

Article

# Vacation Queueing Model for Performance Evaluation of Multiple Access Information Transmission Systems without Transmission Interruption

Alexander Dudin <sup>1,2,\*</sup> , Sergei Dudin <sup>1</sup>, Valentina Klimenok <sup>1</sup> and Yuliya Gaidamaka <sup>2</sup> 

<sup>1</sup> Department of Applied Mathematics and Computer Science, Belarusian State University, 4, Nezavisimosti Ave., 220030 Minsk, Belarus; dudins@bsu.by (S.D.); vklimenok@yandex.ru (V.K.)

<sup>2</sup> Applied Mathematics and Communications Technology Institute, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya St, 117198 Moscow, Russia; gaydamaka-yuv@rudn.ru

\* Correspondence: dudin@bsu.by

**Abstract:** We consider a *MAP/PH/1*-type queueing system with server vacations as a model that is useful for the analysis of multiple access systems with polling discipline without transmission interruption. Vacation of the server corresponds to the service providing competitive information flows to the polling system. In this paper, we consider a vacation queueing model under pretty general assumptions about the probabilistic distributions describing the behavior of the system and the realistic assumption, in many real-world systems, that ongoing service cannot be terminated ahead of schedule. We derive the criterion of the stable operation of the system and the stationary distributions of the system states and the waiting time. An illustrative numerical example is presented.



**Citation:** Dudin, A.; Dudin, S.; Klimenok, V.; Gaidamaka, Y. Vacation Queueing Model for Performance Evaluation of Multiple Access Information Transmission Systems without Transmission Interruption. *Mathematics* **2021**, *9*, 1508. <https://doi.org/10.3390/math9131508>

Academic Editor: Manuel Alberto M. Ferreira

Received: 10 June 2021  
Accepted: 23 June 2021  
Published: 28 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** queueing model; random multiple access; time-limited service; polling system; stability

## 1. Introduction

A popular discipline for service provision, in particular information transmission, in many systems with multiple access, is polling, see, for example, [1–3]. Under this discipline, different flows of information, which should be transmitted in the system, sequentially obtain time slots for access to the transmission thread. Queueing theory is one of the most powerful tools for the analysis of polling systems. The polling systems are usually described by very complicated multi-dimensional stochastic processes and queueing models with server vacations are very useful for the analysis of these processes. Vacation queueing models have been studied in a huge number of works, for example, in [4–8].

In this paper, we consider a vacation queueing model with an exhaustive time-limited service. This means that the server takes a vacation when it becomes idle or when the so-called maximum attendance time expires, whichever occurs first. An important feature of the considered model is the assumption that the ongoing service cannot be interrupted when the maximum attendance time expires. A similar model has previously been considered, for example, in [9–11]. In contrast to these papers, we consider a system with a more general Markovian Arrival process, which much better describes the real-world information flows than the stationary Poisson process. We also assume that the maximum attendance time has the so-called phase-type (*PH*) distribution, while only exponential or degenerate distributions have been previously considered in the literature. The exponential distribution is one of the special cases of the phase-type distribution. As follows from [12], to approximate any arbitrary distribution in the sense of weak convergence, a *PH*-type distribution can be used.

Consideration of the model analyzed in this paper was originated during the implementation of applied research to optimize the work of the inter-banking processing center

of the Republic of Belarus, which handles all money transactions between the banks. The results of analysis of a similar queueing model have recently been reported in [13]. That paper presents the motivation to consider these types of queueing system, their practical importance, the state-of-the-art, and the relevant literature. The principal difference between the models under study in [13] and in this paper is the following: in the model considered in [13], it was assumed that the customers are impatient. This means that any customer waiting in the queue departs from the system without service after a random amount of time. We assume that this time depends on the state of the server and has an exponential distribution with the nonnegative parameter  $\alpha_r$ . The server can be in three states:  $r = 0$  if the server has a vacation; the state  $r = 1$  corresponds to the case when the server is busy and the maximum server attendance time is not finished; and  $r = 2$  if the server is also busy but the maximum server attendance time is expired. It is assumed in [13] that  $\alpha_r > 0$  for at least one of the values of  $r$ ,  $r = 0, 1, 2$ .

The main difficulty in the analysis of the model with impatient customers in [13] consists of the fact that the process of the system is a multi-dimensional, level-dependent Markov chain. This excludes the possibility of using the well-known results for the level-independent Quasi-Birth-and-Death processes; see, for example, [14], for an analysis of the system. Therefore, more complicated results for the so-called Asymptotically Quasi-Toeplitz Markov chains (see [15]) were used in [13] for the analysis of a queueing model. In particular, it was proven that if  $\alpha_r > 0$  for at least one of the values of  $r$ ,  $r = 0, 1, 2$ , then the Markov chain describing the vacation queueing model under study is ergodic for any values of other parameters of the system, in particular of the pattern of the arrival process and its rate, distributions of service, vacation and maximum attendance times. The impatience of customers is the inherent feature of many real-world systems and this explains its account in the paper [13]. However, in certain systems, the customers are absolutely patient and depart from the system only after receiving the service. In particular, in modeling the inter-banking processing center noted above, it was suggested that any financial transaction accepted for processing in the center must be implemented and committed. Therefore, the parameter  $\alpha_r = 0$  for all values of  $r$ ,  $r = 0, 1, 2$ , and the ergodicity condition of the Markov chain (stability condition of the queueing model) has to be derived. Fulfillment of this condition has to be verified before the stationary distribution of the system states will be computed. Such a condition is presented in this paper. The practical importance of this condition consists of the possibility of its use for planning the throughput of the system, customers' admission discipline, and the proper choice of the equipment required for providing the desired quality of service.

This paper has the following structure: The mathematical model of the considered system is presented in Section 2. In Section 3, the behavior of the system under study is described by a continuous-time multi-dimensional Markov chain. This chain belongs to the class of Quasi-Birth-and-Death processes. The generator of this chain is given. The ergodicity condition is derived and the vectors of the stationary probabilities of the system states are computed in the matrix geometric form in Section 4. In Section 5, the formulas for the main performance measures of the system are written down. Section 6 presents the analysis of the waiting time distribution including the derivation of the formula for computing the mean waiting time. Section 7 contains the major findings and conclusions.

## 2. Mathematical Model

A single-server queueing system is analyzed. The system has an infinite buffer. We assume the Markovian arrival flow (MAP) of customers. Arrivals in the MAP are described by an underlying process  $\nu_t$ ,  $t \geq 0$ , that is an irreducible continuous-time Markov chain with a finite state space  $\{0, \dots, W\}$ . The intensities of this chain transitions that are accompanied by the generation of one customer (not accompanied by customer arrivals) are combined into the square matrix  $\mathbf{D}_1(\mathbf{D}_0)$ , of size  $W + 1$ . The infinitesimal generator of the process  $\nu_t$ ,  $t \geq 0$  is the matrix  $\mathbf{D}(1) = \mathbf{D}_0 + \mathbf{D}_1$ .

The vector of the stationary distribution (invariant vector)  $\theta$  of the process  $v_t$  is a unique solution of the following system

$$\theta \mathbf{D}(1) = \mathbf{0}, \theta \mathbf{e} = 1.$$

Hereinafter,  $\mathbf{e}$  is a column vector consisting of ones and  $\mathbf{0}$  is a zero row vector of appropriate size. The size of a vector is indicated using a lower index if it is not clear from context, for instance  $\mathbf{e}_{\bar{W}}$  means the unit column vector of size  $\bar{W} = W + 1$ .

