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

2024, 32 (3) 271–282
http://journals.rudn.ru/miph

Computer science

Research article

UDC 519.21
DOI: 10.22363/2658-4670-2024-32-3-271–282

EDN: BAUGIT

Analysis of a queuing system of a single capacity with phase-type distributions and queue updating

Sergey I. Matyushenko, Konstantin E. Samouylov, Nikolai Yu. Gritsenko

RUDN University, 6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation

(received: September 14, 2023; revised: September 30, 2023; accepted: October 1, 2023)

Abstract. In this paper, we study a queuing system with a single-capacity storage device and queue updating. An update is understood as the following mechanism: an application that enters the system and finds another application in the drive destroys it, taking its place in the drive. It should be noted that systems with one or another update mechanism have long attracted the attention of researchers, since they have important applied significance. Recently, interest in systems of this kind has grown in connection with the tasks of assessing and managing the age of information. A system with a queue update mechanism similar to the one we are considering has already been studied earlier in the works of other authors. However, in these works we were talking about the simplest version of the system with Poisson flow and exponential maintenance. In this paper, we consider a phase-type flow and maintenance system. As a result of our research, we developed a recurrent matrix algorithm for calculating the stationary distribution of states of a Markov process describing the stochastic behavior of the system in question, and obtained expressions for the main indicators of its performance.

Key words and phrases: queuing system, phase-type distribution, queue update mechanism

For citation: Matyushenko, S. I., Samouylov, K. E., Gritsenko, N. Y. Analysis of a queuing system of a single capacity with phase-type distributions and queue updating. *Discrete and Continuous Models and Applied Computational Science* **32** (3), 271–282. doi: 10.22363/2658-4670-2024-32-3-271–282. edn: BAUGIT (2024).

1. Introduction

The tasks related to the assessment and management of information age, which were initiated in [1–13], revived interest in the study of systems with various kinds of updating mechanisms. One of these systems is a system with a queue update mechanism [14], the essence of which is that an application entering the system and finding another application in the drive “kills” it and takes its place in the drive. This ensures that the information transmitted by the application is updated as quickly as possible, which is extremely important for real technical systems implementing service complexes for which the time factor plays the most important role. A system with this queue update mechanism was considered in [7, 15, 16]. However, the authors of these papers considered a system with Poisson flow and exponential maintenance, which, according to Kendall’s notation, is usually encoded as $M/M/1/1$. It is known that such models of queuing systems allow us to obtain only rough estimates of the characteristics of real technical systems, since single-parameter distributions do not make it possible to take into account all the features of the protocols of modern dispatch control and data collection systems, random multiple access from several remote sender nodes, multistep information transmission routes, etc. Therefore, in this paper we have followed the path of generalization, assuming that the time intervals between the receipt of applications and the duration of their service are random variables with phase-type distributions. This circumstance will allow us to subsequently use the universality of phase-type distributions to build more accurate models of real technical systems.

2. Description of the model

A single-line queuing system (QS) with a single-capacity storage device is considered, which receives a recurrent flow of applications with a phase-type distribution function (DF) $A(t)$:

$$A(t) = 1 - \boldsymbol{\alpha}^T e^{\boldsymbol{\Lambda} t}\mathbf{1}, \quad t \geq 0, \quad \boldsymbol{\alpha}^T\mathbf{1} = 1,$$

admitting an irreducible PH representation $(\boldsymbol{\alpha}, \boldsymbol{\Lambda})$ of order l [17]. The duration of the application service has a phase type DF $B(t)$ with an irreducible PH representation of the order m:

$$B(t) = 1 - \boldsymbol{\beta}^T e^{\boldsymbol{M} t}\mathbf{1}, \quad t \geq 0, \quad \boldsymbol{\beta}^T\mathbf{1} = 1.$$

Consider the QS with queue update. This means that an application that enters the system and finds the drive busy displaces the application from the drive and takes its place. The repressed application leaves the system and does not return to it anymore. In accordance with Kendall's notation, the QS in question will be encoded as $PH/PH/1/1$ with queue update (Fig. 1).

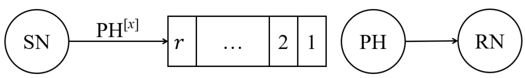

Figure 1. QS $PH/PH/1/1$ with queue update

3. Mathematical model

Based on the probabilistic interpretation of the PH distributions, the functioning of the QS under consideration is described by a homogeneous Markov process (MP) $\{X(t),\ t \geq 0\}$ over the state space

$$\mathscr{X} = \bigcup_{k=0}^{2} \mathscr{X}_k,$$

where $\mathscr{X}_0 = \left\{(i,0), i = \overline{1,l}\right\}$, $\mathscr{X}_k = \left\{(k,i,j), i = \overline{1,l}, j = \overline{1,m}\right\}$, $k = 1,2$.

Here $X(t) = (i,0)$ if at time t the system is empty and the process of generating a new application is in phase i. In turn, equality $X(t) = (k,i,j)$ means that in the system of k applications, the process of generating a new application is in phase i, and maintenance is in phase j.

