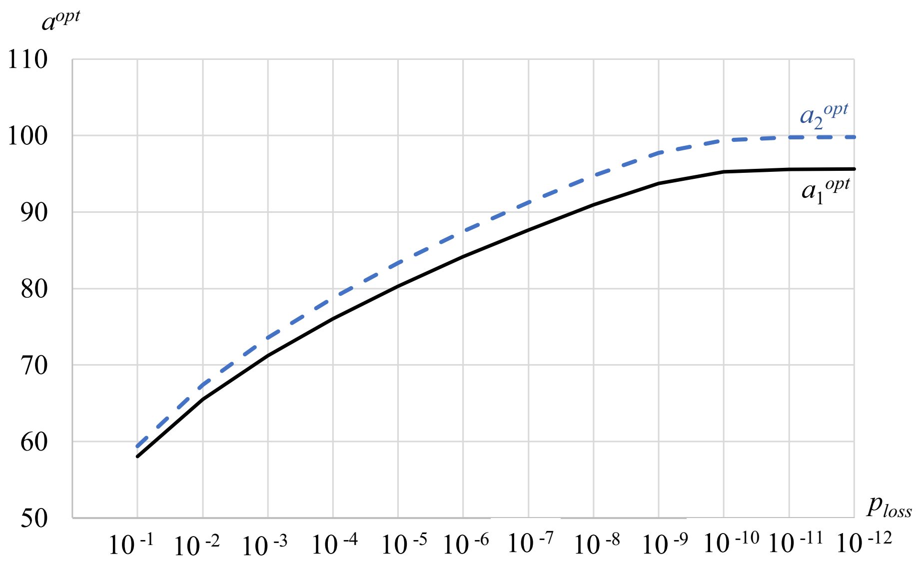

Fig. 5 Optimal amounts of resources a_1^{opt} and a_2^{opt} for channels of a finite-resource system when various values of losing probability p_{loss} is choosing

enough. Then we obtain the following mean vector and covariance matrix for occupied resources in both channels:

$$\mathbf{a} = [40 \quad 40], \quad \mathbf{K} \approx \begin{bmatrix} 70.535 & 25.138 \\ 25.138 & 81.512 \end{bmatrix}.$$

Applying formula (25), if we choose value $p_{loss} = 10^{-3}$ then $r \approx 3.717$ and we calculate (rounding up): $a_1^{opt} = 72$, $a_2^{opt} = 74$. These values are the total amounts of resources in the corresponding channels that will be enough to a new request arrived in the system in any time with probability $1 - p_{loss}$ finds enough resources in both channels to be transmitted in parts via the channels. We may interpret these amounts as the number of radio transmission bands used in the channels of the system. It is interesting that if we choose $p_{loss} = 10^{-6}$ decreasing losing probability by a thousand, we obtain $r \approx 5.256$ and $a_1^{opt} = 85$, $a_2^{opt} = 88$ which means that we need enlarge amount of available radio resources only by 20% approximately. Figure 5 depicts changing of a_1^{opt} and a_2^{opt} while p_{loss} decreases.

6 Conclusions

According to Cisco, it is expected that by 2022 the number of devices connected to wireless networks will reach about 3.5 devices per person. Video content will be the main part of the traffic. The amount of allocated radio resources for such traffic does not affect the duration of the sessions but affects the quality of the video image. The 3GPP Project Coordination Group is considering now supporting active links between multiple Radio Access Technologies simultaneously to ensure session continuity in fifth-generation wireless networks. One of the solutions may be a splitting of transmitted data into two streams one of which uses 5G New Radio and another uses 4G Long-Term Evolution technology.

In the paper, we have considered the problem of analyzing the total amount of resources allocated in both channels. To do this, we have used an infinite-server resource queueing model with the splitting of incoming requests into two service units (transmitting channels) that have unlimited resources assuming the channels have great enough bandwidth. We have found the solution under the asymptotic condition of equivalent growth of service times in the units which has a practical interpretation: during transmitting high-quality streaming traffic, resources are used to provide quality of the content, and service time corresponds to the duration of transmitting process. We have found the solution in the form of the distribution approximation and then we have analyzed the applicability area of it. Due to the resulting distribution (approximation) is two-dimensional Gaussian, we have shown for example how to estimate the reserve amount of resources for the system with finite limitations of resources.