The fundamental intensity  $\lambda$  of the *MAP* arrival process can be calculated by the following formula

$$\lambda = \theta \mathbf{D}_1 \mathbf{e}.$$

The rate  $\lambda$  defines the average number of arriving customers per unit of time.

Detailed information about the *MAP* arrival process and its usefulness in modeling customers flows in telecommunication systems can be found in, for example, [16–21]. An example of when *MAP* with a special matrix structure is used to simulate traffic in real networks can be found in a recent paper [22].

We assume that the busy periods of the server alternate with the vacation periods. The service period starts immediately after completion of the vacation period conditioned on the fact that the system is not empty. If the system is empty, the server takes another vacation. Since the arrival of a customer does not imply the beginning of a busy period, each arriving customer joins the buffer. We assume the First In–First Out service discipline.

The duration of the vacation time has the *PH* distribution with the parameters  $(\gamma, \Gamma)$ . The vacation time is defined as the time until the continuous-time Markov chain  $\zeta_t, t \geq 0$ , having  $R$  transient states and one absorbing state, transits to the absorbing state conditioned on the fact that the initial state is chosen among transient states. The initial state of the process  $\zeta_t, t \geq 0$ , is randomly chosen according to the probability distribution defined by the probabilistic row vector  $\gamma = (\gamma_1, \dots, \gamma_R)$ . The intensities of the transitions of the Markov chain  $\zeta_t, t \geq 0$ , within the set of transient states, are defined by the  $R$ -dimension square irreducible matrix  $\Gamma$ . The intensities of transitions to the absorbing state are defined by a column vector  $\Gamma_0 = -\Gamma \mathbf{e}$ . The distribution function of the vacation time can be found as  $1 - \gamma e^{\Gamma x} \mathbf{e}$ . The Laplace–Stieltjes transform of this distribution function is  $\gamma(sI - \Gamma)^{-1} \Gamma_0, \operatorname{Re} s > 0$ . The mean vacation time can be found as  $v_1 = \gamma(-\Gamma)^{-1} \mathbf{e}$ .

We assume that the service period also has the *PH* distribution with the parameters  $(\tau, \mathbf{T})$ . The service period is defined as the time until the continuous-time Markov chain  $\chi_t, t \geq 0$ , having  $K$  transient states and one absorbing state, transits to the absorbing state conditioned on the fact that the initial state is chosen among transient states. Let us denote  $\mathbf{T}_0 = -\mathbf{T} \mathbf{e}$ . The mean service period can be found as  $\tau_1 = \tau(-\mathbf{T})^{-1} \mathbf{e}$ .

When the service period starts, the first customer from the buffer is chosen for service. The service time of any customer also has the *PH* distribution with the parameters  $(\beta, \mathbf{S})$ . An underlying process of this time is denoted as  $\eta_t, t \geq 0$ . The Markov chain  $\eta_t, t \geq 0$ , has  $M$  transient states. Let us denote  $\mathbf{S}_0 = -\mathbf{S} \mathbf{e}$ . The mean service time can be found as  $b_1 = \beta(-\mathbf{S})^{-1} \mathbf{e}$ .

If, during the service completion epoch, the buffer is idle, the vacation period immediately begins. In standard  $T$ -limited service (see e.g., [8]) it is assumed that when the service period is expired, the vacation of the server immediately begins and the customer who obtains service during the service period completion epoch is lost or returns to the buffer. In the model under study, we assume that the service of a customer cannot be terminated. After the service period completion the vacation period starts only after the customer service completion.

The main goal of the analysis is to obtain the condition of the existence of the stationary distribution of the system states and to compute this distribution as well as the distribution of the waiting time of an arbitrary customer.

### 3. The Process of the System States

Let, at the moment  $t, t \geq 0,$

- $i_t$  be the number of customers in the system,  $i_t \geq 0;$
- $r_t$  be the current status of the server:  $r_t = 0$  if the server is on vacation,  $r_t = 1$  if the server is busy and the service period is not finished, and  $r_t = 2$  if the server is busy but the maximum server attendance time is finished;
- $v_t$  be the state of the underlying process of the MAP,  $v_t = \overline{0, W};$
- $m_t$  be the state of the underlying process  $\zeta_t$  of the vacation period in the case  $r_t = 0,$   $m_t = (\chi_t, \eta_t)$  (the pair of the states of the underlying processes of the service period and the service time) in the case  $r_t = 1,$  and  $m_t = \eta_t$  if  $r_t = 2.$

One can verify that the behavior of the system under study is described by the multi-dimensional process  $\zeta_t = \{i_t, r_t, v_t, m_t\}, t \geq 0,$  which is an irreducible continuous-time Markov chain. This Markov chain has the following state space

$$\Omega = \{(i, 0, v, \xi), i \geq 0, 0 \leq v \leq W, 1 \leq \xi \leq R\} \\ \cup \{(i, 1, v, \chi, \eta), i \geq 1, 0 \leq v \leq W, 1 \leq \chi \leq K, 1 \leq \eta \leq M\} \\ \cup \{(i, 2, v, \eta), i \geq 1, 0 \leq v \leq W, 1 \leq \eta \leq M.\}$$

Let us denote the generator of the Markov chain  $\zeta_t, t \geq 0,$  by  $\mathbf{Q}.$  The generator is a square matrix with all negative diagonal entries. Moduli of the diagonal entries define the intensity of the leaving of the Markov chain in the corresponding states. The non-diagonal entries of the generator are non-negative. These entries define the transition intensities of the process  $\zeta_t, t \geq 0,$  from one state to another.

To obtain the generator  $\mathbf{Q}$  in the block-matrix form, we enumerate the states of the process  $\zeta_t$  in the direct lexicographic order of the components and assume that all states of the Markov chain  $\zeta_t$  having values  $(i, r)$  of the first two components form the *sub-level*  $(i, r).$  One can verify that the number of states of the level  $(i, 0)$  is  $\overline{WR},$  the number of states of the level  $(i, 1)$  is  $\overline{WKM}$  and the number of states of the level  $(i, 2)$  is  $\overline{WM}.$  Then, we form level  $i, i \geq 0,$  from the sub-levels  $(i, r), r = 0, 1, 2.$

**Lemma 1.** *The generator  $\mathbf{Q}$  of the Markov chain  $\zeta_t$  has the following block-tridiagonal structure:*

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{0,0} & \mathbf{Q}_{0,1} & \mathbf{O} & \mathbf{O} & \dots \\ \mathbf{Q}_{1,0} & \mathbf{Q}^0 & \mathbf{Q}^+ & \mathbf{O} & \dots \\ \mathbf{O} & \mathbf{Q}^- & \mathbf{Q}^0 & \mathbf{Q}^+ & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where the non-zero blocks  $\mathbf{Q}_{i,j}$  are defined as follows:

