

GI/Geom/1/N/MWV queue with changeover time and searching for the optimum service rate in working vacation period

Miaomiao Yu^{a,b,*}, Yinghui Tang^{a,**}, Yonghong Fu^c, Lemeng Pan^d

^a School of Mathematics & Software Science, Sichuan Normal University, 610068, Chengdu, China

^b School of Science, Sichuan University of Science and Engineering, 643000, Zigong, China

^c School of Computer, Sichuan University of Science and Engineering, 643000, Zigong, China

^d Department of Mathematics, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Article history:

Received 10 October 2009

Received in revised form 17 February 2010

Keywords:

Working vacation
Discrete-time queue
Finite buffer
Parabolic method
Optimum value

ABSTRACT

In this paper, we consider a finite buffer size discrete-time multiple working vacation queue with changeover time. Employing the supplementary variable and embedded Markov chain techniques, we derive the steady state system length distributions at different time epochs. Based on the various system length distributions, the blocking probability, probability mass function of sojourn time and other performance measures along with some numerical examples have been discussed. Then, we use the parabolic method to search the optimum value of the service rate in working vacation period under a given cost structure.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

In recent years, great attention has been paid to discrete-time queueing systems because of their abundant application in communications. Classical examples are synchronous communication systems (slotted ALOHA) or packet switching systems with time slotting. Currently, ATM multiplexers in the broadband integrated services digital network (B-ISDN), circuit-switched time-division multiple access (TDMA) systems and slotted carrier-sense multiple access (CSMA) protocols have become a powerful incentive to the study of discrete-time queueing systems. For an extensive analysis of a wide variety of discrete-time queueing models, see the books in [1–3].

During the past decade, discrete-time queues with vacations have been widely used in the performance analysis of communication systems. In the classical queueing systems with vacations the authors often assume that the server stops the service completely during the vacation period. Recently, motivated by the analysis of a reconfigurable wavelength-division multiplexing optical access network, Servi and Finn [4] introduced a kind of semi-vacation policy: the server will not completely remain inactive during the vacation period rather it will render service to the queue with a low service rate, when a vacation ends and if there are customers in the queue, a normal busy period begins and the server serves the queue with its original service rate, otherwise on return from a vacation if there is no customer in the queue, the server gives for another vacation and continues to do so till on return from a vacation it finds at least one customer. Such a type of vacation is called a multiple working vacation. To better understand the notion of a working vacation, we may suppose that the queueing system can be staffed with a substitute server during the period that the main server is taking vacations,

* Corresponding author at: School of Mathematics & Software Science, Sichuan Normal University, 610068, Chengdu, China.

** Corresponding author.

E-mail addresses: mmyu75@163.com (M. Yu), tangyh@uestc.edu.cn (Y. Tang).

and the service rate of the substitute server is different from (usually lower than) that of the main server. Servi and Finn studied the $M/M/1/MWV$ queue (MWV stands for multiple working vacation), and obtained the transform formulae for the distributions of the number of customers in the system and the average sojourn time in the steady state. Subsequently, this research was extended to the $GI/M/1/MWV$ queue in [5] and to the $M/G/1/MWV$ queue in [6]. Further, not long ago, Chae et al. [7] studied another important working vacation policy, namely, single working vacation policy, using a novel method, they derived various system performance measures in the steady state. Parallel to the continuous-time working vacation queue, the discrete-time working vacation queue has also been analyzed by many authors. Tian et al. [8] studied the $Geom/Geom/1/MWV$ queue, Li et al. [9] considered the $GI/Geom/1/MWV$ queue and Yi et al. [10] investigated the $Geom/GI/1/MWV$ queue with disasters. More recently, Goswami and Selvaraju [11] studied a discrete-time working vacation queue with a discrete Markovian arrival process.

Although, a lot of working vacation queueing models with infinite buffer size have been studied extensively in the past years, their finite buffer size counterparts received very little attention in the literature. To the best of our knowledge, the only work about a general input working vacation queueing model with finite buffer size can be found in [12,13]. But Refs. [12,13] have been limited to the analysis of a continuous-time queue, very few papers are known about the discrete-time working vacation queue with general input and finite buffer size. Also, in references about a $GI/Geom/1$ working vacation queue (see [9,14]), the system length distributions at arbitrary and an outside observer's observation epoch are not considered. While, as a very important operating characteristic, system length distribution at an outside observer's observation epoch plays a key role in evaluating performance measures. For example, we can employ it to calculate the average sojourn time through Little's rule. Moreover, in literature presented before, authors explore the working vacation queue under the assumption that the times to change the service rate are negligible. But in real life situations, when the system becomes empty, the main server has to remain in the system for a period of time because the substitute server usually does some preparatory work before starting its service to get accustomed to the working circumstances. Based on the problems mentioned above, in this paper, we will consider a $GI/Geom/1/N$ multiple working vacation queue with changeover time. Using both the supplementary variable and embedded Markov Chain techniques [2,15], the system length distributions at arbitrary, pre-arrival and an outside observer's observation epoch with some important performance measures are obtained. Furthermore, it is worth noting that the probability mass function (p.m.f.) and the histogram of an accepted customer's sojourn time are also reported in our paper. With this quantity, we can know the performance measures of the queueing system and improve the quality of service by changing the system parameters.

In order to economize the operating cost, this paper also discusses the optimum value of service rate in the working vacation period. In Ref. [16], under a given cost structure, the authors pointed out that there is an optimum vacation service rate η to minimize the cost, but the value of the optimum η is not given. According to the thought of Refs. [17,18], we develop the steady state expected operating cost function per unit time and employ the parabolic method to find the optimum value of service rate in the working vacation period.

The remainder of this paper is organized as follows. Section 2 presents a description of the model and some notations used in this paper. In Section 3, the supplementary variable and embedded Markov chain techniques are employed to analyze three kinds of system distributions at different time epochs. In Section 4, various performance measures such as the probability of blocking, p.m.f. and the Z transform of the sojourn time etc. are obtained. Further, we also give some numerical results in the form of tables and graphs. Then, in Section 5, we establish a cost structure and consider a cost minimization problem through parabolic method. Finally, in Section 6, we present our conclusions and two topics for future research.

2. Model description and some notations

We consider a finite buffer $GI/Geom/1$ a queue of size N (including the one in service) where the time axis is divided into equal intervals called slots and all queueing activities occur at the slot boundaries. Let the time axis be marked by $0, 1, \dots, t, \dots$. Here, we discuss the early arrival system (EAS) so that a potential customer arrival occurs in (t, t^+) and a potential departure takes place in (t^-, t) . The inter-arrival times $\{T_r, r \geq 1\}$ of customers are independently identically distributed (i.i.d.) random variables with a general p.m.f. $a_m = P\{T_r = m\}$, $m \geq 1$, the corresponding probability generating function (p.g.f.) $a(z) = \sum_{i=1}^{\infty} a_i z^i$ and mean inter-arrival time $1/\lambda$. The service times $\{S_l, l \geq 1\}$ in a normal busy period follow geometric distribution with parameter μ ($0 < \mu < 1$), namely, $P\{S_l = k\} = \mu \bar{\mu}^{k-1}$, $k \geq 1$, where we use symbol $\bar{x} = 1 - x$, for any real number x ($0 < x < 1$). After serving all the customers, the server will remain in the system for a period of time, referred as the changeover time. The changeover time D also follows a geometric distribution with parameter ξ , that is, $P\{D = k\} = \xi \bar{\xi}^{k-1}$, $k \geq 1$. If a customer arrives during the changeover time, the server starts to serve the customer immediately with normal service rate μ . Otherwise, after the changeover time, the server will take a working vacation immediately. The working vacation time V is geometrically distributed with p.m.f. $P\{V = k\} = \theta \bar{\theta}^{k-1}$, $k \geq 1$. During any working vacation the server will serve customers at a rate which is different from the rate of service in normal busy period. The service times $\{\tilde{S}_l, l \geq 1\}$ in working vacation period follow a geometric distribution with parameter η ($0 < \eta < \mu < 1$), namely, $P\{\tilde{S}_l = k\} = \eta \bar{\eta}^{k-1}$, $k \geq 1$. Furthermore, on return from a working vacation if the server finds the system nonempty it will change to another service rate, say the normal service rate, and the service interrupted at the end of vacation restarts from the beginning; otherwise, another working vacation is taken. To make it clear, the various time epochs at which events

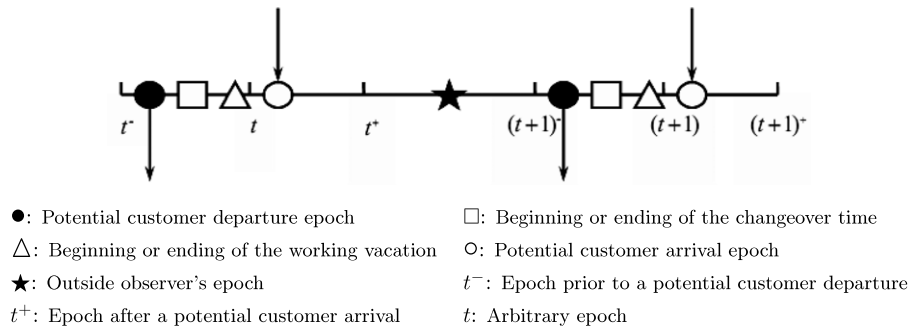


Fig. 1. Various time epochs in an early arrival system.

occur are shown in a self-explanatory figure (see Fig. 1). Finally, we assume that various stochastic processes involved in the system are assumed to be independent of each other.

The state of the system at time t is described by the following random variables:

- N_t : the number of customers in the system (including the one in service) at time t ;
- U_t : remaining inter-arrival time for the next arrival at time t ;

$$Y_t = \begin{cases} 0, & \text{if the server is in a working vacation at time } t, \\ 1, & \text{if the server is in a normal busy period or changeover time at time } t. \end{cases}$$

It is then clear that the joint probabilities of the system length, the remaining inter-arrival time and the state of the server $\{N_t, U_t, Y_t\}$ is a Markov process. Let us define joint probabilities by

$$P_{i,0}(t, u) = P\{N_t = i, U_t = u, Y_t = 0\}, \quad u \geq 0, 0 \leq i \leq N;$$

$$P_{i,1}(t, u) = P\{N_t = i, U_t = u, Y_t = 1\}, \quad u \geq 0, 0 \leq i \leq N.$$

In the steady state, i.e., when $t \rightarrow \infty$, the above probabilities will be denoted by $P_{i,j}(u)$, $i = 0, 1, \dots, N, j = 0, 1$.