It follows from the irreducibility of PH distributions [17] that all states of the process $\{X(t),\ t \geq 0\}$ are reported, the process itself is ergodic, and the limiting probabilities

$$p_{i0} = \lim_{t\to\infty} P\{X(t) = (i,0)\},$$
$$p_{ikj} = \lim_{t\to\infty} P\{X(t) = (i,k,j)\},$$

$$\begin{tikzcd} \mathscr{X}_1 \arrow[d, "(\boldsymbol{I} \otimes \boldsymbol{\mu})\varDelta"] \\ \mathscr{X}_0 \arrow[loop left, "{[\boldsymbol{\Lambda} \oplus \boldsymbol{M} - diag(\boldsymbol{\Lambda} \oplus \boldsymbol{M})]\varDelta}"] \arrow[loop right, "{\boldsymbol{I} - [-diag(\boldsymbol{\Lambda} \oplus \boldsymbol{M})]\varDelta}"] \end{tikzcd}$$

Figure 2. Diagram of transitions of MP $X(t)$ to states $\mathscr{X}_0$ for $\varDelta$

strictly positive, independent of the initial distribution, and consistent with stationary probabilities.

Let's introduce vectors:

$$\boldsymbol{p}_0^T = (p_{10}, \dots, p_{l0}),$$
$$\boldsymbol{p}_k^T = (p_{1k1}, \dots, p_{1km}, \dots, p_{lk1}, \dots, p_{lkm}), \quad k = 1, 2.$$

Stationary probabilities $\{\boldsymbol{p}_k,\ k = 0, 1, 2\}$ are the only solution to the system of equilibrium equations (SEE):

$$\mathbf{0}^T = \boldsymbol{p}_0^T \boldsymbol{\Lambda} + \boldsymbol{p}_1^T (\boldsymbol{I} \otimes \boldsymbol{\mu}), \tag{1}$$

$$\mathbf{0}^T = \boldsymbol{p}_0^T (\boldsymbol{\lambda} \boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) + \boldsymbol{p}_1^T (\boldsymbol{\Lambda} \oplus \boldsymbol{M}) + \boldsymbol{p}_2^T (\boldsymbol{I} \otimes \boldsymbol{\mu} \boldsymbol{\beta}^T), \tag{2}$$

$$\mathbf{0}^T = \boldsymbol{p}_1^T (\boldsymbol{\lambda} \boldsymbol{\alpha}^T \otimes \boldsymbol{I}) + \boldsymbol{p}_2^T (\boldsymbol{\Lambda} \oplus \boldsymbol{M} + \boldsymbol{\lambda} \boldsymbol{\alpha}^T \otimes \boldsymbol{I}), \tag{3}$$

with the condition of normalization

$$\sum_{k=0}^{2} \boldsymbol{p}_k^T \mathbf{1} = 1. \tag{4}$$

Hereafter $\boldsymbol{\lambda} = -\boldsymbol{\Lambda}\mathbf{1}$, $\boldsymbol{\mu} = -\boldsymbol{M}\mathbf{1}$, the sign $\otimes$ means the Kronecker product, and the sign $\oplus$ means the Kronecker sum of matrices.

Let's explain the conclusion of the SEE (1)–(3) using the MP transition scheme $X(t)$ on the interval $(t, t + \varDelta)$, where $\varDelta$ is a "small" time interval.

The subset of states $\mathscr{X}_0$ can be accessed from the subset $\mathscr{X}_1$ due to the end of the application service on the device with an intensity characterized by the vector $\boldsymbol{\mu}$ (Fig. 2). In addition to this transition, Fig. 2 shows two more situations in which the process does not go beyond the subset $\mathscr{X}_0$: during $\varDelta$ there will be no change in the generation phases, or vice versa — the generation phases change.

The first situation is reflected by the elements of the main diagonal of the matrix $\boldsymbol{\Lambda}$, taken with the opposite sign, which we will denote $diag(\boldsymbol{\Lambda})$, and the second is the non-diagonal elements of this matrix, which we will denote $\boldsymbol{\Lambda} - diag(\boldsymbol{\Lambda})$. In a subset of states $\mathscr{X}_1$, during $\varDelta$, it is possible to get from the subset $\mathscr{X}_0$ due to the receipt of a new application, which occurs with an intensity characterized by the vector $\boldsymbol{\lambda}$. At the same time, we take into account that the generation of the next application immediately begins, and the choice of the initial phase occurs in accordance with the probability vector $\boldsymbol{\alpha}$. In addition, the subset $\mathscr{X}_1$ during $\varDelta$ can be accessed from the subset $\mathscr{X}_2$ by the end of the service with an intensity determined by the vector $\boldsymbol{\mu}$. In this case, the initial phase of the service of the next application is selected in accordance with the probabilistic vector $\boldsymbol{\beta}$. The third possibility is to remain in this subset due to the fact that the passage of the current generation and maintenance phases will not be completed during $\varDelta$. This possibility is reflected by the intensities equal to the elements of the main diagonal of the matrix $\boldsymbol{\Lambda} \oplus \boldsymbol{M}$, taken with the opposite sign. Either due to Δ there will be changes in the phases of generation or maintenance, which reflect intensities equal to the non-diagonal elements of the matrix $\boldsymbol{\Lambda} \oplus \boldsymbol{M}$ (Fig. 3).

