

Analysis of a model of batch arrival single server queue with random vacation policy

Choudhury Gautam, Kalita Priyanka & Selvamuthu Dharmaraja

To cite this article: Choudhury Gautam, Kalita Priyanka & Selvamuthu Dharmaraja (2021) Analysis of a model of batch arrival single server queue with random vacation policy, Communications in Statistics - Theory and Methods, 50:22, 5314-5357, DOI: [10.1080/03610926.2020.1728329](https://doi.org/10.1080/03610926.2020.1728329)

To link to this article: <https://doi.org/10.1080/03610926.2020.1728329>



Published online: 22 Feb 2020.



Submit your article to this journal [↗](#)



Article views: 105



View related articles [↗](#)



View Crossmark data [↗](#)



Analysis of a model of batch arrival single server queue with random vacation policy

Choudhury Gautam^a, Kalita Priyanka^a, and Selvamuthu Dharmaraja^b 

^aMathematical Sciences Division, Institute of Advanced Study in Science and Technology, Guwahati, Assam, India; ^bIndian Institute of Technology Delhi, New Delhi, India

ABSTRACT

This article deals with batch arrival single server queue with random vacation policy, in which the server takes the maximum number of random vacations till it finds some customers waiting in a queue at a vacation completion epoch. If no arrival occurs after completing maximum number of random vacations, the server stays dormant in the system and waits for the upcoming arrival. Here, it is assumed that the customer's arrival in batches conforms to compound Poisson process, although service time and vacation time are generally distributed. Explicit expressions are obtained for steady state queue size distribution at service completion point and steady state system size probabilities. Some significant measures such as a mean number of batches served during the busy period, Laplace-Stieltjes transform of waiting time, unfinished work and its corresponding mean values are also obtained. A cost optimal policy is developed in terms of the average cost function to determine a locally optimal random vacation policy at a lower cost. Finally, various numerical results are presented for the above system performance measures.

ARTICLE HISTORY

Received 14 May 2019

Accepted 6 February 2020

KEYWORDS

$M^X/G/1$ queue; random vacation policy; embedded Markov chain; Markov regenerative process; cost optimal policy

MATHEMATICS SUBJECT CLASSIFICATION (2000)

60K25; 90B22; 60K20

1. Introduction

Queueing models with server vacations have been effectively studied by many researchers in the last few decades and successfully applied in the various problems such as computer communication networks, modern telecommunication systems, production, manufacturing systems, etc. An excellent survey on the vacation queueing models have been documented in Doshi (1986, 1990), Takagi (1991), Tian and Zhang (2006) and others. There are two basic vacation queueing models namely, multiple vacation queueing model and single vacation queueing model. In multiple vacation queueing models, the server keeps on taking consecutive vacations until it finds some customers waiting in a queue at a vacation completion epoch. However, in single vacation queueing models, the server takes exactly one vacation between two consecutive busy periods. These two types of vacation models were first introduced by Levy and Yechiali (1975).

Batch arrival in a queueing systems is common as well as important in many practical situations for an effective and efficient utilization of resources. Baba (1996) was first to study the batch arrival queueing system with multiple vacation policy. He discussed the

queue size distribution, waiting time distribution and busy period distribution for an $M^X/G/1$ queueing system with multiple vacations using the well known supplementary variable technique. Later, Bratiichuk (2000) derived the probability generating function (PGF) of a number of customers served during a busy period, the exact formula for a busy period and queue size distribution of an $M^X/G/1$ type queueing system but with a finite buffer capacity. Choudhury (2002a) studied a batch arrival queueing system with single vacation policy and obtained an explicit expression for steady state queue size distribution at different points in time, as well as steady state distributions of the busy period and unfinished work. Ke, Huang, and Chu (2010) investigated batch arrival queue with N-policy and atmost J vacations. If the system is empty by the end of J^{th} vacation, the server remains idle in the system until the number of arrivals in the queue reaches N . Kempa (2011, 2014) also investigated these types of batch arrival queueing systems with both multiple and single vacation policies. In the same paper, the departure process in transient state and Laplace transform of the queue size distribution are also obtained.