3. Analysis of the model

One of the important operating characteristics in a queueing system is the system length distribution. In the following subsections, using different analysis methods, we will consider three kinds of system length at different time epochs, that is, the steady state system length distribution at arbitrary, pre-arrival and an outside observer's observation epoch.

3.1. Steady state system length distribution at an arbitrary epoch

To obtain system length distribution at an arbitrary epoch, we develop the usual Kolmogorov-type difference equations by employing the remaining inter-arrival times as the supplementary variable. Relating the state of the system at time t and $t + 1$, and using probabilistic arguments, we get a set of difference equations:

$$P_{0,0}(t + 1, u - 1) = P_{0,0}(t, u) + \xi P_{0,1}(t, u) + \eta P_{1,0}(t, u) + \eta a_u P_{0,0}(t, 0), \tag{1}$$

$$P_{i,0}(t + 1, u - 1) = \bar{\eta} \bar{\theta} P_{i,0}(t, u) + \eta \bar{\theta} P_{i+1,0}(t, u) + \bar{\eta} \bar{\theta} a_u P_{i-1,0}(t, 0) + \eta \bar{\theta} a_u P_{i,0}(t, 0), \quad i = 1, 2, \dots, N - 2, \tag{2}$$

$$P_{N-1,0}(t + 1, u - 1) = \bar{\eta} \bar{\theta} P_{N-1,0}(t, u) + \eta \bar{\theta} P_{N,0}(t, u) + \bar{\eta} \bar{\theta} a_u P_{N-2,0}(t, 0) + \eta \bar{\theta} a_u P_{N-1,0}(t, 0) + \eta \bar{\theta} a_u P_{N,0}(t, 0), \tag{3}$$

$$P_{N,0}(t + 1, u - 1) = \bar{\eta} \bar{\theta} P_{N,0}(t, u) + \bar{\eta} \bar{\theta} a_u P_{N-1,0}(t, 0) + \bar{\eta} \bar{\theta} a_u P_{N,0}(t, 0), \tag{4}$$

$$P_{0,1}(t + 1, u - 1) = \bar{\xi} P_{0,1}(t, u) + \mu P_{1,1}(t, u) + \mu a_u P_{0,1}(t, 0), \tag{5}$$

$$P_{i,1}(t + 1, u - 1) = \bar{\mu} P_{i,1}(t, u) + \mu P_{i+1,1}(t, u) + \bar{\eta} \bar{\theta} P_{i,0}(t, u) + \eta \bar{\theta} P_{i+1,0}(t, u) + \mu a_u P_{i,1}(t, 0) + \bar{\mu} a_u P_{i-1,1}(t, 0) + \eta \bar{\theta} a_u P_{i,0}(t, 0) + \bar{\eta} \bar{\theta} a_u P_{i-1,0}(t, 0), \quad i = 1, \dots, N - 2, \tag{6}$$

$$P_{N-1,1}(t + 1, u - 1) = \bar{\mu} P_{N-1,1}(t, u) + \mu P_{N,1}(t, u) + \bar{\eta} \bar{\theta} P_{N-1,0}(t, u) + \eta \bar{\theta} P_{N,0}(t, u) + \bar{\mu} a_u P_{N-2,1}(t, 0) + \mu a_u P_{N-1,1}(t, 0) + \mu a_u P_{N,1}(t, 0) + \bar{\eta} \bar{\theta} a_u P_{N-2,0}(t, 0) + \eta \bar{\theta} a_u P_{N-1,0}(t, 0) + \eta \bar{\theta} a_u P_{N,0}(t, 0), \tag{7}$$

$$P_{N,1}(t + 1, u - 1) = \bar{\mu} P_{N,1}(t, u) + \bar{\eta} \bar{\theta} P_{N,0}(t, u) + \bar{\mu} a_u P_{N-1,1}(t, 0) + \bar{\mu} a_u P_{N,1}(t, 0) + \bar{\eta} \bar{\theta} a_u P_{N-1,0}(t, 0) + \bar{\eta} \bar{\theta} a_u P_{N,0}(t, 0). \tag{8}$$

Since we shall discuss our model in steady state, we let $t \rightarrow \infty$ in (1)–(8) and those equations can be written as

$$P_{0,0}(u - 1) = P_{0,0}(u) + \xi P_{0,1}(u) + \eta P_{1,0}(u) + \eta a_u P_{0,0}(0), \tag{9}$$

$$P_{i,0}(u - 1) = \bar{\eta} \bar{\theta} P_{i,0}(u) + \eta \bar{\theta} P_{i+1,0}(u) + \bar{\eta} \bar{\theta} a_u P_{i-1,0}(0) + \eta \bar{\theta} a_u P_{i,0}(0), \quad i = 1, \dots, N - 2, \tag{10}$$

$$P_{N-1,0}(u - 1) = \bar{\eta} \bar{\theta} P_{N-1,0}(u) + \eta \bar{\theta} P_{N,0}(u) + \bar{\eta} \bar{\theta} a_u P_{N-2,0}(0) + \eta \bar{\theta} a_u P_{N-1,0}(0) + \eta \bar{\theta} a_u P_{N,0}(0), \tag{11}$$

$$P_{N,0}(u - 1) = \bar{\eta} \bar{\theta} P_{N,0}(u) + \bar{\eta} \bar{\theta} a_u P_{N-1,0}(0) + \bar{\eta} \bar{\theta} a_u P_{N,0}(0), \tag{12}$$

$$P_{0,1}(u - 1) = \bar{\xi} P_{0,1}(u) + \mu P_{1,1}(u) + \mu a_u P_{0,1}(0), \tag{13}$$

$$P_{i,1}(u - 1) = \bar{\mu} P_{i,1}(u) + \mu P_{i+1,1}(u) + \bar{\eta} \theta P_{i,0}(u) + \eta \theta P_{i+1,0}(u) + \mu a_u P_{i,1}(0) + \bar{\mu} a_u P_{i-1,1}(0) + \eta \theta a_u P_{i,0}(0) + \bar{\eta} \theta a_u P_{i-1,0}(0), \quad i = 1, \dots, N - 2, \tag{14}$$

$$P_{N-1,1}(u - 1) = \bar{\mu} P_{N-1,1}(u) + \mu P_{N,1}(u) + \bar{\eta} \theta P_{N-1,0}(u) + \eta \theta P_{N,0}(u) + \bar{\mu} a_u P_{N-2,1}(0) + \mu a_u P_{N-1,1}(0) + \mu a_u P_{N,1}(0) + \bar{\eta} \theta a_u P_{N-2,0}(0) + \eta \theta a_u P_{N-1,0}(0) + \eta \theta a_u P_{N,0}(0), \tag{15}$$

$$P_{N,1}(u - 1) = \bar{\mu} P_{N,1}(u) + \bar{\eta} \theta P_{N,0}(u) + \bar{\mu} a_u P_{N-1,1}(0) + \bar{\mu} a_u P_{N,1}(0) + \bar{\eta} \theta a_u P_{N-1,0}(0) + \bar{\eta} \theta a_u P_{N,0}(0). \tag{16}$$

Let us define the p.g.f. of $P_{i,j}(u)$ and a_u as follows:

$$P_{i,j}^*(z) = \sum_{u=0}^{\infty} P_{i,j}(u) z^u, \quad i = 0, 1, \dots, N, \quad j = 0, 1, \tag{17}$$

$$a(z) = \sum_{u=1}^{\infty} a_u z^u. \tag{18}$$

As our objective is to obtain $P_{i,j}$ ($i = 0, 1, \dots, N, j = 0, 1$), we have from (17)

$$P_{i,j} \equiv P_{i,j}^*(1) = \sum_{u=0}^{\infty} P_{i,j}(u).$$

Now multiplying (9)–(16) by z^u and summing over u from 1 to ∞ , we obtain after using (17) and (18)

$$(z - 1)P_{0,0}^*(z) = \xi P_{0,1}^*(z) + \eta P_{1,0}^*(z) + \eta a(z)P_{0,0}(0) - P_{0,0}(0) - \xi P_{0,1}(0) - \eta P_{1,0}(0), \tag{19}$$

$$(z - \bar{\eta} \bar{\theta})P_{i,0}^*(z) = \eta \bar{\theta} P_{i+1,0}^*(z) + \bar{\eta} \bar{\theta} a(z)P_{i-1,0}(0) + \eta \bar{\theta} a(z)P_{i,0}(0) - \bar{\eta} \bar{\theta} P_{i,0}(0) - \eta \bar{\theta} P_{i+1,0}(0), \quad i = 1, \dots, N - 2, \tag{20}$$

$$(z - \bar{\eta} \bar{\theta})P_{N-1,0}^*(z) = \eta \bar{\theta} P_{N,0}^*(z) + \bar{\eta} \bar{\theta} a(z)P_{N-2,0}(0) + \eta \bar{\theta} a(z)P_{N-1,0}(0) + \eta \bar{\theta} a(z)P_{N,0}(0) - \bar{\eta} \bar{\theta} P_{N-1,0}(0) - \eta \bar{\theta} P_{N,0}(0), \tag{21}$$

$$(z - \bar{\eta} \bar{\theta})P_{N,0}^*(z) = \bar{\eta} \bar{\theta} a(z)P_{N-1,0}(0) + \bar{\eta} \bar{\theta} a(z)P_{N,0}(0) - \bar{\eta} \bar{\theta} P_{N,0}(0), \tag{22}$$

$$(z - \bar{\xi})P_{0,1}^*(z) = \mu P_{1,1}^*(z) + \mu a(z)P_{0,1}(0) - \bar{\xi} P_{0,1}(0) - \mu P_{1,1}(0), \tag{23}$$