$$\begin{tikzcd}
\mathscr{X}_2 \arrow[d, "(\boldsymbol{I} \otimes \boldsymbol{\mu})\Delta"] \\
\mathscr{X}_1 \arrow[loop left, "{[\boldsymbol{\Lambda} \oplus \boldsymbol{M} - diag(\boldsymbol{\Lambda} \oplus \boldsymbol{M})]\Delta}"] \arrow[loop right, "{\boldsymbol{I} - [-diag(\boldsymbol{\Lambda} \oplus \boldsymbol{M})]\Delta}"] \\
\mathscr{X}_0 \arrow[u, "(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T)\Delta"']
\end{tikzcd}$$

Figure 3. Diagram of transitions of MP $X(t)$ to states $\mathscr{X}_1$ for Δ

Now let's explain the derivation of equation (3). In the subset $\mathscr{X}_2$ for the time Δ you can get in by receiving a new application as from a subset $\mathscr{X}_1$, while remaining inside a subset of $\mathscr{X}_2$. In both cases, the given transition occurs with intensities equal to the corresponding coordinates of the vector $\boldsymbol{\lambda}$, and ends with the choice of a new generation phase in accordance with the vector $\boldsymbol{\alpha}$ (Fig. 4). In addition, as in the previous case, there are two possibilities to remain in the subset $\mathscr{X}_2$: due to the fact that there are no changes in Δ it will happen, or only a change of generation or maintenance phases will occur.

$$\begin{tikzcd}
\mathscr{X}_2 \arrow[loop left, "{[\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{I} + \boldsymbol{\Lambda} \oplus \boldsymbol{M} - diag(\boldsymbol{\Lambda} \oplus \boldsymbol{M})]\Delta}"] \arrow[loop right, "{\boldsymbol{I} - [-diag(\boldsymbol{\Lambda} \oplus \boldsymbol{M})]\Delta}"] \\
\mathscr{X}_1 \arrow[u, "(\boldsymbol{\lambda}\boldsymbol{\alpha}^T \otimes \boldsymbol{I})"']
\end{tikzcd}$$

Figure 4. Diagram of transitions of MP $X(t)$ to states $\mathscr{X}_1$ for Δ

4. Solution of the SEE

Let's move on to solving the SEE (1)–(4), noting that we are interested not in numerical, but in the analytical solution of a system of equations, that is, analytical expressions in explicit form both to determine the performance indicators of the system itself and to obtain similar results for numerous special cases of the system under consideration. Before proceeding to the direct solution of the SEE, we introduce notation and prove the validity of a number of auxiliary correlations.

Let's introduce the matrices

$$\boldsymbol{V}_0 = \boldsymbol{\Lambda} \otimes (\boldsymbol{1}\boldsymbol{\beta}^T - \boldsymbol{I}) - \boldsymbol{I} \otimes \boldsymbol{M}, \tag{5}$$

$$\boldsymbol{V}_1 = \boldsymbol{\Lambda} \oplus \boldsymbol{M} + \lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I}, \tag{6}$$

$$\boldsymbol{W}_0 = -(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)\boldsymbol{V}_0^{-1}, \tag{7}$$

$$\boldsymbol{W}_1 = -(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I})\boldsymbol{V}_1^{-1}, \tag{8}$$

$$\boldsymbol{W} = [(\boldsymbol{I} + \boldsymbol{W}_0 + \boldsymbol{W}_0\boldsymbol{W}_1)(\boldsymbol{I} \otimes \mathbf{1})]. \tag{9}$$

Let's prove that the following lemmas are valid.

Lemma 1.

$$\boldsymbol{V}_0(\boldsymbol{I} \otimes \mathbf{1}\boldsymbol{\beta}^T) = \boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T. \tag{10}$$

Proof. Given (5), we get:

$$\boldsymbol{V}_0(\boldsymbol{I} \otimes \mathbf{1}\boldsymbol{\beta}^T) = \big(\boldsymbol{\Lambda} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I}) - \boldsymbol{I} \otimes \boldsymbol{M}\big)(\boldsymbol{I} \otimes \mathbf{1}\boldsymbol{\beta}^T) = \big(\boldsymbol{\Lambda} \otimes \mathbf{1}\boldsymbol{\beta}^T\mathbf{1}\boldsymbol{\beta}^T\big) - (\boldsymbol{\Lambda} \otimes \mathbf{1}\boldsymbol{\beta}) - (\boldsymbol{I} \otimes \boldsymbol{M}\mathbf{1}\boldsymbol{\beta}^T).$$

