Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (2) 105–119
http://journals.rudn.ru/miph

Research article

UDC 519.872
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-2-105-119
EDN: VUBLKP

Heterogeneous queueing system with Markov renewal arrivals and service times dependent on states of arrival process

Evgeny P. Polin[1,2], **Svetlana P. Moiseeva**[1], **Alexander N. Moiseev**[1]

[1] *National Research Tomsk State University, 36, Lenin Avenue, Tomsk, 634050, Russian Federation*
[2] *National Research Tomsk Polytechnic University, 30, Lenin Avenue, Tomsk, 634050, Russian Federation*

(received: April 14, 2023; revised: April 25, 2023; accepted: June 26, 2023)

Abstract. In the proposed work, we consider a heterogeneous queueing system with a Markov renewal process and an unlimited number of servers. The service time for requests on the servers is a positive random variable with an exponential probability distribution. The service parameters depend on the state of the Markov chain nested over the renewal moments. It should be noted that these parameters do not change their values until the end of maintenance. Thus, the devices in the system under consideration are heterogeneous. The object of the study is a multidimensional random process — the number of servers of each type being served with different intensities in the stationary regime. The method of asymptotic analysis under the condition of equivalent growing of service times in the units of servers is applied for the study. The method of asymptotic analysis is implemented in the construction of a sequence of asymptotic of increasing order, in which the asymptotic of the first order determines the asymptotic mean value of the number of occupied servers. The second-order asymptotic allows one to construct a Gaussian approximation of the probability distribution of the number of occupied servers in the system. It is shown that this approximation coincides with the Gaussian distribution.

Key words and phrases: queuing system, random environment, Markov renewal process, asymptotic analysis method

1. Introduction

Queueing theory is a field of applied mathematics that deals with the study and analysis of processes in various service, production, management, and communication systems in which homogeneous events are repeated many times. Examples of such systems include consumer services; systems for receiving, processing, and transmitting information, automatic production lines, telecommunication systems, and others [1].

The independence of processes in queueing systems is generally assumed when developing queueing models. However, real systems often involve several process dependencies, and failure to consider these can lead to a serious under errors in the estimation of the performance measures. Semi-Markov processes are used in modeling stochastic control problems arising in Markovian dynamic systems where the sojourn time in each state is a general continuous random variable. They are powerful, natural tools for the optimization of queues, production scheduling, reliability/maintenance [2, 3]. For example, in a machine replacement problem with deteriorating performance over time, a decision-maker, after observing the current state of the machine, decides whether to continue its usage, initiate maintenance (preventive or corrective) repair or replace the machine.

Semi-Markov Processes include renewal processes and continuous-time Markov chains as special cases. In a semi-Markov process similar to Markov chains, state changes occur according to the Markov property, i.e., states in the future do not depend on the states in the past given the present. However, the sojourn time in a state is a continuous random variable with distribution depending on that state and the next state. A renewal process is a generalization of a Poisson process that allows arbitrary holding times. Its applications include such as planning for replacing worn-out machinery in a factory. A Markov renewal process is a generalization of a renewal process that the sequence of holding times is not independent and identically distributed. Their distributions depend on the states in a Markov chain. The Markov renewal processes were studied by Pyke in the 1960s [4, 5].

In the proposed work, we consider a heterogeneous queueing system (QS) with a Markov renewal process (MRP) for the process of its arrival and an unlimited number of servers. The service time for requests have an exponential probability distribution. Parameter of the service depends on the state of the Markov chain nested over the renewal moments. It should be noted that these parameters do not change their values until the end of maintenance. Thus, the devices in the system under consideration are heterogeneous.This problem for the Queueing System $M|M|\infty$ in a Markov Random Enviroment was addressed in [6–8].

The objects of the study are the number of servers of each type being served in the stationary regime. Such a QS can be attributed to the class of non-homogeneous QS operating in a random environment.

Currently, a significant part of the information, telecommunication, and other systems operate in a changing environment. The impact of a random environment can be expressed, for example, in a change in the parameters of the functioning of the system. In this regard, questions arise about the stability of such systems to external influences. Therefore, the study of systems operating in a random environment is an urgent task. In various works devoted to the study of systems in Markov and semi-Markov random environments, various variants of the system's response to a change in the state of the external environment were considered in [9–11].

In this paper, we consider the case assuming that the service mode of claims does not change until they leave the system. The method of asymptotic analysis under the condition of equivalent growing of service times in the units of servers is applied for the study. This asymptotic condition means proportional growth of the average service times in both service units and it is taken from practice. The method of asymptotic analysis is implemented in the construction of a sequence of asymptotic of increasing order, in which the asymptotic of the first order determines the asymptotic mean value of the number of occupied servers. The second-order asymptotic allows to construct an approximation of the probability distribution of the number of occupied servers in the system. It is shown that this approximation coincides with the Gaussian distribution.

2. Markov renewal process

A renewal process is a generalization of a Poisson process that allows arbitrary waiting time between events. Its applications include such as planning for replacing worn-out machinery in a factory. A Markov renewal process is a generalization of a renewal process that the sequence of holding times is not independent and identically distributed. Their distributions depend on the states in a Markov chain. The Markov renewal processes were studied by Pyke [4, 5] in 1960s.