$$(z - \bar{\mu})P_{i,1}^*(z) = \mu P_{i+1,1}^*(z) + \bar{\eta} \theta P_{i,0}^*(z) + \eta \theta P_{i+1,0}^*(z) + \mu a(z)P_{i,1}(0) + \bar{\mu} a(z)P_{i-1,1}(0) + \eta \theta a(z)P_{i,0}(0) + \bar{\eta} \theta a(z)P_{i-1,0}(0) - \bar{\mu} P_{i,1}(0) - \mu P_{i+1,1}(0) - \bar{\eta} \theta P_{i,0}(0) - \eta \theta P_{i+1,0}(0), \quad i = 1, \dots, N - 2, \tag{24}$$

$$(z - \bar{\mu})P_{N-1,1}^*(z) = \mu P_{N,1}^*(z) + \bar{\eta} \theta P_{N-1,0}^*(z) + \eta \theta P_{N,0}^*(z) + \bar{\mu} a(z)P_{N-2,1}(0) + \mu a(z)P_{N-1,1}(0) + \mu a(z)P_{N,1}(0) + \bar{\eta} \theta a(z)P_{N-2,0}(0) + \eta \theta a(z)P_{N-1,0}(0) + \eta \theta a(z)P_{N,0}(0) - \bar{\mu} P_{N-1,1}(0) - \mu P_{N,1}(0) - \bar{\eta} \theta P_{N-1,0}(0) - \eta \theta P_{N,0}(0), \tag{25}$$

$$(z - \bar{\mu})P_{N,1}^*(z) = \bar{\eta} \theta P_{N,0}^*(z) + \bar{\mu} a(z)P_{N-1,1}(0) + \bar{\mu} a(z)P_{N,1}(0) + \bar{\eta} \theta a(z)P_{N-1,0}(0) + \bar{\eta} \theta a(z)P_{N,0}(0) - \bar{\mu} P_{N,1}(0) - \bar{\eta} \theta P_{N,0}(0). \tag{26}$$

Adding (19)–(26) over all possible values of i and j , we obtain after some simplification

$$\sum_{i=0}^N P_{i,0}^*(z) + \sum_{i=0}^N P_{i,1}^*(z) = \frac{a(z) - 1}{z - 1} \left[\sum_{i=0}^N P_{i,0}(0) + \sum_{i=0}^N P_{i,1}(0) \right].$$

Taking limit as $z \rightarrow 1$ in the above term, we have

$$\sum_{i=0}^N P_{i,0}^*(1) + \sum_{i=0}^N P_{i,1}^*(1) = \frac{1}{\lambda} \left[\sum_{i=0}^N P_{i,0}(0) + \sum_{i=0}^N P_{i,1}(0) \right].$$

Then, using the normalization condition $\sum_{i=0}^N P_{i,0}^*(1) + \sum_{i=0}^N P_{i,1}^*(1) = \sum_{i=0}^N P_{i,0} + \sum_{i=0}^N P_{i,1} = 1$, we can get

$$\sum_{i=0}^N P_{i,0}(0) + \sum_{i=0}^N P_{i,1}(0) = \lambda. \tag{27}$$

Let $P_{i,0}^- (0 \leq i \leq N)$ and $P_{i,1}^- (0 \leq i \leq N)$ denote the pre-arrival epoch probabilities, namely, an arrival sees i customers in the system at an arrival epoch when the server is in normal busy period or changeover time and in working vacation period, respectively. Applying Bayes' theorem, we obtain

$$\begin{aligned} P_{i,0}^- &= \lim_{t \rightarrow \infty} P \{N_t = i, Y_t = 0 | U_t = 0\} = \lim_{t \rightarrow \infty} \frac{P \{N_t = i, U_t = 0, Y_t = 0\}}{P \{U_t = 0\}} \\ &= \frac{P_{i,0}(0)}{\sum_{i=0}^N P_{i,1}(0) + \sum_{i=0}^N P_{i,0}(0)}, \quad 0 \leq i \leq N, \end{aligned}$$

$$\begin{aligned} P_{i,1}^- &= \lim_{t \rightarrow \infty} P \{N_t = i, Y_t = 1 | U_t = 0\} = \lim_{t \rightarrow \infty} \frac{P \{N_t = i, U_t = 1, Y_t = 0\}}{P \{U_t = 0\}} \\ &= \frac{P_{i,1}(0)}{\sum_{i=0}^N P_{i,1}(0) + \sum_{i=0}^N P_{i,0}(0)}, \quad 0 \leq i \leq N. \end{aligned}$$

Further, using (27), we have

$$P_{i,0}^- = \frac{1}{\lambda} P_{i,0}(0), \quad 0 \leq i \leq N, \tag{28}$$

$$P_{i,1}^- = \frac{1}{\lambda} P_{i,1}(0), \quad 0 \leq i \leq N. \tag{29}$$

Setting $z = 1$ in Eqs. (20)–(26) and using (28) and (29), after simplification, we can get relations between pre-arrival and arbitrary epoch probabilities

$$P_{N,0} = \frac{\bar{\eta}\bar{\theta}\lambda}{1 - \bar{\eta}\bar{\theta}} P_{N-1,0}^-, \tag{30}$$

$$P_{N-1,0} = \frac{1}{1 - \bar{\eta}\bar{\theta}} [\eta\bar{\theta}P_{N,0} + \bar{\eta}\bar{\theta}\lambda P_{N-2,0}^- + \eta\bar{\theta}\lambda P_{N-1,0}^- - \bar{\eta}\bar{\theta}\lambda P_{N-1,0}^-], \tag{31}$$

$$P_{i,0} = \frac{1}{1 - \bar{\eta}\bar{\theta}} [\eta\bar{\theta}P_{i+1,0} + \bar{\eta}\bar{\theta}\lambda P_{i-1,0}^- + \eta\bar{\theta}\lambda P_{i,0}^- - \bar{\eta}\bar{\theta}\lambda P_{i,0}^- - \eta\bar{\theta}\lambda P_{i+1,0}^-], \quad 1 \leq i \leq N - 2, \tag{32}$$

$$P_{N,1} = \frac{1}{1 - \bar{\mu}} [\bar{\eta}\theta P_{N,0} + \bar{\mu}\lambda P_{N-1,1}^- + \bar{\eta}\theta\lambda P_{N-1,0}^-], \tag{33}$$

$$\begin{aligned} P_{N-1,1} &= \frac{1}{1 - \bar{\mu}} [\mu P_{N,1} + \bar{\eta}\theta P_{N-1,0} + \eta\theta P_{N,0} + \bar{\mu}\lambda P_{N-2,1}^- + \mu\lambda P_{N-1,1}^- + \bar{\eta}\theta\lambda P_{N-2,0}^- + \eta\theta\lambda P_{N-1,0}^- \\ &\quad - \bar{\mu}\lambda P_{N-1,1}^- - \bar{\eta}\theta\lambda P_{N-1,0}^-], \end{aligned} \tag{34}$$

$$\begin{aligned} P_{i,1} &= \frac{1}{1 - \bar{\mu}} [\mu P_{i+1,1} + \bar{\eta}\theta P_{i,0} + \eta\theta P_{i+1,0} + \mu\lambda P_{i,1}^- + \bar{\mu}\lambda P_{i-1,1}^- + \eta\theta\lambda P_{i,0}^- + \bar{\eta}\theta\lambda P_{i-1,0}^- - \bar{\mu}\lambda P_{i,1}^- \\ &\quad - \mu\lambda P_{i+1,1}^- - \bar{\eta}\theta\lambda P_{i,0}^- - \eta\theta\lambda P_{i+1,0}^-], \quad 1 \leq i \leq N - 2, \end{aligned} \tag{35}$$

$$P_{0,1} = \frac{1}{1 - \bar{\xi}} [\mu P_{1,1} + \mu\lambda P_{0,1}^- - \bar{\xi}\lambda P_{0,1}^- - \mu\lambda P_{1,1}^-]. \tag{36}$$

One may note that $P_{0,0}$ can be obtained using the normalizing condition $\sum_{i=0}^N P_{i,1} + \sum_{i=0}^N P_{i,0} = 1$. Meanwhile, we see that through computing the pre-arrival epoch probabilities, the arbitrary time epoch probabilities can be obtained from Eqs. (30)–(36). So, in the next subsection, using the embedded Markov chain technique, we will investigate the system length distribution at a pre-arrival epoch.

3.2. Steady state system length distribution at a pre-arrival epoch

Let L_r be the number of customers in the system just before the r th arrival instant τ_r , and let

$$J_r = J(\tau_r) = \begin{cases} 0, & \tau_r \text{ occurs during the working vacation period,} \\ 1, & \tau_r \text{ occurs during the normal busy period or changeover time.} \end{cases}$$

Thus $\{(L_r, J_r), r \geq 1\}$ is a two-dimensional Markov chain with state space $\Omega = \{(i, j), i = 0, 1, \dots, N, j = 0, 1\}$. Now we develop the transition probabilities of the two-dimensional Markov chain. Let

$$P_{(i,j),(m,n)} = P\{L_{r+1} = m, J_{r+1} = n | L_r = i, J_r = j\}.$$

To obtain the one step transition probability matrix (T.P.M.) \mathbb{P} , first, we introduce some notations:

$$\begin{aligned} \delta(k) &= \begin{cases} 1, & k = 0, \\ 0, & k = 1, 2, \dots, \end{cases} \\ \alpha_h &= \sum_{k=0}^h \sum_{m=h+1+\delta(k)}^{\infty} a_m \sum_{n=\max(1,k)}^{m-(h+1)+k} \theta \bar{\theta}^{n-1} \binom{n}{k} \eta^k \bar{\eta}^{n-k} \sum_{g=h+1-k}^{m-n} \binom{g-1}{h-k} \mu^{h+1-k} \bar{\mu}^{g-(h+1-k)} \bar{\xi}^{m-n-g}, \\ & \quad h = 0, 1, \dots, N-1, \\ \beta_h &= \sum_{m=h+1}^{\infty} a_m \sum_{g=h+1}^m \binom{g-1}{h} \mu^{h+1} \bar{\mu}^{g-(h+1)} \bar{\xi}^{m-g}, \quad h = 0, 1, \dots, N-1, \\ \Gamma_h &= \sum_{m=\max(1,h)}^{\infty} a_m \binom{m}{h} \eta^h \bar{\eta}^{m-h} \bar{\theta}^m, \quad h = 0, 1, \dots, N-1, \\ \Delta_h &= \sum_{k=0}^h \sum_{m=h+\delta(k)}^{\infty} a_m \sum_{n=\max(1,k)}^{m-h+k} \theta \bar{\theta}^{n-1} \binom{n}{k} \eta^k \bar{\eta}^{n-k} \binom{m-n}{h-k} \mu^{h-k} \bar{\mu}^{m-n-(h-k)}, \quad h = 0, 1, \dots, N-1, \\ \sigma_h &= \sum_{m=\max(1,h)}^{\infty} a_m \binom{m}{h} \mu^h \bar{\mu}^{m-h}, \quad h = 0, 1, \dots, N-1. \end{aligned}$$

Remark 1. If $j > i$ then $\sum_{k=j}^i = 0$.