Further, noting that $\boldsymbol{\beta}^T\mathbf{1} = 1$ and $\boldsymbol{M}\mathbf{1} = -\boldsymbol{\mu}$, we obtain the right part (10). □

Lemma 2.

$$-(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)(\mathbf{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T). \tag{11}$$

Proof. Obviously, if we consider that $\boldsymbol{\Lambda}\mathbf{1} = -\boldsymbol{\lambda}$. □

Lemma 3.

$$-\boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) = \boldsymbol{p}_1^T\boldsymbol{V}_0. \tag{12}$$

Proof. Multiplying equation (2) on the right by the matrix $\boldsymbol{I} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})$, we get:

$$\mathbf{0}^T = \boldsymbol{p}_0^T\left(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T\right)\left(\boldsymbol{I} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right) + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M})\left(\boldsymbol{I} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right) + \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T)\left(\boldsymbol{I} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right).$$

Let's consider each term of the right part separately:

$$\begin{aligned}\boldsymbol{p}_0^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T)\left(\boldsymbol{I} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right) &= \boldsymbol{p}_0^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T\mathbf{1}\boldsymbol{\beta}^T) - \boldsymbol{p}_0^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) = \\ &= \boldsymbol{p}_0^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) - \boldsymbol{p}_0^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{\beta}^T) = \mathbf{0}^T,\end{aligned}$$

$$\begin{aligned}\boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M})\left(\boldsymbol{I} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right) &= \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \otimes \boldsymbol{I} + \boldsymbol{I} \otimes \boldsymbol{M})\left(\boldsymbol{I} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right) = \\ &= \boldsymbol{p}_1^T\left(\boldsymbol{\Lambda} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I}) + \boldsymbol{I} \otimes \boldsymbol{M}(\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right) = \\ &= \boldsymbol{p}_1^T\left((\boldsymbol{\Lambda} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})) - \boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T - \boldsymbol{I} \otimes \boldsymbol{M}\right) = \boldsymbol{p}_1^T\left(\boldsymbol{V}_0 - \boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T\right),\end{aligned}$$

$$\boldsymbol{p}_2^T(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T)\left(\mathbf{1} \otimes (\mathbf{1}\boldsymbol{\beta}^T - \boldsymbol{I})\right) = \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T\mathbf{1}\boldsymbol{\beta}^T) - \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T) = \mathbf{0}^T.$$

So,

$$\mathbf{0}^T = \boldsymbol{p}_1^T(\boldsymbol{V}_0 - \boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T).$$

Therefore,

$$\boldsymbol{p}_1^T\boldsymbol{V}_0 = \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}).$$

Next, multiplying equation (1) on the right by $(\boldsymbol{I} \otimes \boldsymbol{\beta}^T)$, we obtain

$$\mathbf{0}^T = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) + \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \mu\boldsymbol{\beta}^T).$$

Given the obtained equality in the expression for $\boldsymbol{p}_1^T\boldsymbol{V}_0$, we come to the formula (12). Thus, the lemma is proved. □

The main result of this section is formulated in the form of a theorem.

Theorem 5. *The stationary distribution* $\{p_x, x \in \mathscr{X}\}$ *is determined by the formulas:*

$$\boldsymbol{p}_0^T = \boldsymbol{q}^T \boldsymbol{W}^{-1}, \tag{13}$$

$$\boldsymbol{p}_1 = \boldsymbol{p}_0^T \boldsymbol{W}_0, \tag{14}$$

$$\boldsymbol{p}_2^T = \boldsymbol{p}_1^T \boldsymbol{W}_1, \tag{15}$$

where $\boldsymbol{q}$ *is the only solution to the system of equations*

$$\boldsymbol{q}^T(\boldsymbol{\Lambda} + \lambda \boldsymbol{\alpha}^T) = \boldsymbol{0}^T, \tag{16}$$

$$\boldsymbol{q}^T \boldsymbol{1} = 1. \tag{17}$$

Proof. First, we substitute (14) and (15) into the equations (1)–(3) and let's make sure that after substitution they turn into identities.

Let's start with equation (1). Multiply it on the right by the matrix $(\boldsymbol{I} \otimes \boldsymbol{\beta}^T)$. As a result, we get:

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) + \boldsymbol{p}_0^T \boldsymbol{W}_0(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T).$$

Considering (7) and (10), we arrive at the identity for equation (1):

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T) - \boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)\boldsymbol{V}_0^{-1}\boldsymbol{V}_0(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T).$$

Next, consider equation (2), which, taking into account (11) and (15), is written as:

$$\boldsymbol{0}^T = -\boldsymbol{p}_0^T(\boldsymbol{\Lambda} \otimes \boldsymbol{\beta}^T)(\boldsymbol{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M}) + \boldsymbol{p}_1^T \boldsymbol{W}_1(\boldsymbol{I} \otimes \boldsymbol{\mu}\boldsymbol{\beta}^T).$$

Next, taking into account (10) and (12), we get:

$$\boldsymbol{0}^T = \boldsymbol{p}_1^T \boldsymbol{V}_0(\boldsymbol{1}\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \oplus \boldsymbol{M}) + \boldsymbol{p}_1^T \boldsymbol{W}_1 \boldsymbol{V}_0(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T).$$

And finally, multiplying both parts of the obtained ratio on the right by $(\boldsymbol{I} \otimes \boldsymbol{1}\boldsymbol{\beta}^T)$, we arrive at the identity for equation (2).