2.1. Mathematical model of the Markov renewal process

Consider a two-dimensional homogeneous Markov random process $\{\xi(n), \tau(n)\}$ with discrete time $n = 1, 2, 3, ...$, where $\xi(n)$ takes values from some discrete set $\xi(n) = k = 1, 2, 3, ...$ and $\tau(n)$ takes on non-negative values.
We denote

$$F(k_2, x; k_1, y) = P\{\xi(n+1) = k_2,\ \tau(n+1) < x | \xi(n) = k_1,\ \tau(n) = y\} = \\ = F(k_2, x; k_1) = P_{k_1 k_2} A_{k_2}(x).$$

A random stream of homogeneous events $t_1 < ... < t_n < t_{n+1} < ...$ will be called the Markov renewal flow or MR-flow given by the matrix of transition probabilities $\mathbf{P}$ and functions $A_k(x)$ distribution of interval lengths $\tau_{n+1} = t_{n+1} - t_n$, for which the equalities $\tau_{n+1} = \tau_n$ hold.

To study the MR-flow, we define the process $z(t)$ as the length of the interval from the time t to the time t_{n+1} of the next event in the considered flow and the process

$$k(t) = \xi(n), \quad t_n \leqslant t < t_{n+1},$$

that is, the process $k(t)$ on the interval $t_n \leqslant t < t_{n+1}$ retains the value that it received at the beginning of this interval and which coincides with the value $\xi(n)$ of the embedded Markov chain.

For a Markov renewal flow, the three-dimensional process $\{k(t), z(t), m(t)\}$ is Markov, therefore, for its probability distribution

$$P(k, z, m, t) = P\{k(t) = k, z(t) < z, m(t) = m\}$$

by the formula of total probability we obtain the equality

$$P(k, z - \Delta t, m, t + \Delta t) =$$
$$= P(k, z, m, t) - P(k, \Delta t, m, t) + \sum_{\nu} P(\nu, \Delta t, m-1, t) P_{\nu k} A_k(z) + o(\Delta t)$$

from which it follows that the probability distribution $P(k, z, m, t)$ is a solution to the Kolmogorov equations

$$\frac{\partial P(k,z,m,t)}{\partial t} =$$
$$= \frac{\partial P(k,z,m,t)}{\partial z} - \frac{\partial P(k,0,m,t)}{\partial z} + \sum_{\nu} \frac{\partial P(\nu,0,m-1,t)}{\partial z} P_{\nu k} A_k(z). \quad (1)$$

By defining the functions

$$H(k, z, u, t) = \sum_{m=0}^{\infty} e^{jum} P(k, z, m, t),$$

the equations (1) can be rewritten as

$$\frac{\partial H(k,z,u,t)}{\partial t} = \frac{\partial H(k,z,u,t)}{\partial z} - \frac{\partial H(k,0,u,t)}{\partial z} + \sum_{\nu} \frac{\partial H(\nu,0,u,t)}{\partial z} e^{ju} P_{\nu k} A_k(z).$$

The basic equation for a semi-Markov flow has the form

$$\frac{\partial \mathbf{h}(z,u,t)}{\partial t} = \frac{\partial \mathbf{h}(z,u,t)}{\partial z} + \frac{\partial \mathbf{h}(0,u,t)}{\partial z} \left(e^{ju} \mathbf{P} \mathbf{A}(z) - \mathbf{I} \right), \quad (2)$$

where $\mathbf{P}$ is the matrix of transition probabilities, $\mathbf{A}(z) = \operatorname{diag}[A_k(z)]$, $\mathbf{I}$ is identity diagonal matrix. To find its particular solution, we define the initial condition in the form

$$\mathbf{h}(z, u, 0) = \mathbf{r}(z),$$

where $\mathbf{r}(z)$ — stationary probability distribution of the values of a two-dimensional random process $\{k(t), z(t)\}$.

2.2. Finding the distribution $\mathbf{r}(z)$

Vector $\mathbf{r}(z)$ is a solution to the equation obtained from (2)

$$\frac{\partial \mathbf{r}(z)}{\partial z} + \frac{\partial \mathbf{r}(0)}{\partial z} (\mathbf{P} \mathbf{A}(z) - \mathbf{I}) = 0,$$

therefore it can be written as

$$\mathbf{r}(z) = \int_0^z \frac{\partial \mathbf{r}(0)}{\partial z} (\mathbf{I} - \mathbf{P} \mathbf{A}(x)) \, dx. \quad (3)$$

Since $r(k,z) = P\{k(t) = k,\ z(t) < z\}$ then $\mathbf{r} = \mathbf{r}(\infty)$. Therefore, we obtain

$$\mathbf{r} = \int_0^\infty \frac{\partial \mathbf{r}(0)}{\partial z} \left(\mathbf{I} - \mathbf{P}\mathbf{A}(x)\right) dx. \tag{4}$$

By virtue of the necessary condition for the convergence of the improper integral, we can write down the equality to zero of the integrand at $x \to \infty$, we obtain the system of equations

$$\frac{\partial \mathbf{r}(0)}{\partial z} \left(\mathbf{I} - \mathbf{P}\right) = 0 \tag{5}$$

for $\dfrac{\partial \mathbf{r}(0)}{\partial z}$, where $\mathbf{P} = \mathbf{A}(\infty)$.