Through probabilistic arguments, some entries $P_{(i,j),(m,n)}$ of the T.P.M. \mathbb{P} are given as follows:

(1) First, the transition from $(i, 0)$ to $(0, 1)$ ($i = 0, 1, \dots, N-1$) occurs if the residual working vacation ends sometime during an inter-arrival time and there are $i+1$ service completions— k ($k = 0, 1, \dots, i$) completions before the working vacation ends and $i+1-k$ after the working vacation ends. Furthermore, the $(r+1)$ th customer arrives at the system during the changeover time. Thus, we have

$$\begin{aligned} P_{(i,0),(0,1)} &= \sum_{k=0}^i P \left\{ \sum_{l=1}^k \tilde{S}_l \leq V < \sum_{l=1}^{k+1} \tilde{S}_l, V + \sum_{l=k+1}^{i+1} S_l \leq T_r < V + \sum_{l=k+1}^{i+1} S_l + D \right\} \\ &= \sum_{k=0}^i \sum_{m=i+1+\delta(k)}^{\infty} a_m P \left\{ \sum_{l=1}^k \tilde{S}_l \leq V < \sum_{l=1}^{k+1} \tilde{S}_l, V + \sum_{l=k+1}^{i+1} S_l \leq m < V + \sum_{l=k+1}^{i+1} S_l + D \right\} \\ &= \sum_{k=0}^i \sum_{m=i+1+\delta(k)}^{\infty} a_m \sum_{n=\max(1,k)}^{m-(i+1)+k} \theta \bar{\theta}^{n-1} \binom{n}{k} \eta^k \bar{\eta}^{n-k} P \left\{ \sum_{l=k+1}^{i+1} S_l \leq m-n < \sum_{l=k+1}^{i+1} S_l + D \right\} \\ &= \sum_{k=0}^i \sum_{m=i+1+\delta(k)}^{\infty} a_m \sum_{n=\max(1,k)}^{m-(i+1)+k} \theta \bar{\theta}^{n-1} \binom{n}{k} \eta^k \bar{\eta}^{n-k} \sum_{g=i+1-k}^{m-n} \binom{g-1}{i-k} \mu^{i+1-k} \bar{\mu}^{g-(i+1-k)} P \{D > m-n-g\} \\ &= \sum_{k=0}^i \sum_{m=i+1+\delta(k)}^{\infty} a_m \sum_{n=\max(1,k)}^{m-(i+1)+k} \theta \bar{\theta}^{n-1} \binom{n}{k} \eta^k \bar{\eta}^{n-k} \sum_{g=i+1-k}^{m-n} \binom{g-1}{i-k} \mu^{i+1-k} \bar{\mu}^{g-(i+1-k)} \bar{\xi}^{m-n-g} \\ &= \alpha_i. \end{aligned}$$

(2) Second, considering a normal busy period, the transition from $(i, 1)$ to $(0, 1)$ ($i = 0, 1, \dots, N-1$) occurs if there are $i+1$ service completions during an inter-arrival time and the $(r+1)$ th customer arrives at the system during the changeover time. Therefore, we have

$$\begin{aligned} P_{(i,1),(0,1)} &= P \left\{ \sum_{l=1}^{i+1} S_l \leq T_r < \sum_{l=1}^{i+1} S_l + D \right\} \\ &= \sum_{m=i+1}^{\infty} a_m P \left\{ \sum_{l=1}^{i+1} S_l \leq m < \sum_{l=1}^{i+1} S_l + D \right\} \end{aligned}$$

$$\begin{aligned}
 &= \sum_{m=i+1}^{\infty} a_m \sum_{g=i+1}^m \binom{g-1}{i} \mu^{i+1} \bar{\mu}^{g-(i+1)} P\{D > m-g\} \\
 &= \sum_{m=i+1}^{\infty} a_m \sum_{g=i+1}^m \binom{g-1}{i} \mu^{i+1} \bar{\mu}^{g-(i+1)} \bar{\xi}^{m-g} \\
 &= \beta_i.
 \end{aligned}$$

(3) Third, the transition from $(i, 0)$ to $(j, 0)$ ($i = 0, 1, \dots, N - 1; j = 1, 2, \dots, i + 1$) occurs if the working vacation time is greater than an inter-arrival time and there are $i + 1 - j$ service completions during an inter-arrival time. So, we have

$$\begin{aligned}
 P_{(i,0),(j,0)} &= P\left\{ \sum_{l=1}^{i+1-j} \tilde{S}_l \leq T_r < \sum_{l=1}^{i+2-j} \tilde{S}_l, T_r < V \right\} \\
 &= \sum_{m=\max(1,i+1-j)}^{\infty} a_m P\left\{ \sum_{l=1}^{i+1-j} \tilde{S}_l \leq m < \sum_{l=1}^{i+2-j} \tilde{S}_l, m < V \right\} \\
 &= \sum_{m=\max(1,i+1-j)}^{\infty} a_m \binom{m}{i+1-j} \eta^{i+1-j} \bar{\eta}^{m-(i+1-j)} \bar{\theta}^m \\
 &= \Gamma_{i+1-j}.
 \end{aligned}$$

(4) Fourth, the transition from $(i, 0)$ to $(j, 1)$ ($i = 0, 1, \dots, N - 1; j = 1, 2, \dots, i + 1$) occurs if there are $i + 1 - j$ service completions during a working vacation time and a customer arrives. Assuming that there are k service completions during a working vacation time and $i + 1 - j - k$ service completions during a normal busy period. Then, we have

$$\begin{aligned}
 P_{(i,0),(j,1)} &= \sum_{k=0}^{i+1-j} P\left\{ \sum_{l=1}^k \tilde{S}_l \leq V < \sum_{l=1}^{k+1} \tilde{S}_l, V + \sum_{l=k+1}^{i+1-j} S_l \leq T_r < V + \sum_{l=k+1}^{i+2-j} S_l \right\} \\
 &= \sum_{k=0}^{i+1-j} \sum_{m=i+1-j+\delta(k)}^{\infty} a_m P\left\{ \sum_{l=1}^k \tilde{S}_l \leq V < \sum_{l=1}^{k+1} \tilde{S}_l, V + \sum_{l=k+1}^{i+1-j} S_l \leq m < V + \sum_{l=k+1}^{i+2-j} S_l \right\} \\
 &= \sum_{k=0}^{i+1-j} \sum_{m=i+1-j+\delta(k)}^{\infty} a_m \sum_{n=\max(1,k)}^{m-(i+1-j)+k} \theta \bar{\theta}^{n-1} \binom{n}{k} \eta^k \bar{\eta}^{n-k} P\left\{ \sum_{l=k+1}^{i+1-j} S_l \leq m-n < \sum_{l=k+1}^{i+2-j} S_l \right\} \\
 &= \sum_{k=0}^{i+1-j} \sum_{m=i+1-j+\delta(k)}^{\infty} a_m \sum_{n=\max(1,k)}^{m-(i+1-j)+k} \theta \bar{\theta}^{n-1} \binom{n}{k} \eta^k \bar{\eta}^{n-k} \binom{m-n}{i+1-j-k} \mu^{i+1-j-k} \bar{\mu}^{m-n-(i+1-j-k)} \\
 &= \Delta_{i+1-j}.
 \end{aligned}$$

(5) At last, considering the normal busy period, transition from $(i, 1)$ to $(j, 1)$ ($i = 0, 1, \dots, N - 1; j = 1, 2, \dots, i + 1$) means that $i + 1 - j$ customers are served during an inter-arrival time. We have

$$\begin{aligned}
 P_{(i,1),(j,1)} &= P\left\{ \sum_{l=1}^{i+1-j} S_l \leq T_r < \sum_{l=1}^{i+2-j} S_l \right\} = \sum_{m=\max(1,i+1-j)}^{\infty} a_m \binom{m}{i+1-j} \mu^{i+1-j} \bar{\mu}^{m-(i+1-j)} \\
 &= \sigma_{i+1-j}.
 \end{aligned}$$

Using the lexicographical sequence for the states, the structure of the T.P.M. \mathbb{P} is given by

$$\mathbb{P} = \begin{pmatrix} A_0 & B_0 & 0 & 0 & \cdots & 0 & 0 \\ A_1 & B_1 & B_0 & 0 & \cdots & 0 & 0 \\ A_2 & B_2 & B_1 & B_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{N-1} & B_{N-1} & B_{N-2} & B_{N-3} & \cdots & B_1 & B_0 \\ A_{N-1} & B_{N-1} & B_{N-2} & B_{N-3} & \cdots & B_1 & B_0 \end{pmatrix},$$

where

$$A_i = \begin{pmatrix} 1 - \alpha_i - \sum_{k=0}^i (\Gamma_k + \Delta_k) & \alpha_i \\ 1 - \beta_i - \sum_{k=0}^i \sigma_k & \beta_i \end{pmatrix}, \quad i = 0, 1, \dots, N - 1,$$

$$B_i = \begin{pmatrix} \Gamma_i & \Delta_i \\ 0 & \sigma_i \end{pmatrix}, \quad i = 0, 1, \dots, N - 1.$$

Let $\mathbf{P}^- = \{P_{0,0}^-, P_{0,1}^-, P_{1,0}^-, P_{1,1}^-, \dots, P_{N-1,0}^-, P_{N-1,1}^-, P_{N,0}^-, P_{N,1}^-\}$ be the row vector of the pre-arrival epoch probabilities. Using one of the available software packages such as MAPLE or Matlab, the stationary probability vector \mathbf{P}^- can be obtained by solving the system equations $\mathbf{P}^- \mathbb{P} = \mathbf{P}^-$ with $\mathbf{P}^- \mathbf{e} = 1$, where \mathbf{e} is a column vector of dimension $2N + 2$ with all entries equal to one.