Substitute (14) and (15) in equation (3):

$$\boldsymbol{0}^T = \boldsymbol{p}_0^T \boldsymbol{W}_0(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) + \boldsymbol{p}_0^T \boldsymbol{W}_0 \boldsymbol{W}_1(\boldsymbol{\Lambda} \oplus \boldsymbol{M} + \lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I}).$$

Considering (6) and (8), we obtain the identity for equation (3).

Next, we multiply the equations (1)–(3) to the right, look at the matrix $(\boldsymbol{I} \otimes \boldsymbol{1})$ of the corresponding dimension and sum up the obtained equalities. As a result, we get

$$\begin{aligned}
\boldsymbol{0}^T = \boldsymbol{p}_0^T \boldsymbol{\Lambda} + \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_0^T \lambda \boldsymbol{\alpha}^T + \boldsymbol{p}_1^T(\boldsymbol{\Lambda} \otimes \boldsymbol{1}) -{}& \\
- \boldsymbol{p}_1^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_1^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{1}) + \boldsymbol{p}_2^T(\boldsymbol{\Lambda} \otimes \boldsymbol{1}) -{}& \\
- \boldsymbol{p}_2^T(\boldsymbol{I} \otimes \boldsymbol{\mu}) + \boldsymbol{p}_2^T(\lambda\boldsymbol{\alpha}^T \otimes \boldsymbol{I}) = \boldsymbol{p}_0^T(\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T) +{}& \\
+ \boldsymbol{p}_1^T(\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T)(\boldsymbol{I} \otimes \boldsymbol{1}) + \boldsymbol{p}_2^T(\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T)(\boldsymbol{I} \otimes \boldsymbol{1}) ={}& \\
= \left[\boldsymbol{p}_0^T + (\boldsymbol{p}_1^T + \boldsymbol{p}_2^T)(\boldsymbol{I} \otimes \boldsymbol{1})\right](\boldsymbol{\Lambda} + \lambda\boldsymbol{\alpha}^T).&
\end{aligned} \tag{18}$$

Substitute (14) and (15) into equations (18). As a result, we get:

$$\mathbf{0}^T = \left[\boldsymbol{p}_0^T + (\boldsymbol{p}_0^T \boldsymbol{W}_0 + \boldsymbol{p}_0^T \boldsymbol{W}_0 \boldsymbol{W}_1)(\boldsymbol{I} \otimes \mathbf{1})\right](\boldsymbol{\Lambda} + \lambda \boldsymbol{\alpha}^T). \tag{19}$$

Considering (9), we write the system (19) in the form (16):

$$\mathbf{0}^T = \boldsymbol{p}_0^T \boldsymbol{W}(\boldsymbol{\Lambda} + \lambda \boldsymbol{\alpha}^T),$$

i.e. the vector $\boldsymbol{p}_0^T \boldsymbol{W}$ is satisfies the system (16) and, therefore,

$$\boldsymbol{p}_0^T \boldsymbol{W} = C \boldsymbol{q}^T. \tag{20}$$

Since the matrix of coefficients $\boldsymbol{\Lambda} + \lambda \boldsymbol{\alpha}^T$ of the system (16) is indissoluble due to the irreducibility of the PH representation $(\boldsymbol{\alpha}, \boldsymbol{\Lambda})$, this system, taking into account (17) has a unique solution [18].

It remains for us to define C. According to (4), we have

$$\boldsymbol{p}_0^T \mathbf{1} + \boldsymbol{p}_1 \mathbf{1} + \boldsymbol{p}_2^T \mathbf{1} = 1,$$

which, taking into account (14) and (15), we write in the form:

$$\boldsymbol{p}_0^T \mathbf{1} + \boldsymbol{p}_0^T \boldsymbol{W}_0 \mathbf{1} + \boldsymbol{p}_0^T \boldsymbol{W}_0 \boldsymbol{W}_1 \mathbf{1} = 1,$$

The resulting equality, taking into account (9) and (20), has the form

$$\boldsymbol{p}_0^T \boldsymbol{W} \mathbf{1} = C \boldsymbol{q}^T \mathbf{1} = 1.$$

However, according to (17) $\boldsymbol{q}^T \mathbf{1} = 1$. Therefore, $C = 1$. So, we have shown that $\boldsymbol{p}_0^T \boldsymbol{W}$ coincides with the vector $\boldsymbol{q}$, i.e. to determine $\boldsymbol{p}_0$, we can use formula (13), having previously determined $\boldsymbol{q}$ from the system (16)–(17). We were convinced of the validity of formulas (14) and (15) earlier by substituting them in SEE (1)–(3) and turning the equations of the system into identities.