$$\mathbf{Q}_{0,0} = \mathbf{D}_0 \oplus (\mathbf{\Gamma} + \mathbf{\Gamma}_0 \gamma), \quad \mathbf{Q}_{0,1} = (\mathbf{D}_1 \otimes \mathbf{I}_R \quad | \mathbf{O} \quad | \mathbf{O}), \quad \mathbf{Q}_{1,0} = \begin{pmatrix} \mathbf{O} \\ \mathbf{I}_{\overline{W}} \otimes \mathbf{e}_K \otimes \mathbf{S}_0 \otimes \gamma \\ \mathbf{I}_{\overline{W}} \otimes \mathbf{S}_0 \otimes \gamma \end{pmatrix}, \\ \mathbf{Q}_{i,i-1} = \mathbf{Q}^- = \begin{pmatrix} \mathbf{O} & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{\overline{W}} \otimes \mathbf{I}_K \otimes \mathbf{S}_0 \otimes \beta & \mathbf{O} \\ \mathbf{I}_{\overline{W}} \otimes \mathbf{S}_0 \otimes \gamma & \mathbf{O} & \mathbf{O} \end{pmatrix}, \\ \mathbf{Q}_{i,i} = \mathbf{Q}^0 = \begin{pmatrix} \mathbf{D}_0 \oplus \mathbf{\Gamma} & \mathbf{I}_{\overline{W}} \otimes \mathbf{\Gamma}_0 \otimes \tau \otimes \beta & \mathbf{O} \\ \mathbf{O} & \mathbf{D}_0 \oplus \mathbf{T} \oplus \mathbf{S} & \mathbf{I}_{\overline{W}} \otimes \mathbf{T}_0 \otimes \mathbf{I}_M \\ \mathbf{O} & \mathbf{O} & \mathbf{D}_0 \oplus \mathbf{S} \end{pmatrix}, \\ \mathbf{Q}_{i,i+1} = \mathbf{Q}^+ = \begin{pmatrix} \mathbf{D}_1 \otimes \mathbf{I}_R & \mathbf{O} & \mathbf{O} \\ \mathbf{O} & \mathbf{D}_1 \otimes \mathbf{I}_K \otimes \mathbf{I}_M & \mathbf{O} \\ \mathbf{O} & \mathbf{O} & \mathbf{D}_1 \otimes \mathbf{I}_M \end{pmatrix}, \quad i \geq 1.$$

Here,  $\mathbf{I}$  denotes the identity matrix, and  $\mathbf{O}$  denotes a zero matrix, and  $\otimes, \oplus$  are the symbols of the Kronecker product and the sum of the matrices correspondingly, see [23]. If it is necessary, the size of the matrix  $\mathbf{I}$  is given as the low index, that is,  $\mathbf{I}_K$  represents the identity matrix of size  $K$ .

Proof of Lemma 1 is implemented using careful analysis of all possible transitions in the Markov chain  $\zeta_t, t \geq 0$ , during an infinitesimally small time and further forming the corresponding intensities of transitions into the block matrices. The Kronecker product and sum of matrices are used here for the description of the intensities of the simultaneous transition of several independent Markov chains. One can see that the blocks  $\mathbf{Q}_{i,j}$  (except the boundary blocks) do not depend on the levels  $i$  and  $j$  separately but depend on the difference  $j - i$ . The Markov chain  $\zeta_t$ , having such a property according to [14], belongs to the class of level-independent Quasi-Birth-and-Death processes. The boundary blocks  $\mathbf{Q}_{0,0}, \mathbf{Q}_{0,1}, \mathbf{Q}_{1,0}$  of the generator have fewer sub-blocks than other sub-blocks, since the process  $r_t$  can only be in state 0 when the system is idle.

#### 4. Ergodicity Condition

The criterion for the ergodicity of the Markov chain  $\zeta_t$  is presented in the following statement.

**Theorem 1.** *The Markov chain  $\zeta_t$  is ergodic if and only if the inequality is fulfilled:*

$$-\mathbf{x}_1(\mathbf{T} \oplus \mathbf{S})\mathbf{e} > \lambda. \tag{1}$$

Here, the vector  $\mathbf{x}_1$  is computed as

$$\mathbf{x}_1 = (1, 0, \dots, 0)\tilde{\mathbf{A}}^{-1}, \tag{2}$$

where the non-singular matrix  $\tilde{\mathbf{A}}$  is given by

$$\tilde{\mathbf{A}} = \mathbf{A}\hat{\mathbf{I}} + \bar{\mathbf{I}},$$

$$\mathbf{A} = \mathbf{I}_K \otimes (\mathbf{S}_0 \otimes \boldsymbol{\beta} + \mathbf{S}) + \mathbf{T} \otimes \mathbf{I}_M + (\mathbf{T}_0 \otimes \mathbf{e}_M)(\boldsymbol{\tau} \otimes \boldsymbol{\beta}), \tag{3}$$

$$\hat{\mathbf{I}} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \bar{\mathbf{I}} = (\mathbf{b} \mid \mathbf{O}),$$

$$\mathbf{b} = \mathbf{e}_{KM} + \mathbf{T}_0 \otimes (-\mathbf{S})^{-1}\mathbf{e}_M + v_1\mathbf{T}_0 \otimes \mathbf{e}_M. \tag{4}$$

**Proof.** According to [14], the Markov chain  $\zeta_t$  has a stationary mode if and only if the following inequality holds:

$$\mathbf{y}\mathbf{Q}^-\mathbf{e} > \mathbf{y}\mathbf{Q}^+\mathbf{e}. \tag{5}$$

Here, the vector  $\mathbf{y}$  can be found as the unique solution of the system

$$\mathbf{y}(\mathbf{Q}^- + \mathbf{Q}^0 + \mathbf{Q}^+) = \mathbf{0}, \tag{6}$$

$$\mathbf{y}\mathbf{e} = 1. \tag{7}$$

One can see that the matrix  $\mathbf{Q}^- + \mathbf{Q}^0 + \mathbf{Q}^+$  has the following form

$$\begin{aligned} & \mathbf{Q}^- + \mathbf{Q}^0 + \mathbf{Q}^+ = \\ & = \begin{pmatrix} \mathbf{D}_0 \oplus \boldsymbol{\Gamma} + \mathbf{D}_1 \otimes \mathbf{I}_R & \mathbf{I}_{\bar{W}} \otimes \boldsymbol{\Gamma}_0 \otimes \boldsymbol{\tau} \otimes \boldsymbol{\beta} & \mathbf{O} \\ \mathbf{O} & \mathbf{D}_0 \oplus \mathbf{T} \oplus \mathbf{S} + (\mathbf{D}_1 \otimes \mathbf{I}_K) \oplus (\mathbf{S}_0 \otimes \boldsymbol{\beta}) & \mathbf{I}_{\bar{W}} \otimes \mathbf{T}_0 \otimes \mathbf{I}_M \\ \mathbf{I}_{\bar{W}} \otimes \mathbf{S}_0 \otimes \boldsymbol{\gamma} & \mathbf{O} & \mathbf{D}_0 \oplus \mathbf{S} + \mathbf{D}_1 \otimes \mathbf{I}_M \end{pmatrix}. \end{aligned} \tag{8}$$

Looking at the structure of this matrix given by formula (8), we can assume that the solution to system (6) has the form:

$$\mathbf{y} = (\boldsymbol{\theta} \otimes \mathbf{x}_0, \boldsymbol{\theta} \otimes \mathbf{x}_1, \boldsymbol{\theta} \otimes \mathbf{x}_2). \tag{9}$$

Remember that  $\boldsymbol{\theta}$  is the invariant vector of the MAP underlying process and it satisfies the system

$$\boldsymbol{\theta}(\mathbf{D}_0 + \mathbf{D}_1) = \mathbf{0}, \boldsymbol{\theta}\mathbf{e} = 1. \tag{10}$$