Since the system (4) coincides with the system of Kolmogorov equations for the stationary probability distribution $\mathbf{r}$ of values of the embedded Markov chain, then

$$\frac{\partial \mathbf{r}(0)}{\partial z} = \lambda \mathbf{r}, \tag{6}$$

where λ is some multiplicative constant, the value of which is found as follows.

Substituting (6) into (4), we obtain

$$\mathbf{r} = \lambda \int_0^\infty \mathbf{r} \left(\mathbf{P} - \mathbf{A}(x)\right) dx.$$

Since $\mathbf{r}\mathbf{e} = 1$ then

$$\lambda = \frac{1}{\int_0^\infty \mathbf{r}\left(\mathbf{P} - \mathbf{A}(x)\right)\mathbf{e}\, dx} = \frac{1}{\int_0^\infty \left(1 - F(x)\right) dx}. \tag{7}$$

Equalities (7), (6) and (3) solve the problem of finding the probability distribution $\mathbf{r}(z)$.

3. Mathematical model

Consider a queueing system $MRP|M|\infty$ with an unlimited number of servers of different types, operating in a semi-Markov random environment (see the figure 1). Arrivals are determined as Markov renewal process, interarrival periods have cumulative distribution functions $A_1(x), A_2(x), \dots A_K(x)$ and the matrix of transition probabilities $\mathbf{P} = [p_{ij}]$, $i, j = 1, 2, \dots, K$ — embedded in the moments of occurrence of events Markov chains with a finite number of states $k(t) = 1, 2, \dots, K$. The service discipline is defined as follows: if the embedded Markov chain is in the state $k(t) = i$, then the incoming customer will be serviced on the i-th type server during a random time, exponentially distributed $F_i(x) = 1 - e^{-\mu_i x}$.

The problem is to study a multidimensional random process — numbers occupied servers of different types in the system at time t, which is denoted by $\mathbf{i}(t) = [i_1(t), i_2(t), \dots, i_K(t)]$. The process $\mathbf{i}(t)$ is not Markov. For clarity, consider the case when the external environment takes only 2 different states. We define a four-dimensional Markov random process $\{k(t), z(t), i_1(t), i_2(t)\}$, where $z(t)$ is the length of the interval from the time t to the time of the next event in the stream Markov renewal, $k(t)$ is a Markov chain embedded with respect to renewal times.

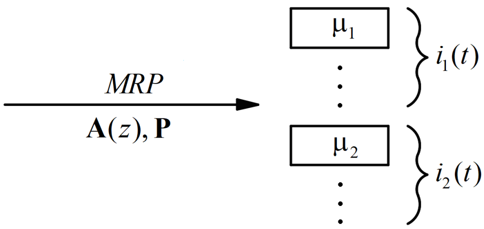

Figure 1. Queueing system $MRP|M|\infty$ in a semi-Markov random environment

For research, we will obtain some characteristics for the number of events occurring in the MR stream.

For the probability distribution

$$P(k, z, i_1, i_2, t) = P\{k(t) = k,\ z(t) < z,\ i_1(t) = i_1, i_2(t) = i_2\}$$

we write down the Kolmogorov system of differential equations:

$$\frac{\partial P(1,z,i_1,i_2,t)}{\partial t} = \frac{\partial P(1,z,i_1,i_2,t)}{\partial z} - \frac{\partial P(1,0,i_1,i_2,t)}{\partial z} - \\ -(i_1\mu_1 + i_2\mu_2)P(1,z,i_1,i_2,t) + \frac{\partial P(1,0,i_1-1,i_2,t)}{\partial z} p_{11} A_1(z) + \\ + \frac{\partial P(2,0,i_1-1,i_2,t)}{\partial z} p_{21} A_1(z) + P(1,z,i_1+1,i_2,t)(i_1+1)\mu_1 + \\ + P(1,z,i_1,i_2+1,t)(i_2+1)\mu_2,$$

$$\frac{\partial P(2,z,i_1,i_2,t)}{\partial t} = \frac{\partial P(2,z,i_1,i_2,t)}{\partial z} - \frac{\partial P(2,0,i_1,i_2,t)}{\partial z} - \\ -(i_1\mu_1 + i_2\mu_2)P(2,z,i_1,i_2,t) + \frac{\partial P(2,0,i_1,i_2-1,t)}{\partial z} p_{22} A_2(z) + \\ + \frac{\partial P(1,0,i_1,i_2-1,t)}{\partial z} p_{12} A_2(z) + P(2,z,i_1+1,i_2,t)(i_1+1)\mu_1 + \\ + P(2,z,i_1,i_2+1,t)(i_2+1)\mu_2.$$

For a stationary probability distribution, we write this system in the form

$$\frac{\partial P(1,z,i_1,i_2)}{\partial z} - \frac{\partial P(1,0,i_1,i_2)}{\partial z} - \\ -(i_1\mu_1 + i_2\mu_2)P(1,z,i_1,i_2) + \frac{\partial P(1,0,i_1-1,i_2)}{\partial z} p_{11} A_1(z) +$$

$$+\frac{\partial P(2,0,i_1-1,i_2)}{\partial z}p_{21}A_1(z)+P(1,z,i_1+1,i_2)(i_1+1)\mu_1+ \\ +P(1,z,i_1,i_2+1)(i_2+1)\mu_2=0,$$

$$\frac{\partial P(2,z,i_1,i_2)}{\partial z}-\frac{\partial P(2,0,i_1,i_2)}{\partial z}- \\ -(i_1\mu_1+i_2\mu_2)P(2,z,i_1,i_2)+\frac{\partial P(2,0,i_1,i_2-1)}{\partial z}p_{22}A_2(z)+ \\ +\frac{\partial P(1,0,i_1,i_2-1)}{\partial z}p_{12}A_2(z)+P(2,z,i_1+1,i_2)(i_1+1)\mu_1+ \\ +P(2,z,i_1,i_2+1)(i_2+1)\mu_2=0.$$