Thus, the theorem is proved. □

5. Markov chains nested at the moment of entry of applications into the system

Let's build a Markov chain (MC) embedded in MP $X(t)$ at the moments $t + 0$ of receipt of applications to the system over a set of states:

$$\mathscr{X}_A^+ = \bigcup_{k=1}^{2} \mathscr{X}_{A,k}^+,$$

where $\mathscr{X}_{A,k}^+ = \{(k, j), j = \overline{1, m}\}$, $k = 1, 2$.

The state (k, j) means that immediately after the application is received in the system, there are k applications in it and at the same time the service process is in phase j, $j = \overline{1, m}$, $k = 1, 2$. It was received immediately for maintenance and the selection of the maintenance phase was carried out instantly at the time of its receipt in accordance with the initial distribution set by the vector $\boldsymbol{\beta}$.

To determine the stationary probabilities $p_{A,x}^+$, $x \in \mathscr{X}_A^+$, the states of the nested MC will use the results of the work [19]. In accordance with the recommendations of this work, we will differentiate the jumps of MP $X(t)$, considering the jumps associated with the receipt of applications into the system to be "correct". At the same time, it should be noted that in our system, the incoming application cannot be lost, but when the drive is busy, it "kills" the application located in it. Considering the above, and also applying the formulas of [19] to calculate the stationary distribution of states of the nested MC, we come to the following result:

$$\boldsymbol{p}_{A,1}^{+\ T} = \frac{1}{\lambda} \boldsymbol{p}_0^T (\boldsymbol{\lambda} \otimes \boldsymbol{\beta}^T),$$

$$\boldsymbol{p}_{A,2}^{+\,T} = \frac{1}{\lambda}\left[\boldsymbol{p}_1^T + \boldsymbol{p}_2^T\right](\boldsymbol{\lambda} \otimes \boldsymbol{I}) = \frac{1}{\lambda}\left[\mathbf{1}^T - \boldsymbol{p}_0^T\right]\boldsymbol{\lambda}, \tag{21}$$

where $\boldsymbol{p}_{A,k}^{+\,T} = \left(p_{A,(k,1)}^{+}; ...; p_{A,(k,m)}^{+}\right)$, $\lambda = (-\boldsymbol{\alpha}^T\boldsymbol{\Lambda}^{-1}\mathbf{1})^{-1}$.

Next, we will build a MC embedded in the MP $X(t)$ at the moments $t-0$ of the receipt of applications into the system. The set of states of a given MC will be determined:

$$\mathscr{X}_A^- = \bigcup_{k=0}^{2} \mathscr{X}_{A,k}^-,$$

where $\mathscr{X}_{A,0}^- = \{(0)\}$, $\mathscr{X}_{A,k}^- = \left\{(k,j), j = \overline{1,m}\right\}$, $k = 1, 2$.

The state (0) means that immediately before the first application was received into the system, the system was empty, and the state (k, j) means that immediately before the next application was received into the system, there were k applications in it, and at the same time the application on the device was serviced in phase j.

In accordance with the result of [19], we obtain:

$$p_{A,0}^- = \frac{1}{\lambda}\boldsymbol{p}_0^T\boldsymbol{\lambda},$$

$$\boldsymbol{p}_{A,k}^{-\,T} = \frac{1}{\lambda}\boldsymbol{p}_k^T(\boldsymbol{\lambda} \otimes \boldsymbol{I}), \quad k = 1, 2. \tag{22}$$

where $\boldsymbol{p}_{A,k}^{-\,T} = \left(p_{A,(k,1)}^{-}; ...; p_{A,(k,m)}^{-}\right)$.

From (21) and (22) it follows that

$$\boldsymbol{p}_{A,2}^{+\,T} = \boldsymbol{p}_{A,1}^{-\,T} + \boldsymbol{p}_{A,2}^{-\,T}. \tag{23}$$

Formula (23) means that in order for there to be two applications in the system immediately after the receipt of the next application, it is necessary and sufficient that there should be one or two applications immediately before the next application is received. In the first case, the incoming application will take up free space in the drive. In the second case, it will displace ("kill") the application in the drive and take its place.

6. Main indicators of system performance

As noted earlier, there is no loss of applications in the system we are considering due to lack of storage space. However, some applications leave the system without waiting for service. Let's call them "unsuccessful". An application becomes "unsuccessful" if there are two conditions: firstly, at the moment $t-0$ it enters the system, there must be one or two applications in the system, which will automatically turn it into an application waiting for the device to be released, and secondly, its waiting time should be longer before generation of the next application, which will force our application to leave the system by implementing a queue update mechanism. Thus, the time spent by the "unsuccessful" application in the system is equal to the time until the next application is generated. To calculate the probability that the application will be "unsuccessful", consider the following probabilities [20]:

$$\alpha_1 = \boldsymbol{\beta}^T \int_0^\infty e^{\boldsymbol{M}t} dA(t)\mathbf{1},$$

$$\alpha_2 = \boldsymbol{\alpha}^T \int_0^\infty e^{\boldsymbol{\Lambda}t} dB(t)\mathbf{1}.$$

Note that α_1 is a possibility that the service of the application on the device will not be completed during the time before the next request is generated.