The approach presented in the paper may be applied to study resource queues with the splitting of requests with the arrival processes of other various types (Markovian arrival process, semi-Markov process, a mixture of arrival processes, etc.). And also to obtain the main characteristics of splitting of transmitted data into three and more streams (for example, 5G New Radio, 4G Long-Term Evolution and Wi-Fi technology). But we note that the formulas' dimensional will be extra-large.

Acknowledgements Authors are very grateful to the reviewers of the paper for their comments which cardinally improved its quality

Funding Open access funding provided by University of Debrecen.

References

Ageev K, Sopin E, Samouylov K (2018) Simulation of the Limited Resources Queuing System with Signals. In: 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). pp 1–5. https://doi.org/10.1109/ICUMT.2018.8631246

Araniti G, Condoluci M, Cotronei M, Iera A, Molinaro A (2015) A Solution to the Multicast Subgroup Formation Problem in LTE Systems. IEEE Wireless Commun Lett 4(2):149–152. https://doi.org/10.1109/LWC.2014.2387824

Artalejo JR, Gomez-Corral A (2008) The Theory of Queuing Systems with Correlated Flows. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78725-9

Basharin G, Samouylov K, Yarkina N, Gudkova I (2009) A new stage in mathematical teletraffic theory. Autom Remote Control 70(12):1954–1964. https://doi.org/10.1134/S0005117909120030

Benkhelifa F, ElSawy H, McCann JA, Alouini M (2020) Recycling cellular energy for self-sustainable IoT networks: A spatiotemporal study. IEEE Transactions on Wireless Communications. pp 1–14. https://doi.org/10.1109/TWC.2020.2967697

Bushkova T, Danilyuk E, Moiseeva S, Pavlova E (2019) Resource Queueing System with Dual Requests and Their Parallel Service. Commun Comput Inf Sci 1141:364–374. https://doi.org/10.1007/978-3-030-36625-4_29

Buturlin I, Gaidamaka Y, Samouylov A (2012) Utility function maximization problems for two cross-layer optimization algorithms in OFDM wireless networks. In: 2012 IV International Congress on Ultra Modern Telecommunications and Control Systems. pp. 63–65. https://doi.org/10.1109/ICUMT.2012.6459745

Cisco (2018) Cisco VNI Global IP Traffic Forecast, 2017–2022. https://www.cisco.com/c/dam/m/en_us/network-intelligence/service-provider/digital-transformation/knowledge-network-webinars/pdfs/1213-business-services-ckn.pdf

Chu P, Zhang A, Wang X, Fei Z, Fang G, Wang D (2020) Interference characterization and power optimization for automotive radar with directional antenna. IEEE Transactions on Vehicular Technology. pp 1–13. https://doi.org/10.1109/TVT.2020.2968929

Dudin A, Klimenok V, Vishnevsky V (2020) The theory of queuing systems with correlated flows. Springer Nature. https://doi.org/10.1007/978-3-030-32072-0

Falin G, Templeton J (1997) Retrial Queues. Chapman and Hall/CRC

Galileyskaya A, Lisovskaya E, Fedorova E (2019) Resource queueing system with the requirements copying at the second phase. Commun Comput Inf Sci 1141:352–363. https://doi.org/10.1007/978-3-030-36625-4_28

Galinina O, Andreev S, Gerasimenko M, Koucheryavy Y, Himayat N, Yeh S, Talwar S (2014) Capturing Spatial Randomness of Heterogeneous Cellular/WLAN Deployments With Dynamic Traffic. IEEE J Sel Areas Commun 32(6):1083–1099. https://doi.org/10.1109/JSAC.2014.2328172

Galinina O, Andreev S, Turlikov A, Koucheryavy Y (2014) Optimizing Energy Efficiency of a Multi-Radio Mobile Device in Heterogeneous beyond-4G Networks. Perform Eval 78:18–41. https://doi.org/10.1016/j.peva.2014.06.002

Guan W, Wen X, Wang L, Lu Z, Shen Y (2018) A Service-Oriented Deployment Policy of End-to-End Network Slicing Based on Complex Network Theory. IEEE Access 6:19691–19701. https://doi.org/10.1109/ACCESS.2018.2822398

Gudkova I, Samouylov K, Buturlin I, Borodakiy V, Gerasimenko M, Galinina O, Andreev S (2014) Analyzing impacts of coexistence between M2M and H2H Communication on 3GPP LTE System. Lect Notes Comput Sci 8458:162–174. https://doi.org/10.1007/978-3-319-13174-0_13