We introduce partial characteristic functions of the form

$$H(k,z,u_1,u_2)=\sum_{i_1=0}^{\infty}\sum_{i_2=0}^{\infty}e^{ju_1i_1}e^{ju_2i_2}P(k,z,i_1,i_2),\ \text{where } j=\sqrt{-1}.$$

Let us write the system of differential equations for the partial characteristic functions

$$\frac{\partial H(1,z,u_1,u_2)}{\partial z}-\frac{\partial H(1,0,u_1,u_2)}{\partial z}+ \\ +j\mu_1\left(1-e^{-ju_1}\right)\frac{\partial H(1,z,u_1,u_2)}{\partial u_1}+j\mu_2\left(1-e^{-ju_2}\right)\frac{\partial H(1,z,u_1,u_2)}{\partial u_2}+ \\ +\frac{\partial H(1,0,u_1,u_2)}{\partial z}e^{ju_1}p_{11}A_1(z)+\frac{\partial H(2,0,u_1,u_2)}{\partial z}e^{ju_1}p_{21}A_1(z)=0,$$

$$\frac{\partial H(2,z,u_1,u_2)}{\partial z}-\frac{\partial H(2,0,u_1,u_2)}{\partial z}+ \\ +j\mu_1\left(1-e^{-ju_1}\right)\frac{\partial H(2,z,u_1,u_2)}{\partial u_1}+j\mu_2\left(1-e^{-ju_2}\right)\frac{\partial H(2,z,u_1,u_2)}{\partial u_2}+ \\ +\frac{\partial H(1,0,u_1,u_2)}{\partial z}e^{ju_2}p_{12}A_2(z)+\frac{\partial H(2,0,u_1,u_2)}{\partial z}e^{ju_2}p_{22}A_2(z)=0$$

with initial conditions

$$H(k,z,0,0)=r(k,z).$$

In vector-matrix form, this system will take the form

$$\frac{\partial \mathbf{h}(z,u_1,u_2)}{\partial z}+\frac{\partial \mathbf{h}(0,u_1,u_2)}{\partial z}\left(\mathbf{PA}(z)\mathbf{B}(\mathbf{u})-\mathbf{I}\right)+ \\ +j\mu_1\left(1-e^{-ju_1}\right)\frac{\partial \mathbf{h}(z,u_1,u_2)}{\partial u_1}+j\mu_2\left(1-e^{-ju_2}\right)\frac{\partial \mathbf{h}(z,u_1,u_2)}{\partial u_2}=0, \quad (8)$$

with initial conditions

$$\mathbf{h}(z,0,0)=\mathbf{r}(z),$$

where

$$\mathbf{h}(z,u_1,u_2)=\left[H(1,z,u_1,u_2),H(2,z,u_1,u_2)\right],$$

$$\mathbf{B}(\mathbf{u})=\begin{bmatrix} e^{ju_1} & 0 \\ 0 & e^{ju_2}\end{bmatrix}, \quad \mathbf{I}=\begin{bmatrix} 1 & 0 \\ 0 & 1\end{bmatrix}.$$

The resulting system of equations (8) is the main one for further research. Since it is not possible to find an explicit form of a solution to the problem (8), 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.

4. Asymptotic analysis of the first order

We denote $\mu_1=\epsilon$, $\mu_2=q\epsilon$, $q=\text{const}$ (ϵ is an infinitesimal quantity). Then we can write the asymptotic condition of equivalent growing of service times in the units of servers in the form $\mu_1,\mu_2\to 0$. In (8) we perform the replacements

$$\mathbf{h}(z,u_1,u_2)=\mathbf{f}(z,w_1,w_2,\epsilon),\ u_1=\epsilon w_1,\ u_2=\epsilon w_2,$$

we obtain the matrix equation for $\mathbf{f}(z,w_1,w_2,\epsilon)$

$$\begin{gathered}\frac{\partial\mathbf{f}(z,w_1,w_2,\epsilon)}{\partial z}+\frac{\partial\mathbf{f}(0,w_1,w_2,\epsilon)}{\partial z}\left(\mathbf{PA}(z)\mathbf{B}(\mathbf{u},\epsilon)-\mathbf{I}\right)+\\ +j\left(1-e^{-j\epsilon w_1}\right)\frac{\partial\mathbf{f}(z,w_1,w_2,\epsilon)}{\partial w_1}+jq\left(1-e^{-j\epsilon w_2}\right)\frac{\partial\mathbf{f}(z,w_1,w_2,\epsilon)}{\partial w_2}=0,\end{gathered}\tag{9}$$