Remark 2. The system of linear equations

$$\begin{cases} \mathbf{P}^- \mathbb{P} = \mathbf{P}^- \\ \mathbf{P}^- \mathbf{e} = 1 \end{cases}$$

can be directly converted into the following equations

$$\begin{pmatrix} (\mathbb{P} - \mathbf{I}_{2N+2})^T \\ \mathbf{e}^T \end{pmatrix} (\mathbf{P}^-)^T = \begin{pmatrix} \mathbf{0}_{2N+2} \\ 1 \end{pmatrix},$$

where A^T represents the transpose of a matrix or vector A , \mathbf{I}_{2N+2} denotes the identity matrix of dimension $(2N + 2)$ -by- $(2N + 2)$ and $\mathbf{0}_{2N+2}$ is a column vector of dimension $2N + 2$ with all entries equal to zero. Employing a very useful command “linsolve” in Matlab, we can get \mathbf{P}^- immediately.

3.3. Steady state system length distribution at an outside observer’s observation epoch

Steady state system length distribution at an outside observer’s observation epoch plays an important role in evaluating performance measures, for example, in order to use Little’s rule to obtain the average sojourn time in the system, the average number of customers in the queue at an outside observer’s observation epoch is needed.

In EAS, an outside observer’s observation epoch falls in a time interval after a potential customer arrival and before a potential customer departure, the probability $P_{i,1}^o (P_{i,0}^o)$ that an outside observer sees i customers in the system and the server is in a normal busy period or changeover time (in a working vacation period) can be obtained by using the relations

$$P_{N,0} = \bar{\eta} \bar{\theta} P_{N,0}^o, \tag{37}$$

$$P_{i,0} = \bar{\eta} \bar{\theta} P_{i,0}^o + \eta \bar{\theta} P_{i+1,0}^o, \quad i = N - 1, \dots, 1, \tag{38}$$

$$P_{N,1} = \bar{\mu} P_{N,1}^o + \bar{\eta} \theta P_{N,0}^o, \tag{39}$$

$$P_{i,1} = \bar{\mu} P_{i,1}^o + \mu P_{i+1,1}^o + \bar{\eta} \theta P_{i,0}^o + \eta \theta P_{i+1,0}^o, \quad i = N - 1, \dots, 1, \tag{40}$$

$$P_{0,1} = \bar{\xi} P_{0,1}^o + \mu P_{1,1}^o, \tag{41}$$

$$P_{0,0} = P_{0,0}^o + \eta P_{1,0}^o + \xi P_{0,1}^o. \tag{42}$$

The above relations can be obtained by considering arbitrary and an outside observer’s observation epochs in Fig. 1. Solving $P_{i,1}^o (P_{i,0}^o)$ from Eqs. (37)–(42), we have

$$P_{N,0}^o = \frac{1}{\bar{\eta} \bar{\theta}} P_{N,0}, \tag{43}$$

$$P_{i,0}^o = \frac{1}{\bar{\eta} \bar{\theta}} [P_{i,0} - \eta \bar{\theta} P_{i+1,0}^o], \quad i = N - 1, \dots, 1, \tag{44}$$

$$P_{N,1}^o = \frac{1}{\bar{\mu}} [P_{N,1} - \bar{\eta} \theta P_{N,0}^o], \tag{45}$$

$$P_{i,1}^o = \frac{1}{\bar{\mu}} [P_{i,1} - \mu P_{i+1,1}^o - \bar{\eta} \theta P_{i,0}^o - \eta \theta P_{i+1,0}^o], \quad i = N - 1, \dots, 1, \tag{46}$$

$$P_{0,1}^o = \frac{1}{\bar{\xi}} [P_{0,1} - \mu P_{1,1}^o], \tag{47}$$

$$P_{0,0}^o = P_{0,0} - \eta P_{1,0}^o - \xi P_{0,1}^o. \tag{48}$$

4. Performance evaluation

As steady state probabilities at various epochs are known, performance measures of the queueing system can easily be obtained.

4.1. Blocking probability

A particular important measure of the system performance for a finite buffer queueing system is the blocking probability. Let P_{loss} represent the blocking probability of the customer, we have

$$P_{\text{loss}} = P_{N,0}^- + P_{N,1}^- \tag{49}$$

4.2. Sojourn time

In this subsection, we consider sojourn time distribution of an accepted customer in the system under the FCFS queuing discipline. Let us define the random variable W as the sojourn time of an accepted customer in the system and the corresponding p.m.f. $w_m = P\{W = m\}$, $m \geq 1$. Considering various possible cases, the expression of w_m is given below

$$\begin{aligned} w_m &= \frac{1}{1 - P_{\text{loss}}} \left\{ \sum_{i=0}^{N-1} P_{i,1}^- P \left\{ \sum_{j=1}^{i+1} X_j = m \right\} \right. \\ &\quad + P_{0,0}^- \left[P \left\{ \tilde{S}_1 = m | V \geq \tilde{S}_1 \right\} P \left\{ V \geq \tilde{S}_1 \right\} + P \left\{ V + S_1 = m | V < \tilde{S}_1 \right\} P \left\{ V < \tilde{S}_1 \right\} \right] \\ &\quad + \sum_{i=1}^{N-1} P_{i,0}^- P \left\{ \sum_{j=1}^{i+1} \tilde{S}_j = m \mid V \geq \sum_{j=1}^{i+1} \tilde{S}_j \right\} P \left\{ V \geq \sum_{j=1}^{i+1} \tilde{S}_j \right\} \\ &\quad + \sum_{i=1}^{N-1} P_{i,0}^- \sum_{k=0}^i P \left\{ V + \sum_{j=k+1}^{i+1} S_j = m \mid \sum_{j=1}^k \tilde{S}_j \leq V < \sum_{j=1}^{k+1} \tilde{S}_j \right\} P \left\{ \sum_{j=1}^k \tilde{S}_j \leq V < \sum_{j=1}^{k+1} \tilde{S}_j \right\} \Big\} \\ &= \frac{1}{1 - P_{\text{loss}}} \left\{ \sum_{i=0}^{N-1} P_{i,1}^- \binom{m-1}{i} \mu^{i+1} \bar{\mu}^{m-i-1} + P_{0,0}^- \left[\eta \bar{\eta}^{m-1} \bar{\theta}^{m-1} + \sum_{r=1}^{m-1} \theta \bar{\theta}^{r-1} \bar{\eta}^r \mu \bar{\mu}^{m-r-1} \right] \right. \\ &\quad + \sum_{i=1}^{N-1} P_{i,0}^- \binom{m-1}{i} \eta^{i+1} \bar{\eta}^{m-i-1} \bar{\theta}^{m-1} \\ &\quad \left. + \sum_{i=1}^{N-1} P_{i,0}^- \sum_{k=0}^i \sum_{r=\max(1,k)}^{m-(i-k+1)} \theta \bar{\theta}^{r-1} \binom{r}{i-k} \eta^k \bar{\eta}^{r-k} \binom{m-r-1}{i-k} \mu^{i-k+1} \bar{\mu}^{m-r-1-(i-k)} \right\}. \tag{50} \end{aligned}$$

Thus the p.g.f. of the sojourn time is given by

$$\begin{aligned} W(z) &= \sum_{m=1}^{\infty} w_m z^m \\ &= \frac{1}{1 - P_{\text{loss}}} \left\{ \sum_{i=0}^{N-1} P_{i,1}^- \left(\frac{\mu z}{1 - \bar{\mu} z} \right)^{i+1} + P_{0,0}^- \left(\frac{\eta z}{1 - \bar{\eta} \bar{\theta} z} + \frac{\theta \bar{\eta} \mu z^2}{(1 - \bar{\mu} z)(1 - \bar{\eta} \bar{\theta} z)} \right) + \sum_{i=1}^{N-1} P_{i,0}^- \frac{1}{\bar{\theta}} \left(\frac{\eta \bar{\theta} z}{1 - \bar{\eta} \bar{\theta} z} \right)^{i+1} \right. \\ &\quad \left. + \sum_{i=1}^{N-1} P_{i,0}^- \left[\frac{\theta \bar{\eta} \mu^{i+1} z^{i+2}}{(1 - \bar{\mu} z)^{i+1} (1 - \bar{\theta} \bar{\eta} z)} + \sum_{k=1}^i \frac{\theta \bar{\theta}^{k-1} \eta^k z^{i+1}}{(1 - \bar{\theta} \bar{\eta} z)^{k+1}} \left(\frac{\mu}{1 - \bar{\mu} z} \right)^{i+1-k} \right] \right\}. \tag{51} \end{aligned}$$

With the p.g.f. of the sojourn time, we can easily get another performance measure: the expected sojourn time of an accepted customer.