In this paper, we consider an $M^X/G/1$ queueing system with random vacation policy (RVP). When the server finishes the service of all customers in the arriving batches and the system becomes empty, the server leaves for a vacation. The server is allowed to take a maximum number, denoted by Y , successive vacations if the system remains empty after the end of a vacation. After completing Y successive vacations if the system is still empty, the server becomes dormant in the system and waits for the upcoming arrival. The maximum number of Y vacations taken by the server is a discrete random variable (r.v) with probability mass function (PMF) $P\{Y = j\} = y_j; j = 1, 2, \dots$

There have been several models evolved in recent years for the vacation queueing systems with a dormant period. The notion of a dormant period was first introduced by Heyman (1968), where he assumed that the dormant period is a period in which the server is waiting for some customers. Later, Choudhury (2002a) studied the dormant period, which arises between vacation period and busy period. Moreover, Ke, Huang, and Pearn (2010) investigated such a vacation queueing system with dormant period and its application in production and inventory systems.

The optimal control policy has received a great interest in recent years in the vacation queueing systems. The first study of vacation queueing systems with control policy was done by Kella (1989). The variations and expansions of these vacation queueing systems with control policy can be referred to Lee and Srinivasan (1989), Lee et al. (1994, 1995), Zhang, Vickson, and van Eenige (1997), Tadj and Choudhury (2005) and others. Further, Ke (2007) analyzed the operating characteristic of an $M^X/G/1$ system with multiple variant vacation policy and balking. In the same paper, it is presumed that the server takes at most Y number of vacations repeatedly until it finds some customers at a vacation completion epoch and obtained the steady state system size probabilities and steady state queue size distribution at departure point using the well known supplementary variable technique. Recently, Yang, Chang, and Ke (2016) discussed an optimal control policy of an unreliable retrial $M^X/G/1$ system with general repeated attempts and J optional multiple vacations. In this queueing system, when the unreliable server becomes empty, it leaves for the first essential vacation, after which it may either remains idle within the system or take one of J optional vacations. It is also presumed that when the server is busy, it is subject to random breakdowns and repair.

Many researchers have generalized multiple and single vacation queueing model due to its numerous applications in production management, computer communication networks, telecommunication systems, etc. Takagi (1991) and Leung (1992) considered such a model in which they introduced the notion of RVP (a generalization of multiple and single vacation), where the server takes at most Y number of random vacations repeatedly until it finds some customers at a vacation completion epoch. If no arrival occurs by the end of the vacation the server becomes idle. Choudhury (2002b) generalized this model for a batch arrival queueing system as a modification of multiple and single vacation model. Later, Ke, Wu, and Zhang (2010) have made an extensive investigation of RVP queueing systems. However, a further generalization of such a model is also possible by considering Y as a random variable instead of a fixed number (i.e., the maximum number of vacations). Yang, Hou, and Wu (2018) discussed a discrete-time $Geo/Geo/1$ queueing system with multiple working vacations (a generalization of multiple vacations). In the same paper, it is presumed that when the system becomes empty, the server enters into a working vacation i.e., during a working vacation the arriving customers are served by a lower service rate. They also investigated customers equilibrium strategies for joining or balking. Later, Jeyakumar and Senthilnathan (2017) also studied such type of multiple working vacations but with an unreliable $M^X/G/1$ queueing system. They obtained the PGF the system size distribution at a random epoch as well as other completion epoch. More recently, Ke, Chang, and Liu (2019) discussed an $M/M/C$ balking retrial queueing system with both single and multiple vacation policy. In the same paper, some useful formula and various system performance measures are also obtained.

In the literature, the steady state performance measures are studied for various exhaustive service queueing systems with a random vacation policy which focuses on the maximum number of random vacations Y , which is a fixed number. However, no work has been done in a queueing model taking into account batch arrival and random vacation policy together. Thus, in this paper, we propose to investigate an $M^X/G/1$ queueing system with random vacation policy.

This type of model has a potential application in the IEEE 802.16e standard, which supports the mobility of a mobile station (MS) by involving **Sleep/Awake** period (mode) operational sequence to save the battery power of MS. Here, sleep period is modeled as random vacation period (contains maximum Y number of random vacations), warm-up period is modeled as dormant period and awake period is modeled as busy period. The detailed justification of the proposed model with application is illustrated in section 2 and in numerical section 8, we have shown some examples of IEEE802.16e standard. This motivates us to study such a model in the form of random vacation policy in this paper. Hence, in this paper, we proposed to investigate such a type of $M^X/G/1$ queueing system with random vacation policy.

The rest of the paper is arranged as follows: The practical implementation of the queueing model under consideration is presented in Section 2. In Section 3, the mathematical model is described. In Section 4, the queue size distribution at service completion epoch and steady state system size probabilities using embedded Markov chain (EMC) and Markov regenerative process (MRGP) are derived. Some other essential system performance measures of this model are obtained in Section 5. Some important

particular case of this model is demonstrated in Section 6. Section 7 display the cost optimal policy of the proposed model along with average cost function. Finally, in Section 8, some important numerical results are presented followed by concluding remarks in Section 9.