Theorem 1. *The limiting solution for* $\epsilon\to 0$ *to the equation* (9) $\mathbf{f}(z,w_1,w_2,\epsilon)$ *has the form*

$$\mathbf{f}(z,w_1,w_2,\epsilon)=\mathbf{r}(z)\exp\left\{j\lambda\left(r_1w_1+\frac{r_2w_2}{q}\right)\right\},\tag{10}$$

where $\mathbf{r}(z)=[r_1(z),r_2(z)]$ *is the vector of the probability distribution of the values of the embedded Markov chain,* $\mathbf{r}=[r_1,r_2]$ *is vector of stationary probability distribution of the values of the embedded Markov chain.*

Proof. In the equation (9) we carry out the passage to the limit for $\epsilon\to 0$, we obtain that $\mathbf{f}(z,w_1,w_2)$ is a solution to the equation

$$\frac{\partial\mathbf{f}(z,w_1,w_2)}{\partial z}+\frac{\partial\mathbf{f}(0,w_1,w_2)}{\partial z}\left(\mathbf{PA}(z)-\mathbf{I}\right)=0,$$

which defines the vector function $\mathbf{r}(z)$, therefore we will seek the function $\mathbf{f}(z, w_1, w_2, \epsilon)$ in the form of the expansion

$$\mathbf{f}(z, w_1, w_2, \epsilon) = \mathbf{r}(z)\Phi(w_1, w_2) + o(\epsilon). \tag{11}$$

In the equation (9) we carry out the passage to the limit as $z \to \infty$, multiply this equation by the unit column vector $\mathbf{e}$, expand the exponents in a Maclaurin series up to the first order. In the resulting expression, we substitute the expansion (11), divide by ϵ and carry out the passage to the limit at $\epsilon \to 0$, we obtain the equation for the function $\Phi(w_1, w_2)$

$$w_1 \frac{\partial \Phi(w_1, w_2)}{\partial w_1} + q w_2 \frac{\partial \Phi(w_1, w_2)}{\partial w_2} = j \frac{\partial \mathbf{r}(0)}{\partial z} \mathbf{P} \mathbf{W} \mathbf{e} \Phi(w_1, w_2),$$

where $\frac{\partial \mathbf{r}(0)}{\partial z} = \lambda \mathbf{r}$, $\mathbf{r}\mathbf{P} = \mathbf{r}$, $\mathbf{r}\mathbf{e} = 1$, $\lambda = \frac{1}{\int_0^\infty (1 - \mathbf{r}\mathbf{A}(x)\mathbf{e})\, dx}$, $\mathbf{W} = \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix}$.
The solution will have the following form

$$\Phi(w_1, w_2) = \exp\left\{ j\lambda \left(r_1 w_1 + \frac{r_2 w_2}{q} \right) \right\}$$

Substituting the obtained solution into (11), we get (10).

The theorem is proved. □

By substitution and equality (3), we write down the approximate (asymptotic) equality

$$\mathbf{h}(z, u_1, u_2) \approx \mathbf{f}(z, w_1, w_2) = \mathbf{r}(z) \exp\left\{ j\lambda \left(r_1 w_1 + \frac{r_2 w_2}{q} \right) \right\} = \\ = \mathbf{r}(z) \exp\left\{ j\lambda \left(\frac{r_1 u_1}{\mu_1} + \frac{r_2 u_2}{\mu_2} \right) \right\}.$$

Let us define the characteristic of the process $\{i_1(t), i_2(t)\}$ in the stationary mode

$$h(u_1, u_2) = \exp\left\{ j\lambda \left(\frac{r_1 u_1}{\mu_1} + \frac{r_2 u_2}{\mu_2} \right) \right\},$$

which we will call the first-order asymptotics of the characteristic functions of the number of occupied servers in the system.

5. Asymptotic analysis of the second order

In the equation (8) we replace

$$\mathbf{h}(z, u_1, u_2) = \mathbf{h}_2(z, u_1, u_2) \exp\left\{ j\lambda \left(\frac{r_1 u_1}{\mu_1} + \frac{r_2 u_2}{\mu_2} \right) \right\},$$

we obtain the equation for $\mathbf{h}_2(z,u_1,u_2)$

$$\frac{\partial \mathbf{h}_2(z,u_1,u_2)}{\partial z}+\frac{\partial \mathbf{h}_2(0,u_1,u_2)}{\partial z}\left(\mathbf{PA}(z)\mathbf{B}(\mathbf{u})-\mathbf{I}\right)+$$
$$+j\mu_1\left(1-e^{-ju_1}\right)\frac{\partial \mathbf{h}_2(z,u_1,u_2)}{\partial u_1}+j\mu_2\left(1-e^{-ju_2}\right)\frac{\partial \mathbf{h}_2(z,u_1,u_2)}{\partial u_2}- \tag{12}$$
$$-\lambda r_1\left(1-e^{-ju_1}\right)\mathbf{h}_2(z,u_1,u_2)-\lambda r_2\left(1-e^{-ju_2}\right)\mathbf{h}_2(z,u_1,u_2)=0.$$