$$\begin{aligned} E[W] &= \frac{1}{1 - P_{\text{loss}}} \left\{ \sum_{i=0}^{N-1} P_{i,1}^- \frac{(i+1)}{\mu} + P_{0,0}^- \left[\frac{\eta}{(1 - \bar{\eta} \bar{\theta})^2} + \frac{\theta \bar{\eta} (1 - \bar{\theta} \bar{\eta} + \mu)}{\mu (1 - \bar{\eta} \bar{\theta})^2} \right] \right. \\ &\quad + \sum_{i=1}^{N-1} P_{i,0}^- \frac{(i+1)}{\eta \bar{\theta}^2} \left(\frac{\eta \bar{\theta}}{1 - \bar{\eta} \bar{\theta}} \right)^{i+2} + \sum_{i=1}^{N-1} P_{i,0}^- \left[\frac{\theta \bar{\eta} [(i+1)(1 - \bar{\theta} \bar{\eta}) + \mu]}{\mu (1 - \bar{\theta} \bar{\eta})^2} \right. \\ &\quad \left. \left. + \sum_{k=1}^i \left(\frac{\theta \bar{\theta}^{k-1} \eta^k [(i+1)(1 - \bar{\theta} \bar{\eta}) + (k+1) \bar{\theta} \bar{\eta}]}{(1 - \bar{\theta} \bar{\eta})^{k+2}} + \frac{(i+1-k) \theta \bar{\theta}^{k-1} \eta^k \bar{\mu}}{\mu (1 - \bar{\theta} \bar{\eta})^{k+1}} \right) \right] \right\}. \tag{52} \end{aligned}$$

4.3. Mean system length at an outside observer's observation epoch

Three kinds of mean system length at an outside observer's observation epoch are given respectively as follows:

Table 1

System length distribution at various epochs when inter-arrival time is arbitrary with $a_1 = 0.35, a_3 = 0.2, a_5 = 0.25, a_{11} = 0.15, a_{15} = 0.05$ and $\mu = 0.24, \theta = 0.018, \eta = 0.18, \xi = 0.15, N = 15$.

i	$P_{i,1}^-$	$P_{i,0}^-$	$P_{i,1}$	$P_{i,0}$	$P_{i,1}^o$	$P_{i,0}^o$
0	0.0301	0.0535	0.0308	0.0525	0.0242	0.0409
1	0.0450	0.0422	0.0456	0.0419	0.0424	0.0444
2	0.0543	0.0333	0.0548	0.0331	0.0527	0.0350
3	0.0594	0.0262	0.0597	0.0261	0.0586	0.0276
4	0.0615	0.0207	0.0617	0.0206	0.0613	0.0218
5	0.0614	0.0163	0.0615	0.0162	0.0615	0.0172
6	0.0598	0.0129	0.0599	0.0128	0.0602	0.0136
7	0.0571	0.0102	0.0572	0.0102	0.0578	0.0107
8	0.0539	0.0081	0.0539	0.0081	0.0546	0.0085
9	0.0503	0.0065	0.0503	0.0064	0.0511	0.0068
10	0.0465	0.0052	0.0465	0.0052	0.0473	0.0055
11	0.0425	0.0043	0.0426	0.0043	0.0435	0.0045
12	0.0382	0.0037	0.0386	0.0037	0.0395	0.0039
13	0.0339	0.0034	0.0344	0.0034	0.0354	0.0035
14	0.0304	0.0034	0.0304	0.0034	0.0312	0.0034
15	0.0227	0.0032	0.0211	0.0030	0.0277	0.0038
Sum	0.7468	0.2532	0.7491	0.2509	0.7491	0.2509

$$P_{\text{loss}} = 0.0259, E[W] = 29.4315, E[W]_{\text{little}} = 29.4315, L^o = 6.2325, L_1^o = 5.2764, L_0^o = 0.9561.$$

Table 2

System length distribution at various epochs when inter-arrival time is geometric with $\lambda = 0.2174$ and $\mu = 0.24, \theta = 0.018, \eta = 0.18, \xi = 0.15, N = 18$.

i	$P_{i,1}^-$	$P_{i,0}^-$	$P_{i,1}$	$P_{i,0}$	$P_{i,1}^o$	$P_{i,0}^o$
0	0.0283	0.0493	0.0283	0.0493	0.0222	0.0385
1	0.0426	0.0387	0.0426	0.0387	0.0395	0.0410
2	0.0514	0.0305	0.0514	0.0305	0.0495	0.0323
3	0.0562	0.0240	0.0562	0.0240	0.0551	0.0254
4	0.0580	0.0188	0.0580	0.0188	0.0576	0.0200
5	0.0578	0.0148	0.0578	0.0148	0.0578	0.0157
6	0.0562	0.0117	0.0562	0.0117	0.0565	0.0123
7	0.0536	0.0092	0.0536	0.0092	0.0541	0.0097
8	0.0504	0.0072	0.0504	0.0072	0.0511	0.0076
9	0.0469	0.0057	0.0469	0.0057	0.0477	0.0060
10	0.0433	0.0045	0.0433	0.0045	0.0441	0.0047
11	0.0397	0.0035	0.0397	0.0035	0.0405	0.0037
12	0.0362	0.0028	0.0362	0.0028	0.0369	0.0030
13	0.0328	0.0023	0.0328	0.0023	0.0335	0.0024
14	0.0296	0.0019	0.0296	0.0019	0.0303	0.0019
15	0.0266	0.0016	0.0266	0.0016	0.0273	0.0017
16	0.0238	0.0015	0.0238	0.0015	0.0244	0.0015
17	0.0212	0.0015	0.0212	0.0015	0.0218	0.0015
18	0.0147	0.0013	0.0147	0.0013	0.0193	0.0017
Sum	0.7694	0.2306	0.7694	0.2306	0.7694	0.2306

$$P_{\text{loss}} = 0.0161, E[W] = 32.5924, E[W]_{\text{little}} = 32.5924, L^o = 6.9718, L_1^o = 6.0887, L_0^o = 0.8831.$$

(1) The average system length:

$$L^o = \sum_{i=1}^N iP_{i,1}^o + \sum_{i=1}^N iP_{i,0}^o. \tag{53}$$

(2) The average system length when the server is in normal busy period:

$$L_1^o = \sum_{i=1}^N iP_{i,1}^o. \tag{54}$$

(3) The average system length when the server is on a working vacation:

$$L_0^o = \sum_{i=1}^N iP_{i,0}^o. \tag{55}$$

4.4. Numerical examples of operating characteristics

To demonstrate the applicability of the formulae obtained in the previous section, in this subsection, we present some numerical examples in the form of tables and graphs. Under some different cases, distribution of the number of customers in the system at pre-arrival, arbitrary and outside observer’s observation epochs are given in Tables 1 and 2. Also, various

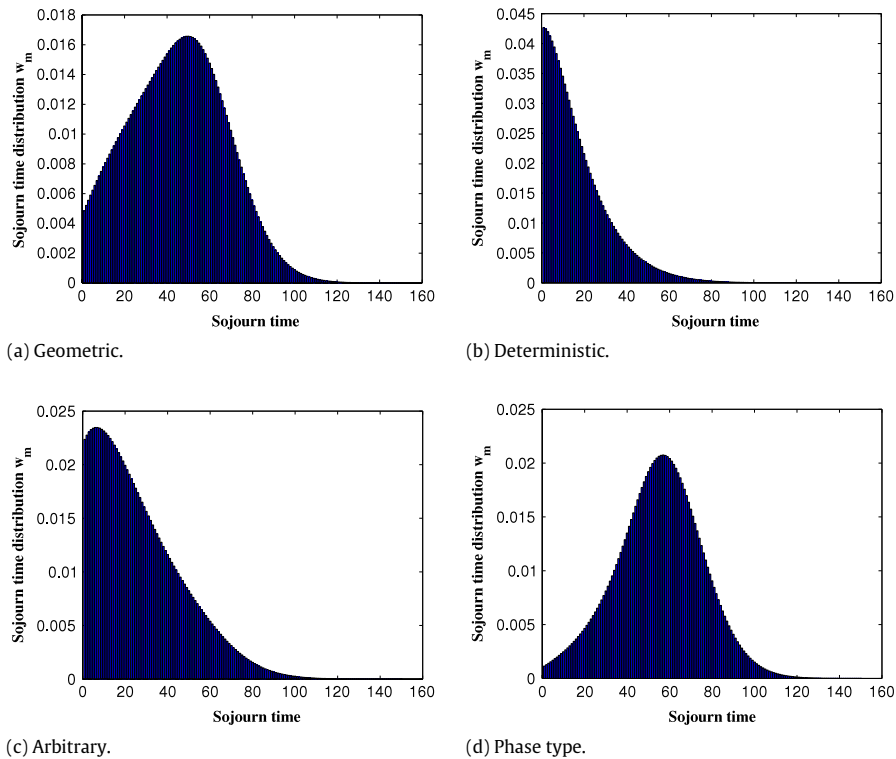


Fig. 2. Histogram of sojourn time for various inter-arrival time distributions.

performance measures such as probability of blocking, three kinds of average system length and average sojourn time in the system are given at the bottom of the tables. All the calculations have been done on the Matlab software package and all the dates are reported here in four decimal places.

Remark 3. The notations used in the tables are the same as those defined earlier in this paper except $E[W]_{\text{little}}$ which denotes the sojourn time in the system of an accepted customer evaluated through Little’s rule.

Remark 4. It can be seen from Table 2 that pre-arrival and arbitrary epoch probabilities are the same (i.e. $P_{ij}^- = P_{ij}$) due to Bernoulli arrivals. It is obvious that the numerical experiment result coincides with the BASTA (i.e. Bernoulli arrivals see time averages) property and indicates the correctness of the theoretical analysis in the previous section and the reliability of the computation program. Furthermore, it is to be noted here that we can also obtain the average sojourn time in the system ($E[W]$) from Little’s rule, $E[W]_{\text{little}} = L^o/\lambda'$, where $\lambda' = \lambda(1 - P_{\text{loss}})$ is the effective arrival rate. By our numerical computation (see the bottom of Tables 1 and 2), we find that the average sojourn time in the system evaluated through Eq. (52) exactly matches with the one obtained from Little’s rule. It also shows that the theoretical analysis and numerical experiment in this paper is correct.