2. Practical implementation of the proposed model

In an IEEE802.16e system, power handling is a significant component of mobility support of a mobile station (MS) by involving **Sleep/Awake** period (mode) operational sequence, see Seo et al. (2004). Sleep period is a state in which MS chooses a period of absence from the serving base station (BS). Before starting the sleep period, an MS in the awake period sends a signal to a being served BS and waits for BS's approval. After receiving the positive approval (i.e., there is no data to be sent to MS), it begins a sleep period. The sleep period consist of many sleep intervals. The sleep intervals are denoted by Sl which is considered to be a generic random variable. After the end of every sleep interval, it wakes up and inspects whether there are any frames or packets for it or not. This task is carried out in the ON interval and is denoted by T_{op} . Here, we are assuming T_{op} to be a constant. Any arrival during sleep interval can only be notified in the following ON interval. Here, the initial sleep interval starts with an ON interval. If no arrival occurs at the end of a sleep interval, it will continue for an upper bound on a number of sleep intervals, whereas if any packet arrived to MS during the ON interval (i.e., after the sleep interval), it starts awake period, otherwise it goes for the last ON interval. The last ON interval is included in the warm-up period and is denoted by T_{w1} (in practice, $T_{op} = T_{w1}$) (see Seo et al. (2004) and Alouf, Altman, and Azad (2008)). Warm-up period is a period where the MS waits for the arrivals after completing a maximum number of sleep.

Now, we define two regeneration cycle namely, Type 1 cycle and Type 2 cycle. The regeneration cycle is defined as the elapsed time from the epoch in which a random vacation period is enabled, until the epoch where the next random vacation period is about to begin. Figures 1 and 2 show the period (mode) of an $M^X/G/1$ system with random vacation policy to the feasible operational mode of an MS during type 1 cycle and type 2 cycle respectively.

Type 1 cycle is defined as the elapsed time from the epoch in which a random vacation period is enabled, until the epoch where the next random vacation period is about to begin i.e., till busy period ends. Therefore in type 1 cycle, if any arrival occurs during the random vacation period (i.e., sleep period), the server starts service immediately at that random vacation completion epoch i.e., busy period starts and as the busy period ends, again random vacation period starts.

Similarly, type 2 cycle is defined as the elapsed time from the epoch in which a random vacation period is enabled, until the epoch where the next random vacation period is about to begin i.e., till busy period ends. After random vacation period, dormant period (warm-up period) starts and after the dormant period, busy period starts. After the end of busy period, again random vacation period is about to begin. Therefore in type 2 cycle, if no arrival occurs during a random vacation period, the server stays dormant (warm-up) in the system and if any batch arrival occurs during this dormant

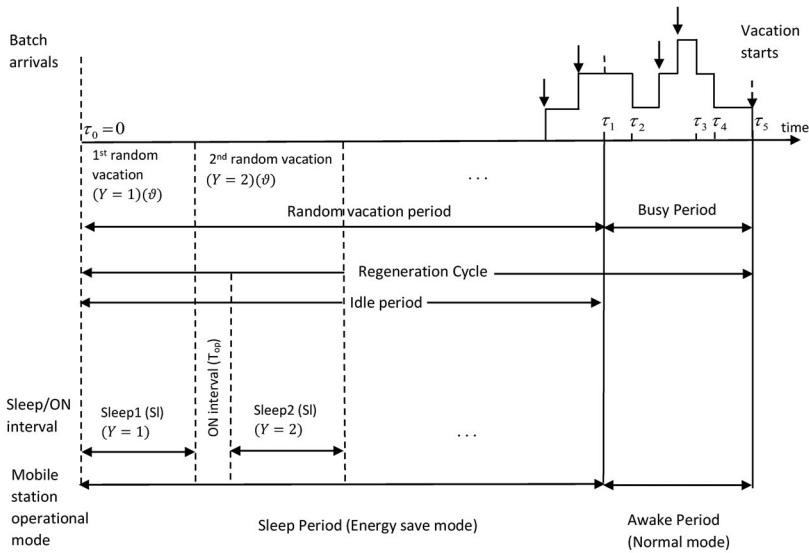


Figure 1. Mapping the $M^X/G/1$ system with random vacation policy to the mobile station operational mode during type 1 cycle.