We denote $\mu_1=\epsilon^2$, $\mu_2=q\epsilon^2$, in (12) we replace

$$\mathbf{h}_2(z,u_1,u_2)=\mathbf{f}_2(z,w_1,w_2,\epsilon),\ u_1=\epsilon w_1,\ u_2=\epsilon w_2,$$

we obtain the equation for $\mathbf{f}_2(z,w_1,w_2,\epsilon)$

$$\frac{\partial \mathbf{f}_2(z,w_1,w_2,\epsilon)}{\partial z}+\frac{\partial \mathbf{f}_2(0,w_1,w_2,\epsilon)}{\partial z}\left(\mathbf{PA}(z)\mathbf{B}(\mathbf{w},\epsilon)-\mathbf{I}\right)+$$
$$+j\epsilon\left(1-e^{-j\epsilon w_1}\right)\frac{\partial \mathbf{f}_2(z,w_1,w_2,\epsilon)}{\partial w_1}+j\epsilon q\left(1-e^{-j\epsilon w_2}\right)\frac{\partial \mathbf{f}_2(z,w_1,w_2,\epsilon)}{\partial w_2}-$$
$$-\lambda r_1\left(1-e^{-j\epsilon w_1}\right)\mathbf{f}_2(z,w_1,w_2,\epsilon)-\lambda r_2\left(1-e^{-j\epsilon w_2}\right)\mathbf{f}_2(z,w_1,w_2,\epsilon)=0. \tag{13}$$

Theorem 2. *The limiting solution for $\epsilon\to 0$ to the equation* (13) $\mathbf{f}_2(z,w_1,w_2)$ *has the form*

$$\mathbf{f}_2(z,w_1,w_2)=\mathbf{r}(z)\exp\left\{\frac{j^2}{2}\left(\lambda\left(r_1w_1^2+r_2\frac{w_2^2}{q}\right)+\right.\right.$$
$$\left.\left.+\kappa\left(r_1^2w_1^2+4r_1r_2\frac{w_1w_2}{q+1}+r_2^2\frac{w_2^2}{q}\right)\right)\right\}, \tag{14}$$

where $\kappa=\lambda^2\int_0^\infty\left(\mathbf{rA}(x)-\mathbf{r}(x)\right)\mathbf{e}dx.$

Proof. We will obtain the solution of the equation (14) in the following form

$$\mathbf{f}_2(z,w_1,w_2,\epsilon)=\Phi(w_1,w_2)\left(\mathbf{r}(z)+j\epsilon(r_1w_1+r_2w_2)\mathbf{f}_2(z)\right)+o^2(\epsilon), \tag{15}$$

where $\mathbf{f}_2(z)$ satisfies the condition $\mathbf{f}_2(\infty)\mathbf{e}=0$. Substitute (15) into (13) and expand the exponents in a series up to the first order. Considering that

$$\frac{\partial \mathbf{r}(z)}{\partial z}+\frac{\partial \mathbf{r}(0)}{\partial z}\left(\mathbf{PA}(z)-\mathbf{I}\right)=0,$$

we obtain the equation for finding the function $\mathbf{f}_2(z)$

$$\mathbf{e}\frac{\partial \mathbf{f}_2(z)}{\partial z}-\lambda\mathbf{e}\mathbf{r}(z)+\mathbf{e}\frac{\partial \mathbf{f}_2(0)}{\partial z}\left(\mathbf{PA}(z)-\mathbf{I}\right)+\lambda\mathbf{A}(z)=0. \tag{16}$$

From the equation (16) we find that

$$\frac{\partial \mathbf{f}_2(0)}{\partial z} = \kappa \mathbf{r}, \quad \text{where} \quad \kappa = \lambda^2 \int_0^\infty (\mathbf{r}\mathbf{A}(x) - \mathbf{r}(x))\, \mathbf{e} dx.$$

Substitute (15) into (13) and expand the exponents in a series up to the second order. Multiply by $\mathbf{e}$ and perform the passage to the limit $z \to \infty$, we obtain the equation for finding the function $\Phi(w_1, w_2)$

$$w_1 \frac{\partial \Phi(w_1,w_2)}{\partial w_1} + w_2 q \frac{\partial \Phi(w_1,w_2)}{\partial w_2} = \\ = \Phi(w_1,w_2)\left(-\lambda\left(r_1 w_1^2 + r_2 w_2^2\right) - \kappa\left(r_1 w_1 + r_2 w_2\right)^2\right). \quad (17)$$

The solution of the equation (17) has the form

$$\Phi(w_1,w_2) = \\ = \exp\left\{\frac{j^2}{2}\left(\lambda\left(r_1 w_1^2 + r_2 \frac{w_2^2}{q}\right) + \kappa\left(r_1^2 w_1^2 + 4 r_1 r_2 \frac{w_1 w_2}{q+1} + r_2^2 \frac{w_2^2}{q}\right)\right)\right\} \quad (18)$$

Substituting the solution (18) into (15) and performing the passage to the limit $\epsilon \to 0$, we obtain (14).