Let's denote by $\bar{\gamma}$ the probability that the next application received by the system will be "unsuccessful". Considering the above, it can be argued that

$$\bar{\gamma} = \left(\boldsymbol{p}_{A,1}^{-}{}^{T} + \boldsymbol{p}_{A,2}^{-}{}^{T}\right)\mathbf{1} \cdot \alpha_1 = (1 - p_{A,0}^{-})\alpha_1, \tag{24}$$

$$\gamma = 1 - \bar{\gamma} = p_{A,0}^{-} + (1 - p_{A,0}^{-})\alpha_2.$$

Let's denote by λ_A the intensity of the flow of "successful" applications, and by λ_D the intensity of the outgoing flow. It is obvious that

$$\lambda_A = \lambda\gamma,$$

$$\lambda_D = \mu(1 - \boldsymbol{p}_0^{+}\mathbf{1}), \quad \text{where } \mu = (-\boldsymbol{\beta}^T\boldsymbol{M}^{-1}\mathbf{1})^{-1}.$$

In stationary mode, $\lambda_A = \lambda_D$, from where we get another not so obvious formula for γ:

$$\gamma = \frac{\mu}{\lambda}(1 - \boldsymbol{p}_0^T\mathbf{1}). \tag{25}$$

Obviously, if you enter the notation for the system load $\rho = \frac{\lambda}{\mu}$ and for the device utilization factor $u = 1 - \boldsymbol{p}_0^T\mathbf{1}$, then the formula (25) can be written as

$$\rho\gamma = u. \tag{26}$$

Considering that only "successful" applications are serviced in the system, formula (26) acquires a quite obvious probabilistic meaning: the utilization factor of the device is equal to the loading of the system with "successful" applications.

7. Conclusion

The paper investigates a single-line service system with queue updates and phase-type distributions. As a result, a recurrent matrix algorithm has been developed to calculate the stationary distribution of states of the Markov process describing the stochastic behavior of the system, and expressions for the main indicators of its performance have been obtained. The considered system is planned to be used as a mathematical model in the tasks of analyzing and managing the age of information. The authors are confident that this study will allow them to obtain sufficiently accurate estimates of the age of information for real technical systems in the future, due to the universality of phase-type distributions.

Author Contributions: Conceptualization, S.I. Matyushenko and K.E. Samouylov; methodology, S.I. Matyushenko; validation, S.I. Matyushenko and N.Yu. Gritsenko; writing—original draft preparation, S.I. Matyushenko; writing—review and editing, S.I. Matyushenko and N.Yu. Gritsenko; visualization, N.Yu. Gritsenko; supervision, K.E. Samouylov; project administration, K.E. Samouylov. 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. Kaul, S., Gruteser, M., Rai, V. & Kenney, J. *Minimizing age of information in vehicular networks* in *2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks* (2011), 350–358. doi:10.1109/SAHCN.2011.5984917.
2. Kaul, S., Yates, R. & Gruteser, M. *Real-time status: How often should one update?* in *2012 Proceedings IEEE INFOCOM* (2012), 2731–2735. doi:10.1109/INFCOM.2012.6195689.
3. Bedewy, A. M., Sun, Y. & Shroff, N. B. *Age-optimal information updates in multihop networks* in *2017 IEEE International Symposium on Information Theory (ISIT)* (2017), 576–580. doi:10.1109/ISIT.2017.8006593.
4. Bojan, T. M., Kumar, U. R. & Bojan, V. M. *An internet of things based intelligent transportation system* in *2014 IEEE International Conference on Vehicular Electronics and Safety* (2014), 174–179. doi:10.1109/ICVES.2014.7063743.
5. Hu, C. & Dong, Y. Age of information of two-way data exchanging systems with power-splitting. *Journal of Communications and Networks* **21,** 295–306. doi:10.1109/JCN.2019.000037 (2019).
6. Costa, M., Codreanu, M. & Ephremides, A. *Age of information with packet management* in *IEEE International Symposium on Information Theory (ISIT)* (2014), 1583–1587.
7. Costa, M., Codreanu, M. & Ephremides, A. On the Age of Information in Status Update Systems With Packet Management. *IEEE Transactions on Information Theory* **62,** 1897–1910. doi:10.1109/TIT.2016.2533395 (2016).
8. Chiariotti, F., Vikhrova, O., Soret, B. & Popovski, P. Peak Age of Information Distribution for Edge Computing With Wireless Links. *IEEE Transactions on Communications* **69,** 3176–3191. doi:10.1109/TCOMM.2021.3053038 (2021).
9. Kadota, I., Uysal-Biyikoglu, E., Singh, R. & Modiano, E. *Minimizing the Age of Information in broadcast wireless networks* in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (2016), 844–851. doi:10.1109/ALLERTON.2016.7852321.
10. Kosta, A., Pappas, N., Ephremides, A. & Angelakis, V. *Non-linear Age of Information in a Discrete Time Queue: Stationary Distribution and Average Performance Analysis* in *ICC 2020 — 2020 IEEE International Conference on Communications (ICC)* (2020), 1–6. doi:10.1109/ICC40277.2020.9148775.
11. Talak, R., Karaman, S. & Modiano, E. Improving Age of Information in Wireless Networks With Perfect Channel State Information. *IEEE/ACM Transactions on Networking* **28,** 1765–1778. doi:10.1109/TNET.2020.2996237 (2020).
12. Tripathi, V., Talak, R. & Modiano, E. *Age of information for discrete time queues* 2019.
13. Wijerathna Basnayaka, C. M., Jayakody, D. N. K., Ponnimbaduge Perera, T. D. & Vidal Ribeiro, M. *Age of Information in an URLLC-enabled Decode-and-Forward Wireless Communication System* in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)* (2021), 1–6. doi:10.1109/VTC2021-Spring51267.2021.9449007.
14. Bocharov, P. P. & Zaryadov, I. S. Stationary probability distribution in a queuing system with an update [Stacionarnoe raspredelenie veroyatnostey v sisteme massovogo obslugivaniya s obnovleniem]. *Vestnik RUDN Seriya Matematika. Informatika. Phizika.* in Russian, 14–23 (2007).
15. Kaul, S. K., Yates, R. D. & Gruteser, M. *Status updates through queues* in *2012 46th Annual Conference on Information Sciences and Systems (CISS)* (2012), 1–6. doi:10.1109/CISS.2012.6310931.
16. Inoue, Y., Masuyama, H., Takine, T. & Tanaka, T. A General Formula for the Stationary Distribution of the Age of Information and Its Application to Single-Server Queues. *IEEE Transactions on Information Theory* **65,** 8305–8324. doi:10.1109/TIT.2019.2938171 (2019).
17. Bocharov, P. P. & Pechinkin, A. V. *Queueing Theory [Teoriya massovogo obsluzhivaniya]* in Russian. 529 pp. (Izd-vo RUDN, Moscow, 1995).