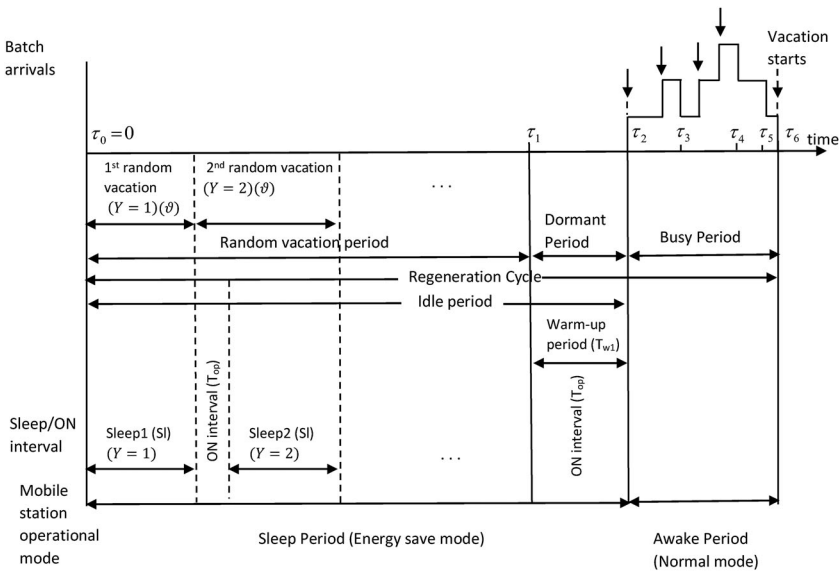


Figure 2. Mapping the $M^X/G/1$ system with random vacation policy to the mobile station operational mode during type 2 cycle.

period, it immediately starts service at that dormant period completion epoch i.e., busy period starts. After the end of busy period the next random vacation period starts.

Seo et al. (2004) proposed an $M/G/1$ queueing system with multiple vacations to evaluate the dropping probability, average waiting time and average delay of the packets in a BS's buffer. Park and Hwang (2007) suggested an $M/G/1/k$ finite buffer queueing system with multiple vacations to analyze the sleep-period operation of a standard IEEE

802.16e system. Moreover, Azad et al. (2009, 2011) formulate an optimization problem for an $M^X/G/1$ queueing system with different vacation policies to analyze the sleep period in an IEEE 802.16e system. Here, we propose to study an $M^X/G/1$ queueing system with a random vacation policy, where each ON and sleep interval (except the first sleep) is modeled as random vacation period and warm-up period is modeled as dormant period, i.e.,

Random vacation period (1^{st} random vacation) = Sleep interval (Sleep 1)(SI)
 Random vacation period (Contains a maximum Y number of random vacations) =
 ON interval (T_{op}) + Sleep interval (maximum Y number of sleep intervals)
 (SI), $Y = 2, 3, \dots$
 Dormant period = Warm-up period (T_{w1}).

The batch arrival process conforms to compound Poisson process and the delivering period of packets is generally (non exponentially) distributed. The Poisson process with batch arrival can be a practical abstraction of video traffic. Any new batch arrival during the random vacation period brings back the MS into the normal service period. These characteristics of the random vacation period in a random vacation policy model can reduce holding time of an arriving packet, which appears at a future stage of the sleep period.

3. Model description

In this section, we consider an $M^X/G/1$ queue with random vacation policy, where the server takes a maximum number of Y random vacations until the server finds a batch of minimum one packet or customer in the queue. The detailed description of the model, based on the information given in section 2, is stated below:

Customers arrive in batches with rate $\lambda > 0$ conforming to a compound Poisson process. Let X be the batch size r.v and $X_1, X_2, \dots$, are corresponding batch sizes of arriving customers which are independently and identically distributed (i.i.d.) random variables (r.vs), with probability mass function (PMF)

$$P\{X_i = k\} = b_k, k = 1, 2, \dots$$

probability generating function (PGF), $P_X(z) = \sum_{k=1}^{\infty} z^k b_k$ and finite moments $E(X^r); r = 1, 2, \dots$. Here, the first moment $E(X)$ is the mean batch size. Let us consider $b_n^{(k)} = P[X_1 + X_2 + \dots + X_k = n]$ is the k -fold convolution of $\{b_n\}$ with itself and $b_n^{(0)}$ is defined to be 1.