The theorem is proved. □

Due to the change, as well as the equality (14) for the function $\mathbf{h}_2(z, u_1, u_2)$ we can write down the approximate (asymptotic) equality

$$\mathbf{h}_2(z,u_1,u_2) \approx \mathbf{f}_2(z,w_1,w_2) = \\ = \mathbf{r}(z) \exp\left\{\frac{j^2}{2}\left(\lambda\left(r_1 \frac{u_1^2}{\mu_1} + r_2 \frac{u_2^2}{\mu_2}\right) + \right.\right. \\ \left.\left. +\kappa\left(r_1^2 \frac{u_1^2}{\mu_1} + 4 r_1 r_2 \frac{u_1 u_2}{\mu_1+\mu_2} + r_2^2 \frac{u_2^2}{\mu_2}\right)\right)\right\}.$$

Thus, the characteristic function of the number of occupied servers in the system under consideration has the form

$$h_2(u_1,u_2) = \exp\left\{j\lambda\left(\frac{r_1 u_1}{\mu_1} + \frac{r_2 u_2}{\mu_2}\right) + \frac{j^2}{2}\left[\lambda\left(r_1 \frac{u_1^2}{\mu_1} + r_2 \frac{u_2^2}{\mu_2}\right) + \right.\right. \\ \left.\left. +\kappa\left(r_1^2 \frac{u_1^2}{\mu_1} + 4 r_1 r_2 \frac{u_1 u_2}{\mu_1+\mu_2} + r_2^2 \frac{u_2^2}{\mu_2}\right)\right]\right\}. \quad (19)$$

6. Numerical example

Let us consider a numerical example where we can illustrate the accuracy of approximating formula (19). Consider queueing system with MRP arrivals, where the Markov renewal process is given by matrices

$$\mathbf{P} = \begin{bmatrix} 0.3 & 0.7 \\ 0.6 & 0.4 \end{bmatrix}, \quad \mathbf{A}(x) = \mathrm{diag}\{A_1(x), A_2(x)\}.$$

Here $A_1(x)$ and $A_2(x)$ are gamma distribution cdf-s with shape and rate parameters α and β that have the following values:

$$\alpha_1 = 0.5,\ \beta_1 = 0.25, \quad \alpha_2 = 1.5,\ \beta_2 = 1.5.$$

The service times are exponentially distributed with service rates

$$\mu_1 = 1 \cdot \varepsilon, \quad \mu_2 = 2 \cdot \varepsilon$$

for the the first and the second types of arrivals respectively. Here parameter ε will be varied to establish the accuracy of approximation (19) accordingly to the asymptotic condition $\varepsilon \to 0$.

To establish the accuracy of the approximation, we use its comparison with the results of simulation modeling of the corresponding system. For the error estimation (difference between the results), we use the Kolmogorov distance

$$\Delta = \max_{i_1, i_2 \in [0,\infty)} \left| F_{\text{approx}}(i_1, i_2) - F_{\text{sim}}(i_1, i_2) \right|,$$

where $F_{\text{approx}}(i_1, i_2)$ is a cdf of Gaussian distribution (19) and $F_{\text{sim}}(i_1, i_2)$ is a cdf built basing on the results of the simulation. The results of the comparison is presented in the table 1. We see that the Kolmogorov distance decreases with decreasing of parameter ε, so, approximation (19) becomes more accurate for small values of this parameter.

Table 1

Kolmogorov distance Δ between the approximation and distribution based on the simulation results for various values of asymptotic parameter ε

ε	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001
Δ	0.1137	0.0501	0.0371	0.0323	0.0253	0.0226	0.0197

For example, if we suppose that error $\Delta \leqslant 0.05$ means that the approximation is accurate enough then we can conclude that for the considered example, Gaussian approximation (19) is applicable for values $\varepsilon < 0.05$.

7. Conclusions

In this paper, the method of asymptotic analysis is used to study a mathematical model of the $MR|M|\infty$ system functioning under the condition of a changing environment. The case is considered when a semi-Markov random environment has 2 different states. It is proved that the asymptotic characteristic function of the number of occupied servers of each type in the considered system is Gaussian with the vector of mathematical expectations

$$\mathbf{a} = \left[\lambda\frac{r_1}{\mu_1}, \lambda\frac{r_2}{\mu_2}\right]$$

and the covariance matrix

$$\mathbf{K} = \begin{bmatrix} \lambda\frac{r_1}{\mu_1} + \kappa\frac{r_1^2}{\mu_1} & 4\kappa\frac{r_1 r_2}{\mu_1+\mu_2} \\ 4\kappa\frac{r_1 r_2}{\mu_1+\mu_2} & \lambda\frac{r_2}{\mu_2} + \kappa\frac{r_2^2}{\mu_2} \end{bmatrix}.$$

References

[1] A. Dudin, V. Klimenok, and V. Vishnevsky, *The Theory of Queuing Systems with Correlated Flow*. Springer Nature, 2020. DOI: 10.1007/978-3-030-32072-0.

[2] V. K. Malinovskii, "Asymptotic expansions in the central limit theorem for recurrent Markov renewal processes," *Theory of Probability & Its Applications*, vol. 51, no. 3, pp. 523–526, 1987. DOI: 10.1137/1131073.

[3] Y. Lim, S. Hur, and J. Seung, "Departure process of a single server queueing system with Markov renewal input and general service time distribution," *Computers & Industrial Engineering*, vol. 51, no. 3, pp. 519–525, 2006. DOI: 10.1016/j.cie.2006.08.011.

[4] R. Pyke, "Markov renewal processes: definitions and preliminary properties," *Ann. Math. Statist.*, vol. 32, pp. 1231–1242, 1961. DOI: 10.1214/aoms/1177704863.

[5] R. Pyke and R. Schaufele, "Stationary measures for Markov renewal processes," *Ann. Math. Statist.*, vol. 37, pp. 1439–1462, 1966. DOI: 10.1214/aoms/1177699138.