Further, using Eq. (50), we have also got the sojourn time distribution of an accepted customer in the system. In Table 3, the computation results of the sojourn time distribution are given for various inter-arrival time distributions with the same system parameters $N = 17, \mu = 0.24, \eta = 0.18, \xi = 0.15, \theta = 0.018$. The inter-arrival time distributions are taken as geometric ($\lambda = 0.25$), deterministic ($a_5 = 1$), arbitrary ($a_2 = 0.6, a_6 = 0.3, a_{20} = 0.1$) and phase type ($\alpha = (0.3, 0.1, 0.6), T = \begin{bmatrix} 0.2 & 0.3 & 0.2 \\ 0.5 & 0.3 & 0 \\ 0.3 & 0 & 0.4 \end{bmatrix}$), respectively. Also, for the above four cases, the probability mass functions of the sojourn time are illustrated in Fig. 2(a)–(d). Moreover, we note that an interesting fact, that is, the probability mass function of the sojourn time becomes very “normal” when the inter-arrival time distribution is taken as a phase type (see Fig. 2(d)).

Fig. 3a compares the buffer size versus blocking probability for various inter-arrival time distributions with the same parameters $\mu = 0.24, \eta = 0.18, \theta = 0.018, \xi = 0.15$. The three curves exhibit a decreasing shape. It can be seen that deterministic distribution yields the lowest blocking probability. Then, under three different inter-arrival time distributions and the same parameters mentioned above, Fig. 3b shows the effect of buffer size on the average sojourn time $E[W]$. We observe that, for each case, the average sojourn time reaches its asymptotic value after certain buffer size N .

In Fig. 4, assume that the inter-arrival time distribution is arbitrary ($a_1 = 0.35, a_3 = 0.2, a_5 = 0.25, a_{11} = 0.15, a_{15} = 0.05$), we display the blocking probability and the expected value of sojourn time as a function of buffer size N . As is to be

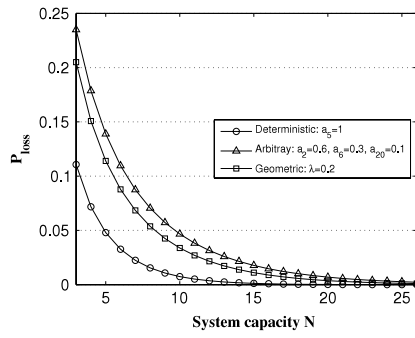


Fig. 3a. Effect of N on blocking probability under different inter-arrival time distributions.

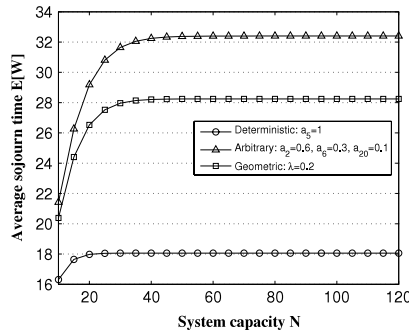


Fig. 3b. Effect of N on average sojourn time under different inter-arrival time distributions.

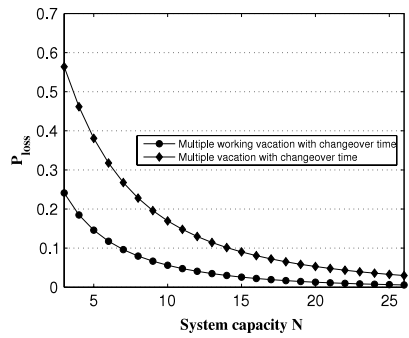


Fig. 4a. Effect of N on blocking probability under different vacation policies.

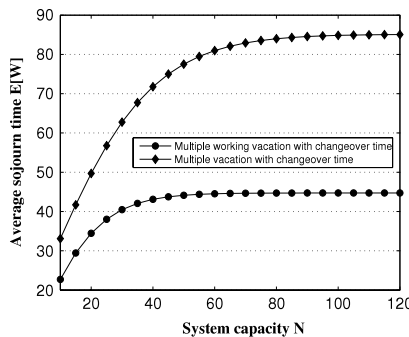


Fig. 4b. Effect of N on average sojourn time under different vacation policies.

expected, here the blocking probability and the average sojourn time are less for a working vacation model in comparison to a non-working vacation.

Table 3

Partial summation of the sojourn time distribution for various inter-arrival time distributions.

$\sum_{m=i}^j w_m$	Geometric	Deterministic	Arbitrary	Phase type
$\sum_{m=1}^{10} w_m$	0.0637	0.3941	0.2316	0.0178
$\sum_{m=11}^{20} w_m$	0.0934	0.2728	0.2157	0.0359
$\sum_{m=21}^{30} w_m$	0.1193	0.1603	0.1759	0.0642
$\sum_{m=31}^{40} w_m$	0.1437	0.0870	0.1338	0.1092
$\sum_{m=41}^{50} w_m$	0.1622	0.0450	0.0970	0.1680
$\sum_{m=51}^{60} w_m$	0.1587	0.0224	0.0663	0.2041
$\sum_{m=61}^{70} w_m$	0.1241	0.0106	0.0409	0.1821
$\sum_{m=71}^{80} w_m$	0.0760	0.0047	0.0220	0.1203
$\sum_{m=81}^{90} w_m$	0.0370	0.0019	0.0102	0.0613
$\sum_{m=91}^{100} w_m$	0.0148	7.3405×10^{-4}	0.0041	0.0251
$\sum_{m=101}^{110} w_m$	0.0051	2.6136×10^{-4}	0.0015	0.0086
$\sum_{m=111}^{120} w_m$	0.0015	8.8319×10^{-5}	5.0576×10^{-4}	0.0025
$\sum_{m=121}^{130} w_m$	4.0576×10^{-4}	2.8636×10^{-5}	1.6016×10^{-4}	6.6641×10^{-4}
$\sum_{m=131}^{140} w_m$	1.0119×10^{-4}	8.9636×10^{-6}	4.8785×10^{-5}	1.5772×10^{-4}
Sum	1.0000	1.0000	1.0000	1.0000

5. The optimum service rate in a working vacation period

In practice, queueing managers are always interested in minimizing operating cost of unit time. In this section, we first develop a steady state expected operating cost function per unit time for the *GI/Geom/1/N* working vacation queue with changeover time, in which η is a decision variable. Our main objective in this section is to determine the optimum value of service rate in working vacation period, so as to minimize the expected operating cost function per unit time.

5.1. Operating cost function per unit time

Let us define the following cost elements:

- C_{WV} \equiv unit time cost of every customer in the system when the server is on a working vacation;
- C_{NB} \equiv unit time cost of every customer in the system when the server is in a normal busy period;
- C_{μ} \equiv fixed service cost per unit time during the normal busy period;
- C_{η} \equiv fixed service cost per unit time during a working vacation period;
- C_{ξ} \equiv fixed cost per unit time during the changeover time.

Based on the definitions of each cost element listed above and its corresponding system performance measures, the cost minimization problem can be illustrated mathematically as:

$$\min_{\eta} : TC(\eta) = C_{WV}L_0^o + C_{NB}L_1^o + C_{\mu}\mu + C_{\eta}\eta + C_{\xi}\xi. \tag{56}$$

It would have been a difficult task to develop analytic results for the optimum value of η because the expected operating cost function is highly non-linear and complex. In addition, we note that the derivatives of operating cost function per unit time are not easily available, because $TC(\eta)$ is provided only in the form of computer program code, that is, the analytic form of $TC(\eta)$ unknown. To solve the optimization problem (56), in the next subsection, we will utilize a kind of direct method to find the optimum value of η , say η^* .

5.2. The parabolic method and some numerical examples

Direct methods are employed when the objective function f is not differentiable or if the computation of its derivatives is a nontrivial task. For the reasons mentioned above, we employ the parabolic method to solve the optimization problem (56).

The essence of the parabolic method is to generate a quadratic function through the evaluated points in each iteration, and then the objective function $f(x)$ is approximated by the quadratic function in generating an estimate of the optimum value. According to the polynomial approximation theory, the unique optimum of the quadratic function agreeing with $f(x)$ at 3-point pattern $\{x^{(lo)}, x^{(mid)}, x^{(hi)}\}$ occurs at

$$x^{(qu)} \triangleq \frac{1}{2} \frac{f^{(lo)}[s^{(mid)} - s^{(hi)}] + f^{(mid)}[s^{(hi)} - s^{(lo)}] + f^{(hi)}[s^{(lo)} - s^{(mid)}]}{f^{(lo)}[x^{(mid)} - x^{(hi)}] + f^{(mid)}[x^{(hi)} - x^{(lo)}] + f^{(hi)}[x^{(lo)} - x^{(mid)}]} [16], \tag{57}$$

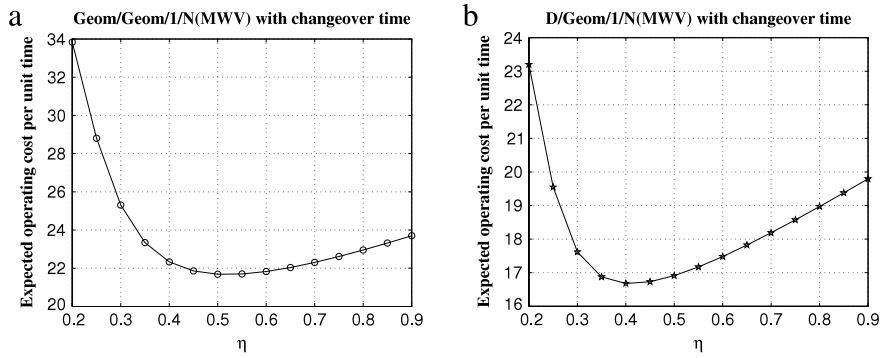


Fig. 5. Effect of η on expected operating cost per unit time.

where $f^{(lo)} \triangleq f(x^{(lo)})$, $f^{(mid)} \triangleq f(x^{(mid)})$, $f^{(hi)} \triangleq f(x^{(hi)})$, $s^{(lo)} \triangleq (x^{(lo)})^2$, $s^{(mid)} \triangleq (x^{(mid)})^2$ and $s^{(hi)} \triangleq (x^{(hi)})^2$. The parabolic method uses this approximation to improve the current 3-point pattern by replacing one of its points with an approximate optimum $x^{(qu)}$. Then, repeating in this way isolates an optimum for $f(x)$ in an ever-narrowing range.