18. Basharin, G. P., Bocharov, P. P. & Kogan, Y. A. *Queue analysis in computer networks [Analiz ocheredey v vichislitelnih setyah]* in Russian (Nauka, Moscow, 1989).
19. Naumov, V. A. *O predelnih veroyatnostyah polumarkovskogo processa [On the limiting probabilities of a semi-Markov process]* in Russian. 35–39 (Universitet drugbi narodov, Moscow, 1975).
20. Gelenbe, E. & Pujolle, G. *Introduction to queueing networks* (John Wiley, New York, 1987).

Information about the authors

Sergey I. Matyushenko—Candidate of Physical and Mathematical Sciences, Assistant professor of Department of Probability Theory and Cyber Security of RUDN University (e-mail: matyushenko-si@rudn.ru, ORCID: 0000-0001-8247-8988)

Konstantin E. Samouylov—Professor, Doctor of Technical Sciences, Head of the Department of Probability Theory and Cyber Security of RUDN University (e-mail: samuylov-ke@rudn.ru, ORCID: 0000-0002-6368-9680)

Nikolai Yu. Gritsenko—PhD student of Department of Probability Theory and Cyber Security of RUDN University (e-mail: 1142221032@rudn.ru)

УДК 519.21
DOI: 10.22363/2658-4670-2024-32-3-271–282 EDN: BAUGIT

Анализ системы обслуживания единичной ёмкости с распределениями фазового типа и обновлением очереди

С. И. Матюшенко, К. Е. Самуйлов, Н. Ю. Гриценко

Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. В данной работе исследуется однолинейная система массового обслуживания с накопителем единичной ёмкости и обновлением очереди. Под обновлением понимается следующий механизм: заявка, поступающая в систему и застающая в накопителе другую заявку, уничтожает её, занимая её место в накопителе. Следует заметить, что системы с тем или иным механизмом обновления давно привлекают внимание исследователей, поскольку имеют важное прикладное значение. В последнее время интерес к системам подобного рода вырос в связи с задачами оценки и управления возрастом информации. Система с механизмом обновления очереди, подобная рассматриваемой нами, уже исследовалась ранее в работах других авторов. Однако в этих работах речь шла о простейшем варианте системы с пуассоновским потоком и экспоненциальным обслуживанием. В данной работе мы рассматриваем систему с потоком и обслуживанием фазового типа. В результате проведённого исследования нами был разработан рекуррентный матричный алгоритм для расчёта стационарного распределения состояний марковского процесса, описывающего стохастическое поведение рассматриваемой системы, и получены выражения для основных показателей её производительности.

Ключевые слова: система массового обслуживания, распределение фазового типа, механизм обновления очереди