[6] J. Sztrik and D. Kouvatsos, "Asymptotic analysis of a heterogeneous multiprocessor system in a randomly changing environment," *IEEE Transactions on Software Engineering*, vol. 17, no. 10, pp. 1069–1075, 1991. DOI: 10.1109/32.99194.

[7] E. P. Polin, S. P. Moiseeva, and S. V. Rozhkova, "Asymptotic analysis of heterogeneous queueing system $M|M|\infty$ in a Markov random enviroment [Asimptoticheskiy analiz neodnorodnoy sistemy massovogo obsluzhivaniya $M|M|\infty$ v markovskoy sluchaynoy srede]," *Tomsk State University Jounal of Control and Computer Science [Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaja tehnika i informatika]*, vol. 47, pp. 75–83, 2019, in Russian. DOI: 10.17223/19988605/47/9.

[8] E. P. Polin, S. P. Moiseeva, and A. N. Moiseev, “Heterogeneous queueing system $MR(S)|M(s)|\infty$ with service parameters depending on the state of the underlying Markov chain [Analiz veroyatnostnykh kharakteristik geterogennoy SMO vida $MR(S)/M(S)/\infty$ s parametrami obsluzhivaniya, zavisyashchimi ot sostoyaniya vlozhennoy tsepi Markova],” *Saratov University News. New Series. Series Mathematics. Mechanics. Informatics [Izv. Saratov Univ. (N. S.), Ser. Math. Mech. Inform.]*, vol. 20, no. 3, pp. 388–399, 2020, in Russian. DOI: 10.18500/1816-9791-2020-20-3-388-399.

[9] B. D’Auria, “$M|M|\infty$ queues in semi-Markovian random environment,” *Queueing Systems*, vol. 58, pp. 221–237, 2008. DOI: 10.1007/s11134-008-9068-7.

[10] H. M. Jansen, “A large deviations principle for infinite-server queues in a random environment,” *Queueing Systems*, vol. 82, pp. 199–235, 2016. DOI: 10.1007/s11134-015-9470-x.

[11] J. Blom, M. Mandjes, and H. Thorsdottir, “Time-scaling limits for Markov-modulated infinite-server queues,” *Stochastic Models*, vol. 29, pp. 112–127, 2012. DOI: 10.1080/15326349.2013.750536.

For citation:

E. P. Polin, S. P. Moiseeva, A. N. Moiseev, Heterogeneous queueing system with Markov renewal arrivals and service times dependent on states of arrival process, Discrete and Continuous Models and Applied Computational Science 31 (2) (2023) 105–119. DOI: 10.22363/2658-4670-2023-31-2-105-119.

Information about the authors:

Polin, Evgeny P. — Assistant of Department of Probability Theory and Mathematical Statistics, National Research Tomsk State University (e-mail: polin_evgeny@mail.ru, phone: +7(923)4480077, ORCID: https://orcid.org/0000-0002-0250-2368)

Moiseeva, Svetlana P. — Doctor in Physics and Mathematics, Professor at Department of Probability Theory and Mathematical Statistics, National Research Tomsk State University (e-mail: smoiseeva@mail.ru, ORCID: https://orcid.org/0000-0001-9285-1555, Scopus Author ID: 56436490300)

Moiseev, Alexander N. — Doctor in Physics and Mathematics, Head of the Department of Software Engineering, National Research Tomsk State University (e-mail: moiseev.tsu@gmail.com, ORCID: https://orcid.org/0000-0003-2369-452X, ResearcherID: N-7189-2014, Scopus Author ID: 55646953800)

УДК 519.872
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-2-105-119

Гетерогенная система массового обслуживания с входящим потоком марковского восстановления и временем обслуживания, зависящими от состояний вложенной цепи Маркова

Е. П. Полин[1,2], **С. П. Моисеева**[1], **А. Н. Моисеев**[1]

[1] *Национальный исследовательский Томский государственный университет, пр. Ленина, д. 36, Томск, 634050, Россия*
[2] *Национальный исследовательский Томский политехнический университет, пр. Ленина, д. 30, Томск, 634050, Россия*

Аннотация. В работе рассматривается гетерогенная система массового обслуживания с входящим потоком марковского восстановления и неограниченным числом серверов. Время обслуживания запросов на серверах является положительной случайной величиной с экспоненциальным распределением вероятностей. Параметры обслуживания зависят от состояния цепи Маркова в моменты восстановления. Следует отметить, что эти параметры не меняют своих значений до окончания обслуживания. Таким образом, устройства в рассматриваемой системе являются неоднородными (гетерогенными). Объектом исследования становится многомерный случайный процесс — количество серверов каждого типа, обслуживаемых с разной интенсивностью в стационарном режиме. Для исследования применён метод асимптотического анализа при условии эквивалентно долгого времени обслуживания. Метод асимптотического анализа реализуется при построении последовательности асимптотик возрастающего порядка, в которой асимптотика первого порядка определяет асимптотическое среднее значение числа занятых серверов. Асимптотика второго порядка позволяет построить гауссовскую аппроксимацию распределения вероятностей числа занятых серверов в системе.

Ключевые слова: система массового обслуживания, случайная среда, поток марковского восстановления, метод асимптотического анализа