The steps of the parabolic method can be described as follows (see Ref. [19]):

Step 1: (Initialization). Choose a starting 3-point pattern $\{x^{(lo)}, x^{(mid)}, x^{(hi)}\}$ along with a stopping tolerance $\varepsilon = 10^{-5}$, and initialize the iteration counter $i = 0$;

Step 2: (Stopping). If $|x^{(qu)} - x^{(mid)}| \leq \varepsilon$, stop and report approximate optimum solution $x^{(mid)}$;

Step 3: (Quadratic Fit). Compute a quadratic fit optimum $x^{(qu)}$ according to the formula (57). Then if $x^{(qu)} \leq x^{(mid)}$, go to Step 4; and if $x^{(qu)} > x^{(mid)}$ go to Step 5.

Step 4: (Left). If $f(x^{(mid)})$ is superior to $f(x^{(qu)})$ (less for a minimize, greater for a maximize), then update

$$x^{(qu)} \rightarrow x^{(lo)}.$$

Otherwise, replace

$$\begin{aligned} x^{(mid)} &\rightarrow x^{(hi)}, \\ x^{(qu)} &\rightarrow x^{(mid)}. \end{aligned}$$

Either way, advance $i = i + 1$, and return to Step 2.

Step 5: (Right). If $f(x^{(mid)})$ is superior to $f(x^{(qu)})$ (less for a minimize, greater for a maximize), then update

$$x^{(qu)} \rightarrow x^{(hi)}.$$

Otherwise, replace

$$\begin{aligned} x^{(mid)} &\rightarrow x^{(lo)}, \\ x^{(qu)} &\rightarrow x^{(mid)}. \end{aligned}$$

Either way, advance $i = i + 1$, and return to Step 2.

Next, to demonstrate the applicability of the parabolic method in our optimization problem, two numerical examples under different cases are provided.

Example 1 (Geom/Geom/1/N). Assume that the inter-arrival time of a customer is geometric with $\lambda = 0.25$. Other parameters of the system are given as follows: $N = 17$, $\mu = 0.67$, $\theta = 0.018$, $\xi = 0.15$, $C_{WV} = 5$, $C_{NB} = 3$, $C_{\mu} = 18$, $C_{\eta} = 10$, $C_{\xi} = 8$ and the stopping tolerance $\varepsilon = 10^{-5}$.

With the information of Fig. 5(a), we select the initial 3-point pattern $\eta^{(lo)} = 0.4$, $\eta^{(mid)} = 0.5$, $\eta^{(hi)} = 0.6$, and then we apply the parabolic method as mentioned above, after four iterations. Table 4 clearly shows that the minimum expected operating cost per unit time converges to the solution $\eta^* = 0.520160$ with value 21.674583.

Example 2 (D/Geom/1/N). Assume that the inter-arrival time of a customer is deterministic with $a_4 = 1$. Other parameters of the system are given as follows: $N = 19$, $\mu = 0.5$, $\theta = 0.02$, $\xi = 0.05$, $C_{WV} = 5$, $C_{NB} = 3$, $C_{\mu} = 18$, $C_{\eta} = 10$, $C_{\xi} = 8$ and the stopping tolerance $\varepsilon = 10^{-5}$.

Similarly, with the information of Fig. 5(b), we select the initial 3-point pattern $\eta^{(lo)} = 0.35$, $\eta^{(mid)} = 0.4$, $\eta^{(hi)} = 0.45$, and also use the parabolic method, after five iterations, Table 5 clearly shows that the minimum expected operating cost per unit time converges to the solution $\eta^* = 0.409928$ with a value 16.673579.

Table 4The parabolic method in searching the optimum solution of $Geom/Geom/1/N$ queue.

No. of iterations	$\eta^{(lo)}$	$\eta^{(mid)}$	$\eta^{(hi)}$	$TC(\eta^{(lo)})$	$TC(\eta^{(mid)})$	$TC(\eta^{(hi)})$	$\eta^{(qu)}$	$TC(\eta^{(qu)})$	Tolerance
0	0.400000	0.500000	0.600000	22.334827	21.687497	21.829058	0.532056	21.678685	0.032056
1	0.500000	0.532056	0.600000	21.687497	21.678685	21.829058	0.521552	21.674641	0.010503
2	0.500000	0.521552	0.532056	21.687497	21.674641	21.678685	0.520517	21.674587	0.001035
3	0.500000	0.520517	0.521552	21.687497	21.674587	21.674641	0.520199	21.674583	3.185107×10^{-4}
4	0.500000	0.520199	0.520517	21.687497	21.674583	21.674587	0.520160	21.674583	3.924987×10^{-5}

Table 5The parabolic method in searching the optimum solution of $D/Geom/1/N$ queue.

No. of iterations	$\eta^{(lo)}$	$\eta^{(mid)}$	$\eta^{(hi)}$	$TC(\eta^{(lo)})$	$TC(\eta^{(mid)})$	$TC(\eta^{(hi)})$	$\eta^{(qu)}$	$TC(\eta^{(qu)})$	Tolerance
0	0.350000	0.400000	0.450000	16.876396	16.677864	16.730082	0.414588	16.674463	0.014588
1	0.400000	0.414588	0.450000	16.677864	16.674463	16.730082	0.410525	16.673594	0.004063
2	0.400000	0.410525	0.414588	16.677864	16.673594	16.674463	0.410038	16.673580	4.866487×10^{-4}
3	0.400000	0.410038	0.410525	16.677864	16.673580	16.673594	0.409943	16.673579	9.540400×10^{-5}
4	0.400000	0.409943	0.410038	16.677864	16.673579	16.673580	0.409930	16.673579	1.280382×10^{-5}
5	0.400000	0.409930	0.409943	16.677864	16.673579	16.673579	0.409928	16.673579	2.401686×10^{-6}

6. Conclusions

In this paper, we have carried out an analysis of a finite buffer discrete-time queue that has potential applications in communication networks. Through two kinds of classical analysis techniques, we have obtained the system length distributions at different time epochs. Utilizing these distributions, we have derived some important performance characteristics and also have considered the cost optimization problem by a numerical method. In the future, research topics such as $GI/D - MSP/1/N/MWV$ and $GI^x/Geom/1/N/MWV$ discrete-time queues can be studied with the same analysis techniques.

Acknowledgements

The authors acknowledge anonymous reviewers for their comments which were very helpful in improving the presentation of this paper. This work was partly supported by the 6th science and technology innovation foundation for the Sichuan normal university students. This research is supported by the National Natural Science Foundation of China (No. 70871084), the Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 200806360001) and the Natural Science Foundation of the Education Department of Sichuan Province (08ZC028).

References

- [1] H. Bruneel, B.G. Kim, Discrete-Time Models for Communication Systems Including ATM, Kluwer, Boston, 1993.
- [2] N.S. Tian, X.L. Xu, Z.Y. Ma, Discrete-Time Queueing Theory, Science Press, Beijing, 2008 (in Chinese).
- [3] H. Takagi, Queueing Analysis: A Foundation of Performance Evaluation, vol. 3, North-Holland, Amsterdam, 1993.
- [4] L.D. Servi, S.G. Finn, $M/M/1$ queues with working vacations ($M/M/1/WV$), Performance Evaluation 50 (2002) 41–52.
- [5] Y. Baba, Analysis of a $GI/M/1$ queue with multiple working vacations, Operations Research Letters 33 (2005) 201–209.
- [6] D. Wu, H. Takagi, $M/G/1$ queue with multiple working vacations, Performance Evaluation 63 (2006) 654–681.
- [7] K.C. Chae, D.E. Lim, W.S. Yang, The $GI/M/1$ queue and the $GI/Geo/1$ queue both with single working vacation, Performance Evaluation 66 (2009) 356–367.
- [8] N.S. Tian, Z.Y. Ma, M.X. Liu, The discrete-time $Geom/Geom/1$ queue with multiple working vacations, Applied Mathematical Modelling 32 (2008) 2941–2953.
- [9] J.H. Li, N.S. Tian, W.Y. Liu, Discrete-time $GI/Geo/1$ queue with multiple working vacations, Queueing System 56 (2007) 53–63.
- [10] X.W. Yi, J.D. Kim, D.W. Choi, K.C. Chae, The $Geo/G/1$ queue with disasters and multiple working vacations, Stochastic Models 23 (2007) 537–549.
- [11] C. Goswami, N. Selvaraju, The discrete-time $MAP/PH/1$ queue with multiple working vacations, Applied Mathematical Modelling 34 (2010) 931–946.
- [12] A.D. Banik, U.C. Gupta, S.S. Pathak, On the $GI/M/1/N$ queue with multiple working vacations—analytic analysis and computation, Applied Mathematical Modelling 31 (2007) 1701–1710.
- [13] M.M. Yu, Y.H. Tang, Y.H. Fu, Steady state analysis and computation of the $GI^{[k]}/M^b/1/L$ queue with multiple working vacations and partial batch rejection, Computers & Industrial Engineering 56 (2009) 1243–1253.
- [14] J.H. Li, N.S. Tian, The discrete time $GI/Geo/1$ queue with working vacations and vacation interruption, Applied Mathematics and Computation 185 (2007) 1–10.
- [15] M.L. Chaudhry, J.G.C. Templeton, U.C. Gupta, Analysis of the discrete-time $GI/Geom(n)/1/N$ queue, Computers & Mathematics with Applications 31 (1996) 59–68.
- [16] J.H. Li, N.S. Tian, Z.Y. Ma, Performance analysis of $GI/M/1$ queue with working vacations and vacation interruption, Applied Mathematical Modelling 32 (2008) 2715–2730.
- [17] K.H. Wang, D.Y. Yang, Controlling arrivals for a queueing system with an unreliable server: Newton-quasi method, Applied Mathematics and Computation 231 (2009) 92–101.
- [18] K.H. Wang, W.L. Chen, D.Y. Yang, Optimal management of the machine repair problem with working vacation: Newtons method, Journal of Computational and Applied Mathematics 233 (2009) 449–458.
- [19] L.R. Ronald, Optimization in Operations Research, Prentice Hall, New Jersey, 1997.