The vectors  $\mathbf{x}_r, r = 0, 1, 2$  have to be defined. Let us substitute the vector  $\mathbf{y}$  in form (9) into the system (6) matrix, which is given by Formula (8). Using the so-called mixed product rule (see properties of the Kronecker product and the sum of matrices given in [23]) and Equation (10), it is possible to verify that this vector satisfies system (6) if the vector  $\mathbf{x} = (\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2)$  satisfies the system

$$\mathbf{x} \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{\Gamma}_0 \otimes \boldsymbol{\tau} \otimes \boldsymbol{\beta} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_K \otimes (\mathbf{S}_0 \otimes \boldsymbol{\beta} + \mathbf{S}) & \mathbf{T}_0 \otimes \mathbf{I}_M \\ \mathbf{S}_0 \otimes \boldsymbol{\gamma} & \mathbf{O} & \mathbf{S} \end{pmatrix} = \mathbf{0}. \tag{11}$$

From normalization condition (7), using the mixed product rule, we easily derive the normalization condition for the vector  $\mathbf{x}$  :

$$\mathbf{x}\mathbf{e} = 1. \tag{12}$$

From Equation (11), we easily obtain the following relations for the sub-vectors  $\mathbf{x}_r, r = 0, 1, 2$  :

$$\mathbf{x}_0 = \mathbf{x}_2 \mathbf{S}_0 \boldsymbol{\gamma} (-\boldsymbol{\Gamma})^{-1}, \tag{13}$$

$$\mathbf{x}_2 = \mathbf{x}_1 (\mathbf{T}_0 \otimes \mathbf{I}_M) (-\mathbf{S})^{-1}, \tag{14}$$

$$\mathbf{x}_1 (\mathbf{I}_K \otimes (\mathbf{S}_0 \otimes \boldsymbol{\beta} + \mathbf{S}) + \mathbf{T} \otimes \mathbf{I}_M) + \mathbf{x}_2 \mathbf{S}_0 (\boldsymbol{\tau} \otimes \boldsymbol{\beta}) = \mathbf{0}. \tag{15}$$

By substituting (14) into (15), we obtain that the vector  $\mathbf{x}_1$  is a solution of the following system

$$\mathbf{x}_1 \mathbf{A} = \mathbf{0}, \tag{16}$$

where the matrix  $\mathbf{A}$  is given by Formula (3).

It is easy to verify that the matrix  $\mathbf{A}$  is an irreducible generator. Therefore, it is singular. The rank of the system is equal to  $KM - 1$ . To obtain the system of a full rank, we use normalization condition (12). Taking into account relations (13)–(15), condition (12) reduces to the equation

$$\mathbf{x}_1 \mathbf{b} = 1, \tag{17}$$

where the column vector  $\mathbf{b}$  is given by Formula (4).

Replacing the first equation of system (16) with Equation (17), we get the equation

$$\mathbf{x}_1 \tilde{\mathbf{A}} = (1, 0, \dots, 0). \tag{18}$$

It can be shown that, if condition (1) is fulfilled, then the matrix  $\tilde{\mathbf{A}}$  is non-singular. Therefore, a solution of system (18) has form (2). Having the vector  $\mathbf{x}_1$  computed, the vectors  $\mathbf{x}_0, \mathbf{x}_2$  are immediately computed from (13) and (14). Thus, we succeeded in explicitly solving systems (11) and (12) for the vector  $\mathbf{x}$  and, then, systems (6) and (7) for the vector  $\mathbf{y}$ . By substituting the vector  $\mathbf{y}$  into (5), performing some algebraic manipulations, we obtain the inequality

$$\mathbf{x}_1 \left[ \mathbf{e}_K \otimes \mathbf{S}_0 + \mathbf{T}_0 \otimes \mathbf{e}_M \right] > \lambda.$$

Taking into account the mixed product rule for the matrices, this inequality can be rewritten in form (1).  $\square$

**Remark 1.** It is well known that the ergodicity of a queueing system means that the system can serve more customers than it receives on average, per unit time conditioned on the fact that the system is overloaded. If the system under study is overloaded, the arrival rate is equal to  $\lambda$  while the rate of customers' departure from the system is equal to the left-hand side of inequality (1). This becomes clearer if this side is rewritten in the equivalent, due to relation (11), form  $\mathbf{x}_1(\mathbf{e}_K \otimes \mathbf{S}_0) + \mathbf{x}_2\mathbf{S}_0$ . The components of the vector  $\mathbf{x}_1$  define the stationary distribution of the underlying processes of the service period and the service time when the state of the process  $r_t$  is 1 and the system is overloaded. The components of the vector  $\mathbf{x}_2$  define the stationary distribution of the underlying process of the service time when the state of the process  $r_t$  is 2 and the system is overloaded. Departures from the system occur only when the state of the process  $r_t$  is 1 or 2. The intensities of departures are defined by the corresponding elements of the vector  $\mathbf{S}_0$  when the states of the underlying process of service time are fixed. Thus, condition (1) is intuitively clear.

**Remark 2.** One of the main performance measures of any queueing system is the throughput, that is, the maximal intensity of flow of customers who obtained successful service. In the system under study, the throughput  $T_{out}$  can be calculated using the following formula:

$$T_{out} = \mathbf{x}_1[\mathbf{e}_K \otimes \mathbf{S}_0 + \mathbf{T}_0 \otimes \mathbf{e}_M] = -\mathbf{x}_1(\mathbf{T} \oplus \mathbf{S})\mathbf{e}_{KM}.$$

**Corollary 1.** If the arrival flow is described by the stationary Poisson process with the rate  $\lambda$ , durations of vacation, maximum attendance time and service time have an exponential distribution with the parameters  $\gamma, \tau$  and  $\mu$ , respectively, and ergodicity condition (1) has the form

$$\lambda < \frac{\tau + \mu}{1 + \frac{\tau}{\mu} + \frac{\tau}{\gamma}}.$$

This condition is easily tractable, as follows. When the system is overloaded, we call the cycle the sequence of three times: vacation time, maximum attendance time, and the residual service time of a customer during service for which the maximum attendance time expires. It is clear that the average length of the cycle is equal to  $\gamma^{-1} + \tau^{-1} + \mu^{-1}$ . Then  $\lambda(\gamma^{-1} + \tau^{-1} + \mu^{-1})$  is the mean number of customers who arrive during the cycle. The mean number of service completions during the maximum attendance time is equal to  $\mu\tau^{-1}$ . One more customer finishes its service after completion of the maximum attendance time. Therefore, the average number of service completions during an arbitrary cycle is equal to  $\mu\tau^{-1} + 1$ . Then, the ergodicity condition requires that the average number of arrivals during the cycle is less than the average number of customer departures during the cycle. This results in the inequality

$$\lambda(\gamma^{-1} + \tau^{-1} + \mu^{-1}) < \mu\tau^{-1} + 1.$$

This coincides with the inequality derived in the considered partial case directly from inequality (1).

If inequality (1) is fulfilled, then the stationary distribution of the Markov chain  $\zeta_t$  exists for any system parameters. Let

$$\pi(i, 0, v, \xi) = \lim_{t \rightarrow \infty} P\{i_t = i, r_t = 0, v_t = v, \xi_t = \xi\}, \quad i \geq 0, v = \overline{0, W}, \xi = \overline{1, R},$$

$$\pi(i, 1, v, \chi, \eta) = \lim_{t \rightarrow \infty} P\{i_t = i, r_t = 1, v_t = v, \chi_t = \chi, \eta_t = \eta\}, \quad i \geq 1, v = \overline{0, W}, \chi = \overline{1, K}, \eta = \overline{1, M},$$

$$\pi(i, 2, v, \eta) = \lim_{t \rightarrow \infty} P\{i_t = i, r_t = 2, v_t = v, \eta_t = \eta\}, \quad i \geq 1, v = \overline{0, W}, \eta = \overline{1, M},$$

be the stationary state probabilities of the Markov chain under study. Then we form the row vectors  $\pi(i, r)$  consisting of state probabilities of the sub-level  $(i, r)$  and the vectors  $\pi_i$  of state probabilities of the level  $i$ :

$$\pi_i = (\pi(i, 0), \pi(i, 1), \pi(i, 2)), \quad i \geq 1, \quad \pi_0 = \pi(0, 0).$$

The following result immediately follows from [14].