Lisovskaya E, Moiseev A, Moiseeva S, Pagano M (2019) Modeling of mathematical processing of physics experimental data in the form of a non-Markovian multi-resource queuing system. Russ Phys J 61(12):2188–2196. https://doi.org/10.1007/s11182-019-01655-6

Lisovskaya E, Moiseeva S, Pagano M (2016) The Total Capacity of Customers in the Infinite-Server Queue with MMPP Arrivals. Commun Comput Inf Sci 678:110–120. https://doi.org/10.1007/978-3-319-51917-3_11

Lisovskaya E, Moiseeva S, Pagano M (2017) Infinite-server tandem queue with renewal arrivals and random capacity of customers. Commun Comput Inf Sci 700:201–216. https://doi.org/10.1007/978-3-319-66836-9_17

Lisovskaya E, Pankratova E, Gaidamaka Y, Moiseeva S, Pagano M (2019) Heterogeneous queueing system $MAP / GI^{(n)} / \infty$ with random customers' capacities. Lect Notes Comput Sci 11965:315–329. https://doi.org/10.1007/978-3-030-36614-8_24

Moiseev A, Nazarov A (2012) Investigation of high intensive general flow. In: 2012 IV International Conference "Problems of Cybernetics and Informatics" (PCI). pp 1–3. https://doi.org/10.1109/ICPCI.2012.6486406

Moiseev A, Nazarov A (2014) Asymptotic analysis of the infinite-server queueing system with high-rate semi-markov arrivals. In: 2014 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). pp 507–513. https://doi.org/10.1109/ICUMT.2014.7002153

Moiseev A, Nazarov A (2016) Queueing network $MAP - (GI/\infty)^K$ with high-rate arrivals. Eur J Oper Res 254(1):161–168. https://doi.org/10.1016/j.ejor.2016.04.011

Naumov V, Samouylov K (2017) Analysis of multi-resource loss system with state-dependent arrival and service rates. Probab Eng Inf Sci 31(4):413–419. https://doi.org/10.1017/S0269964817000079

Naumov V, Samouylov K, Samouylov A (2016) On the total amount of resources occupied by serviced customers. Autom Remote Control 77(8):1419–1427. https://doi.org/10.1134/S0005117916080087

Samouylov A, Moltchanov D, Krupko A, Kovalchukov R, Moskaleva F, Gaidamaka Y (2018) Performance Analysis of Mixture of Unicast and Multicast Sessions in 5G NR Systems. In: 2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). pp 1–7. https://doi.org/10.1109/ICUMT.2018.8631230

Samouylov K, Naumov V, Sopin E, Gudkova I, Shorgin S (2016) Sojourn time analysis for processor sharing loss system with unreliable server. In: Analytical and Stochastic Modelling Techniques and Applications. pp 284–297. https://doi.org/10.1007/978-3-319-43904-4_20

Song C, Zhang M, Zhan Y, Wang D, Guan L, Liu W, Zhang L, Xu S (2019) Hierarchical edge cloud enabling network slicing for 5G optical fronthaul. IEEE/OSA J Opt Commun Networking 11(4):B60–B70. https://doi.org/10.1364/JOCN.11.000B60

Sopin E, Vikhrova O, Samouylov K (2017) LTE network model with signals and random resource requirements. In: 2017 9th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT). pp 101–106. https://doi.org/10.1109/ICUMT.2017.8255155

Tikhonenko O (2010) Queuing system with processor sharing and limited resources. Autom Remote Control 71(5):803–815. https://doi.org/10.1134/S0005117910050073

Authors and Affiliations

Tatyana Bushkova[1] · Svetlana Moiseeva[1] · Alexander Moiseev[1] · János Sztrik[2] · Ekaterina Lisovskaya[1] · Ekaterina Pankratova[3]

Tatyana Bushkova
bushkova70@mail.ru

Svetlana Moiseeva
smoiseeva@mail.ru

Alexander Moiseev
moiseev.tsu@gmail.com

Ekaterina Lisovskaya
ekaterina_lisovs@mail.ru

Ekaterina Pankratova
pankate@sibmail.com

1 Institute of Applied Mathematics and Computer Science, Tomsk State University, 634050 Tomsk, Russia

2 Faculty of Informatics, University of Debrecen, Debrecen, Hungary

3 Laboratory of the Automated Systems of Mass Service, V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, 117342 Moscow, Russia