**Theorem 2.** The vectors  $\pi_i, i \geq 0$ , defining the stationary distribution of the states of the considered queueing system, are calculated by the following formula:

$$\pi_i = \pi_1 \mathbf{R}^{i-1}, \quad i \geq 1. \tag{19}$$

Here, the matrix  $\mathbf{R}$  can be found as the minimal non-negative solution of the following matrix equation:

$$\mathbf{R}^2 \mathbf{Q}^+ + \mathbf{R} \mathbf{Q}^0 + \mathbf{Q}^- = \mathbf{O}.$$

The vectors  $\pi_0$  and  $\pi_1$  satisfy the following system:

$$\begin{aligned} \pi_0 \mathbf{Q}_{0,0} + \pi_1 \mathbf{Q}_{1,0} &= \mathbf{0}, & \pi_0 \mathbf{Q}_{0,1} + \pi_1 (\mathbf{Q}^0 + \mathbf{R} \mathbf{Q}^-) &= \mathbf{0}, \\ \pi_0 \mathbf{e}_{\overline{WR}} + \pi_1 (\mathbf{I} - \mathbf{R})^{-1} \mathbf{e}_{\overline{WKM}} &= 1. \end{aligned} \tag{20}$$

### 5. Performance Measures of the System

When the vectors  $\pi_i, i \geq 0$ , have been calculated, we can compute different performance characteristics of the system.

The mean number  $L$  of customers in the system is

$$L = \sum_{i=1}^{\infty} i \pi_i \mathbf{e} = \pi_1 (\mathbf{I} - \mathbf{R})^{-2} \mathbf{e}.$$

The mean number  $L_q$  of customers in the queue is calculated as

$$L_q = \sum_{i=1}^{\infty} i \pi(i, 0) \mathbf{e}_{\overline{WR}} + \sum_{i=2}^{\infty} (i-1) \pi(i, 1) \mathbf{e}_{\overline{WKM}} + \sum_{i=2}^{\infty} (i-1) \pi(i, 2) \mathbf{e}_{\overline{WM}}.$$

The probability that the server is in the vacation is defined by

$$F^{(0)} = \sum_{i=0}^{\infty} \pi(i, 0) \mathbf{e}.$$

The probability that the server is busy and the maximum attendance time is not finished is defined by

$$F^{(1)} = \sum_{i=1}^{\infty} \pi(i, 1) \mathbf{e}.$$

The probability that the server is busy but the maximum attendance time is finished is defined by

$$F^{(2)} = \sum_{i=1}^{\infty} \pi(i, 2) \mathbf{e}.$$

The mean number of customers in the system, conditioned on the fact that the server is in the state  $r$ , is defined by

$$L^{(r)} = (F^{(r)})^{-1} \sum_{i=1}^{\infty} i \pi(i, r) \mathbf{e}, \quad r = 0, 1, 2.$$

The probability that a customer meets the server, having status  $r$  upon arrival, is defined by

$$P_r = \lambda^{-1} \sum_{i=1}^{\infty} \pi(i, r) (\mathbf{D}_1 \otimes \mathbf{I}_{K^{2-rM}}) \mathbf{e}, \quad r = 0, 1, 2.$$



The intensity of the output flow of successfully serviced customers in the system  $\lambda^{(out)}$  is defined by

$$\lambda^{(out)} = \sum_{i=1}^{\infty} \pi(i, 1)(\mathbf{e}_{\overline{WK}} \otimes \mathbf{S}_0) + \sum_{i=1}^{\infty} \pi(i, 2)(\mathbf{e}_{\overline{W}} \otimes \mathbf{S}_0).$$

It is evident that, because the customers are absolutely patient, the relation  $\lambda^{(out)} = \lambda$  is true. This relation is useful for the control of the accuracy of the computation of the stationary distribution.

The mean number  $J$  of starting service periods per unit of time is defined by

$$J = \sum_{i=1}^{\infty} \pi(i, 0)(\mathbf{e}_{\overline{W}} \otimes \mathbf{\Gamma}_0).$$

### 6. Distribution of Waiting Time

Let us introduce the following notation:

- $W(x)$  is the waiting time distribution function of an arbitrary customer;
- $w(s)$  is the Laplace–Stieltjes transform (LST) of  $W(x)$ . It is evident that  $w(s)$  can be found as  $w(s) = \int_0^{\infty} e^{-sx} dW(x)$ ,  $Re\ s > 0$ ;
- $\mathbf{w}_i^{(r)}(s)$  is the column vectors of the LSTs of the arbitrary customer waiting time conditioned on the fact that this customer arrives during the epoch when the number of customers in the system is  $i$ ,  $i \geq 0$ , the server is in the  $r$ -th state,  $r = 0, 1, 2$ . The states of other components that define the system’s behavior are assumed to be fixed.

**Theorem 3.** *The LST  $w(s)$  can be found as:*

$$w(s) = \frac{1}{\lambda} \left[ \sum_{i=0}^{\infty} \pi(i, 0)(\mathbf{D}_1 \mathbf{e}_{\overline{W}} \otimes \mathbf{I}_R) \mathbf{w}_i^{(0)}(s) + \sum_{i=1}^{\infty} \pi(i, 1)(\mathbf{D}_1 \mathbf{e}_{\overline{W}} \otimes \mathbf{I}_{KM}) \mathbf{w}_i^{(1)}(s) + \sum_{i=1}^{\infty} \pi(i, 2)(\mathbf{D}_1 \mathbf{e}_{\overline{W}} \otimes \mathbf{I}_M) \mathbf{w}_i^{(2)}(s) \right],$$

where

$$\mathbf{w}_i^{(1)}(s) = (\mathbf{\Omega}(s))^i \mathbf{e}_{KM}, \quad i \geq 0, \tag{21}$$

$$\mathbf{\Omega}(s) = (s\mathbf{I} - \mathbf{T} \oplus \mathbf{S})^{-1} \left[ (\mathbf{T}_0 \otimes \mathbf{I}_M)(s\mathbf{I} - \mathbf{S})^{-1} \mathbf{S}_0 \gamma (s\mathbf{I} - \mathbf{\Gamma})^{-1} \mathbf{\Gamma}_0 (\boldsymbol{\tau} \otimes \boldsymbol{\beta}) + \mathbf{I}_K \otimes (\mathbf{S}_0 \boldsymbol{\beta}) \right] \tag{22}$$

and

$$\mathbf{w}_i^{(0)}(s) = (s\mathbf{I} - \mathbf{\Gamma})^{-1} \mathbf{\Gamma}_0 (\boldsymbol{\tau} \otimes \boldsymbol{\beta}) \mathbf{w}_i^{(1)}(s), \quad i \geq 0, \tag{23}$$

$$\mathbf{w}_i^{(2)}(s) = (s\mathbf{I} - \mathbf{S})^{-1} \mathbf{S}_0 \gamma \mathbf{w}_{i-1}^{(0)}(s), \quad i \geq 1. \tag{24}$$

**Proof.** To derive the formulas for LST  $w(s)$ , we use the method of collective marks, (see, [24,25]). To this end, we tag an arbitrary customer that arrives at the system and keep track of its movement in the system. The method of collective marks assumes that a virtual stationary Poisson flow of catastrophes with the intensity  $s$  arrives at the system. In this case,  $w(s)$  defines the probability of no catastrophe arrival during the stay of the tagged customer in the buffer. Note that a catastrophe has no physical interpretation and cannot impact the system under study. It is used only to obtain the probabilistic meaning of the LST. Note that the entries of the vector  $\mathbf{w}_i^{(r)}(s)$  define the probability of no catastrophe arrival during the tagged customer stay in the buffer, conditioned on the fact that the cus-

tomers arrives when the number of customers in the system is  $i$  and all other components that define the system state have the corresponding values.

Taking into account the probabilistic meaning of the conditional LSTs  $\mathbf{w}_i^{(r)}(s)$  and the formula of total probability, the following expressions for the LSTs  $\mathbf{w}_i^{(r)}(s)$ ,  $r = 0, 1, 2$ , can be derived:

$$\mathbf{w}_i^{(0)}(s) = \int_0^\infty e^{(-s\mathbf{I}+\mathbf{\Gamma})t} \mathbf{\Gamma}_0 dt (\boldsymbol{\tau} \otimes \boldsymbol{\beta}) \mathbf{w}_i^{(1)}(s) = (s\mathbf{I} - \mathbf{\Gamma})^{-1} \mathbf{\Gamma}_0 (\boldsymbol{\tau} \otimes \boldsymbol{\beta}) \mathbf{w}_i^{(1)}(s), \quad i \geq 0, \quad (25)$$

$$\mathbf{w}_i^{(2)}(s) = \int_0^\infty e^{(-s\mathbf{I}+\mathbf{S})t} \mathbf{S}_0 dt \boldsymbol{\gamma} \mathbf{w}_{i-1}^{(0)}(s) = (s\mathbf{I} - \mathbf{S})^{-1} \mathbf{S}_0 \boldsymbol{\gamma} \mathbf{w}_{i-1}^{(0)}(s), \quad i \geq 1, \quad (26)$$

$$\begin{aligned} \mathbf{w}_i^{(1)}(s) &= \int_0^\infty e^{(-s\mathbf{I}+\mathbf{T} \oplus \mathbf{S})t} dt \left[ (\mathbf{T}_0 \otimes \mathbf{I}_M) \mathbf{w}_i^{(2)}(s) + (\mathbf{I}_K \otimes (\mathbf{S}_0 \boldsymbol{\beta})) \mathbf{w}_{i-1}^{(1)}(s) \right] = \\ &= (s\mathbf{I} - \mathbf{T} \oplus \mathbf{S})^{-1} \left[ (\mathbf{T}_0 \otimes \mathbf{I}_M) \mathbf{w}_i^{(2)}(s) + (\mathbf{I}_K \otimes (\mathbf{S}_0 \boldsymbol{\beta})) \mathbf{w}_{i-1}^{(1)}(s) \right], \quad i \geq 1, \quad (27) \end{aligned}$$

$$\mathbf{w}_0^{(1)}(s) = \mathbf{e}_K \otimes \mathbf{e}_M. \quad (28)$$

Solving a system of recursive Equations (25)–(27) with initial condition (28), we obtain Expressions (21)–(24). Then, the expression for the LST  $w(s)$  is easily derived from the formula of total probability.  $\square$

**Remark 3.** It is easy to see that

$$\mathbf{\Omega}(0)\mathbf{e} = \mathbf{e}, \quad \mathbf{w}_i^{(r)}(0) = \mathbf{e}, \quad r = 0, 1, 2.$$

**Corollary 2.** The mean waiting time  $W_1$  of an arbitrary customer can be calculated by the following formulas:

$$\begin{aligned} W_1 &= -w'(s)|_{s=0} = \\ &= \frac{1}{\lambda} \left[ \sum_{i=0}^\infty \pi(i, 0) (\mathbf{D}_1 \mathbf{e}_{\overline{W}} \otimes \mathbf{I}_R) \mathbf{W}_i^{(0)} + \sum_{i=1}^\infty \pi(i, 1) (\mathbf{D}_1 \mathbf{e}_{\overline{W}} \otimes \mathbf{I}_{KM}) \mathbf{W}_i^{(1)} + \sum_{i=1}^\infty \pi(i, 2) (\mathbf{D}_1 \mathbf{e}_{\overline{W}} \otimes \mathbf{I}_M) \mathbf{W}_i^{(2)}, \right] \end{aligned}$$

where the column vectors  $\mathbf{W}_i^{(r)} = -(\mathbf{w}_i^{(r)}(s))'|_{s=0}$ ,  $i \geq 1$ ,  $r = 0, 1, 2$ , are computed as follows:

$$\mathbf{W}_i^{(1)} = -((\mathbf{\Omega}(s))^i)'|_{s=0} \mathbf{e}, \quad i \geq 1,$$

where the derivatives  $((\mathbf{\Omega}(s))^i)'|_{s=0}$  can be computed recursively:

$$((\mathbf{\Omega}(s))^i)'|_{s=0} = ((\mathbf{\Omega}(s))^{i-1})'|_{s=0} \mathbf{\Omega}(0) + (\mathbf{\Omega}(0))^{i-1} (\mathbf{\Omega}(s))'|_{s=0}, \quad i \geq 1,$$

with the initial condition

$$\mathbf{\Omega}'(s)|_{s=0} \mathbf{e} = (\mathbf{T} \oplus \mathbf{S})^{-1} \left[ \mathbf{e}_{KM} + \mathbf{T}_0 \otimes ((\mathbf{I}_{V_1} - \mathbf{S}^{-1}) \mathbf{e}_M) \right]$$

and

$$\begin{aligned} \mathbf{W}_0^{(0)} &= \mathbf{\Gamma}^{-1} \mathbf{e}_R, \quad \mathbf{W}_i^{(0)} = \mathbf{\Gamma}^{-1} \mathbf{e}_R + \mathbf{e}_R (\boldsymbol{\tau} \otimes \boldsymbol{\beta}) \mathbf{W}_i^{(1)}, \quad i \geq 1, \\ \mathbf{W}_i^{(2)} &= \mathbf{S}^{-1} \mathbf{e}_M + \mathbf{e}_M \boldsymbol{\gamma} \mathbf{W}_{i-1}^{(0)}, \quad i \geq 1. \end{aligned}$$

**Corollary 3.** *The mean sojourn time  $V_{soj}$  of an arbitrary customer can be calculated by the following formulas:*

$$V_{soj} = W_1 + b_1.$$

## 7. Numerical Example

Numerical realization of the obtained results does not solve essential problems, compared to [13], because the steady-state distribution of the considered queueing system has the matrix geometric distribution. The importance of accounting for the variance of vacation, service and maximum attendance time, revealed in [13] for the system with impatient customers, takes place in the case of the considered system with absolutely patient customers as well.

In this numerical example, we intend to briefly show the impact of the correlation on the arrival process. We assume that the service time of any customer has the *PH* distribution with the parameters given by the vector  $\beta = (1, 0)$  and sub-generator

$$\mathbf{S} = \begin{pmatrix} -100 & 100 \\ 0 & -100 \end{pmatrix}.$$

The mean service time is  $b_1 = 0.02$ .

The service period (maximum attendance time) has the *PH* distribution with the parameters given by the vector  $\tau = (1, 0)$  and sub-generator

$$\mathbf{T} = \begin{pmatrix} -2 & 2 \\ 0 & -2 \end{pmatrix}.$$

The mean duration of the service period is equal to 1. This means that, on average, fifty customers can be served during the service period if the queue is not exhausted.

The duration of the vacation time has the *PH* distribution with the parameters given by the vector  $\gamma = (1, 0)$  and sub-generator

$$\mathbf{\Gamma} = \begin{pmatrix} -0.5 & 0.5 \\ 0 & -0.5 \end{pmatrix}.$$

The mean duration of the vacation time is equal to 4. In possible applications to the analysis of polling systems, this can be approximately interpreted as the fair sharing of the server between five flows of customers. After the unit of time of the customers receiving service from the tagged queue, the service of this queue is suspended. The service is resumed only after four other queues sequentially receive service from the server.

The aim of this example is to show the importance of accounting for the coefficients of correlation and variation of inter-arrival times. To this end, we consider a *MAP* arrival flow, coded as *MAP*, with the coefficient of correlation  $c_{cor} = 0.2$  and the squared coefficient of variation  $c_{cor} = 12.35$  and a stationary Poisson arrival process, coded as *SPAP*, with the coefficient of correlation  $c_{cor} = 0$  and the squared coefficient of variation  $c_{cor} = 1$ . These processes have the same average arrival intensity  $\lambda$ . We will vary the intensity  $\lambda$  over the interval  $[0.1, 8]$  with the step 0.1 and show the dependencies of the several performance measures of the system on this parameter. To obtain the arrival flow with the average intensity  $\lambda$ , we use the base flow with the average intensity 1, and then multiply the corresponding matrices by  $\lambda$ . The base matrices for the *MAP* are fixed as follows:

$$\mathbf{D}_0 = \begin{pmatrix} -1.35223319 & 0 \\ 0 & -0.0438826 \end{pmatrix},$$

$$\mathbf{D}_1 = \begin{pmatrix} 1.34323164 & 0.00900155 \\ 0.0244342 & 0.0194484 \end{pmatrix}.$$

The base matrices for *SPAP* are the following

$$D_0 = (-1),$$

$$D_1 = (1).$$

Figures 1–5 show the dependence of the main performance measures on the average arrival rate  $\lambda$  for the *MAP* and *SPAP*.

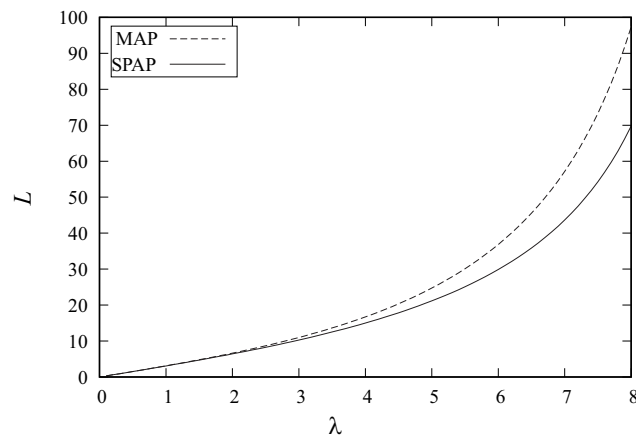


Figure 1. Dependence of the mean number  $L$  of customers in the system on the average arrival rate  $\lambda$ .

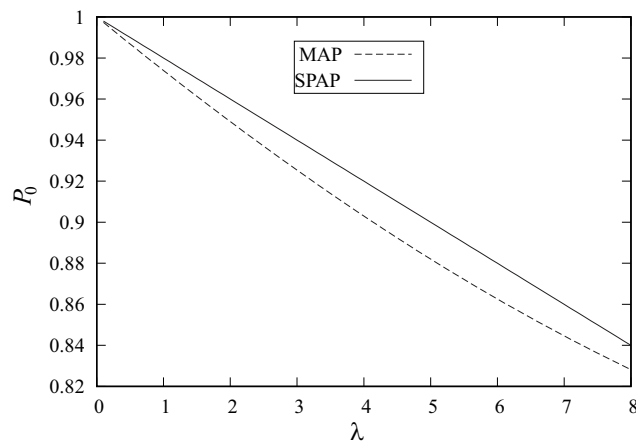
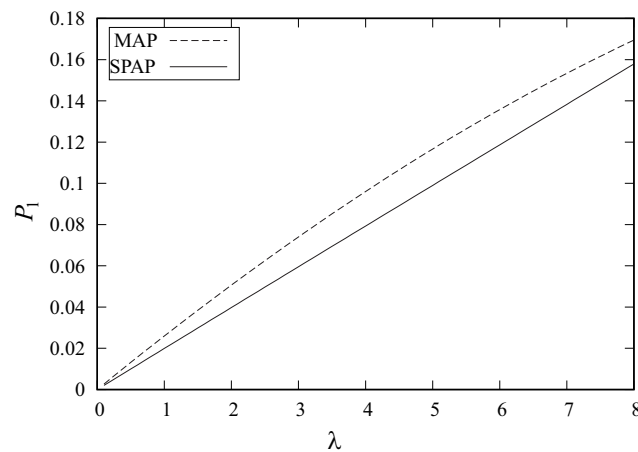
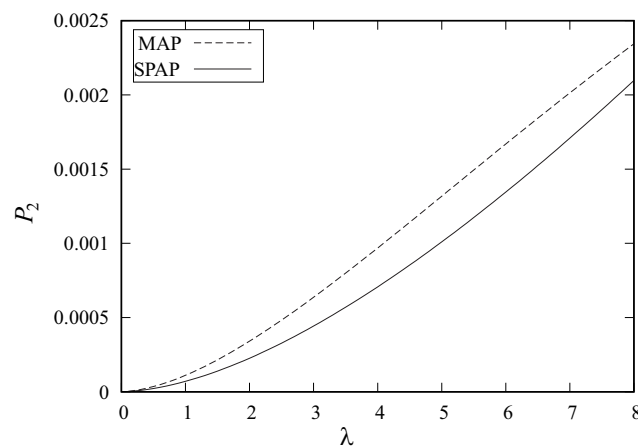


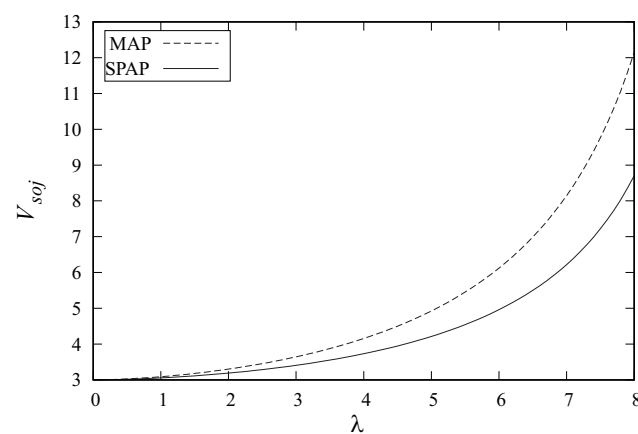
Figure 2. Dependence of the probability  $P_0$  that a customer meets the server, having status 0 (being on vacation) upon arrival, on the average arrival rate  $\lambda$ .



**Figure 3.** Dependence of the probability  $P_1$  that a customer meets the server, having status 1 upon arrival, on the average arrival rate  $\lambda$ .



**Figure 4.** Dependence of the probability  $P_2$  that a customer meets the server, having status 2 upon arrival, on the average arrival rate  $\lambda$ .



**Figure 5.** Dependence of the average sojourn time  $V_{soj}$  on the average arrival rate  $\lambda$ .

It is seen from Figures 1 and 5 that the mean number  $L$  of customers and the average sojourn time  $V_{soj}$  of an arbitrary customer in the system grow when the average arrival rate  $\lambda$  increases. This tendency is evident. Less evident is the fact that the values of  $L$  and  $V_{soj}$  grow more quickly when the arrival flow is the *MAP*. This is explained by the positive correlation of inter-arrival times in the *MAP*. This correlation implies less regular arrival of customers compared to the system with *SPAP*, having the same mean arrival rate as

the *MAP*. Periods of time when customers arrive rarely (and starvation of the server may occur) alternate with periods when customers arrive frequently, which leads to the increase of the queue length and customer sojourn time. These observations are consistent with the behavior of the probabilities  $P_k$ ,  $k = 0, 1, 2$ , illustrated by Figures 2–4. The probability that the server is on vacation is smaller in the case of the *MAP*. A longer queue requires a longer time until the server becomes idle and the probability that the server becomes idle and goes to vacation is smaller compared to the system with non-correlated *SPAP*. Correspondingly, probabilities  $P_1$  and  $P_2$ , that the server is busy, are higher in the case of the *MAP*.

## 8. Conclusions

In this paper, we analyzed a single-server queueing system with vacations. The service period is limited. The arrival flow of customers is defined by the *MAP*. The distributions of the vacation and service times and the maximum server attendance time are of the phase-type. Service is not preemptive: if the maximum server attendance time expires, the currently provided service cannot be interrupted. It has to be performed completely before beginning the vacation. The necessary and sufficient ergodicity condition of the system states is derived. The stationary distribution of the system states, as well as the waiting time distribution of an arbitrary customer, are obtained.

The distributions of the real attendance time of the server, which is the minimum of the busy period of the system, and of the sum of the maximum attendance time and the residual service time of the customer during service for which the maximum attendance time expires, can be found, following [13]. This distribution is interesting from the perspective of the application of the results of the investigation of the considered vacation queueing model to the analysis of polling systems.

**Author Contributions:** Conceptualization, S.D., Y.G. and A.D.; methodology, S.D., V.K. and Y.G.; software, S.D.; validation, S.D. and V.K.; formal analysis, A.D., V.K. and S.D.; investigation, A.D. and V.K.; writing, original draft preparation, Y.G. and A.D.; writing, review and editing A.D. and Y.G.; supervision A.D. and Y.G.; project administration Y.G. and A.D. All authors read and agreed to the published version of the manuscript.

**Funding:** The publication was prepared with the support of the RUDN University Strategic Academic Leadership Program.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

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

## References

1. Borst, S.; Boxma, O. Polling: Past, present, and perspective. *TOP* **2018**, *26*, 335–369. [[CrossRef](#)]
2. Vishnevskii, V.M.; Semenova, O.V. Mathematical methods to study the polling systems. *Autom. Remote Control*. **2006**, *67*, 173–220. [[CrossRef](#)]
3. Vishnevsky, V.; Semenova, O. Polling Systems and Their Application to Telecommunication Networks. *Mathematics* **2021**, *9*, 117. [[CrossRef](#)]
4. Takagi, H. *Queueing Analysis of Polling Models: And update, Stochastic Analysis of Computer and Communication Systems*; Takagi, H., Ed.; North-Holland: Amsterdam, The Netherlands, 1990; pp. 267–318.
5. Takagi, H. *Queueing Analysis of Polling Models: Progress in 1990–1994. Frontiers in Queueing*; Dshalalow, J.H., Ed.; CRC: Boca Raton, FL, USA, 1997; pp. 119–146.
6. Takagi, H. Analysis and applications of polling models. In *Performance Evaluation: Origins and Directions*; Harvig, G., Lindeman, C., Reiser, M., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2000; Volume 1769, pp. 423–442.
7. Takagi, H. *Queueing Analysis: A Foundation of Performance Evaluation*; North-Holland: Amsterdam, The Netherlands, 1991.
8. Tian, N.; Zhang, Z.G. *Vacation Queueing Models: Theory and Applications*; Springer: New York, NY, USA, 2006.
9. Katayama, T. Analysis of a time-limited service priority queueing system with exponential timer and server vacations. *Queueing Syst.* **2007**, *57*, 169–178. [[CrossRef](#)]

10. Katayama, T.; Kobayashi, K. Analysis of a nonpreemptive priority queue with exponential timer and server vacations. *Perform. Eval.* **2007**, *64*, 495–506. [[CrossRef](#)]
11. Leung, K.K. Cyclic-Service Systems with Nonpreemptive, Time-Limited Service. *IEEE Trans. Commun.* **1994**, *42*, 2521–2524. [[CrossRef](#)]
12. Asmussen, S. *Applied Probability and Queues*; Springer: New York, NY, USA, 2003.
13. Kim, C.S.; Dudin, A.; Dudina, O.; Klimenok, V. Analysis of queueing system with non-preemptive time limited service and impatient customers. *Methodol. Comput. Appl. Probab.* **2020**, *22*, 401–432. [[CrossRef](#)]
14. Neuts, M. *Matrix-Geometric Solutions in Stochastic Models*; The Johns Hopkins University Press: Baltimore, MD, USA, 1981.
15. Klimenok, V.I.; Dudin, A.N. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory. *Queueing Syst.* **2006**, *54*, 245–259. [[CrossRef](#)]
16. Chakravarthy, S.R. *The Batch Markovian Arrival Process: A Review and Future Work, Advances in Probability Theory and Stochastic Processes*; Krishnamoorthy, A., Raju, N., Ramaswami, V., Eds.; Notable Publications Inc.: Flint, MI, USA, 2001; pp. 21–29.
17. Dudin, A.N.; Klimenok, V.I.; Vishnevsky, V.M. *The Theory of Queueing Systems with Correlated Flows*; Springer Nature: Berlin/Heidelberg, Germany, 2019; ISBN 978-3-030-32072-0.
18. Heyman, D.P.; Lucantoni, D. Modelling multiple IP traffic streams with rate limits. *IEEE/ACM Trans. Netw.* **2003**, *11*, 948–958. [[CrossRef](#)]
19. Klemm, A.; Lindermann, C.; Lohmann, M. Modelling IP traffic using the batch Markovian arrival process. *Perform. Eval.* **2003**, *54*, 149–173. [[CrossRef](#)]
20. Lucantoni, D. New results on the single server queue with a batch Markovian arrival process. *Commun. Stat. Stochastic Models* **1991**, *7*, 1–46. [[CrossRef](#)]
21. Vishnevski, V.M.; Dudin, A.N. Queueing systems with correlated arrival flows and their applications to modeling telecommunication networks. *Autom. Remote Control* **2017**, *78*, 1361–1403. [[CrossRef](#)]
22. Nazarov, A.; Moiseev, A.; Lapatin, I.; Paul, S.; Lizyura, O.; Pristupa, P.; Peng, X.; Chen, L.; Bai, B. Multi-level MMPP as a Model of Fractal Traffic. *Commun. Comput. Inf. Sci.* **2021**, *1391*, 61–77.
23. Graham, A. *Kronecker Products and Matrix Calculus with Applications*; Ellis Horwood: Chichester, UK, 1981.
24. Kesten, H.; Runnenburg, J.T.; Van Dantzig, D. *Runnenburg, Priority in Waiting Line Problems*; CWI: Amsterdam, The Netherlands, 1956.
25. van Dantzig, D. Chaines de Markof dans les ensembles abstraits et applications aux processus avec regions absorbantes et au probleme des boucles. *Ann. l'Inst. H. Poincare* **1955**, *14*, 145–199.