The service time provided by a single server is an i.i.d. r.v S with cumulative distribution function (CDF) $F_S(x)$, probability density function (PDF) $f_S(x)$, Laplace Stieltjes transform (LST) $S^*(\theta) = E(e^{-\theta S})$ and finite moments $E(S^r); r = 1, 2, \dots$. Here, the first moment $E(S)$ is the average length of service time S . Once the server finishes the service of all customers in the arriving batches, i.e., at the end of each service period, the server enters into a vacation period of random duration ϑ (say). The duration of vacation period ϑ is i.i.d. r.v with CDF $F_{\vartheta}(x)$, the corresponding PDF $f_{\vartheta}(x)$, LST $\vartheta^*(\theta) = E(e^{-\theta \vartheta})$ and finite moments $E(\vartheta^r); r = 1, 2, \dots$. On return from the initial vacation if the server finds no batch waiting in the queue it goes for the second vacation. This procedure will be continued until either

the server finds a batch of minimum one customer at a vacation completion epoch or it has taken a maximum Y number of random vacations (random vacation period). In case of arrivals occurred during a random vacation, the server resumes service immediately at that random vacation completion epoch, which is shown in Figure 1. In case of no arrival occurred during a random vacation period, the server stays dormant in the system and waits to serve the first batch arrival, which is shown in Figure 2. Y is presumed to be a discrete r.v instead of a fixed integer, such that

$$P\{Y = j\} = y_j; j = 1, 2, \dots$$

and $P_Y(z) = E(z^Y)$, its corresponding PGF. Moreover, the system remains idle during a random vacation period and a dormant period, i.e., in case of Figure 1, random vacation period is an idle period and in case of Figure 2, random vacation period and a dormant period together constitute an idle period.

Now, if we define ω_k and ϱ_k as the probabilities that k individual batches arrive during the service time S and random vacation time ϑ respectively, then for $k = 1, 2, \dots$, we have $\omega_k = \int_0^\infty e^{-\lambda x} \left[\frac{(\lambda x)^k}{k!} \right] dF_S(x)$ and $\varrho_k = \int_0^\infty e^{-\lambda x} \left[\frac{(\lambda x)^k}{k!} \right] dF_\vartheta(x)$. Here, $\varrho_0 = \vartheta^*(\lambda)$ is the probability that no individual batch arrives during a random vacation time ϑ .

If we consider two events namely,

γ = No batch arrives during the random vacation period and

φ = At least one batch arrives during the random vacation period, then, the probability of event γ and φ will be

$$P(\gamma) = \sum_{n=1}^\infty (\vartheta^*(\lambda))^n y_n = P_Y(\vartheta^*(\lambda)) \text{ and } P(\varphi) = 1 - P_Y(\vartheta^*(\lambda)).$$

Further, it is presumed that batch arrival process, service time r.v, vacation time r.v and discrete r.v Y are all mutually independent. It is essential to note that here we consider the concept of a warm-up period in an IEEE802.16e system, which refers to the time that is necessary for the server to wait for the first batch arrival after completing a random vacation period and this warm-up period is modeled as a dormant period in the underlying queueing system.

4. Model analysis

4.1. Queue size distribution at service completion epoch

In this subsection, we derive the PGF of queue size distribution at service completion epoch. Let t_m denote the service completion epoch of m^{th} batch. Then the sequence $\Xi_m = Z(t_m^+)$ (where $Z(t_m)$ represents the total number of customers in the arriving batches to the system at time epoch t_m) construct a discrete time Markov chain (DTMC), which is an embedded Markov chain in the underlying queueing system. The sequence $\{\Xi_m, m \geq 0\}$ is a homogeneous Markov chain and it is because of the transition

$$\Xi_{m+1} = \begin{cases} \Xi_m - 1 + L_{m+1} & \text{while } \Xi_m > 0 \\ A_{m+1} + L_{m+1} - 1 & \text{while } \Xi_m = 0 \end{cases}$$

where L_m is the number of batches arrived during the service time of m^{th} batch and A_m is the number of batches arrived either during the m^{th} random vacation period or dormant period.

Clearly, the transition probability matrix $K(\infty) = (P_{hn})$ readily seen to be a so called Δ_2 matrix, which is a special case of $\Delta_{m,n}$ matrix (see Abolnikov and Dukhovny (1991)), which is of the form

$$K(\infty) = (P_{hn}) = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ 0 & P_{21} & P_{22} & \cdots \\ 0 & 0 & P_{31} & \cdots \end{bmatrix}$$

where $P_{hn} = P\{\Xi_{m+1} = n/\Xi_m = h\}$ is the transition probability associated with the Markov Chain $\{\Xi_m, m = 0, 1, \dots\}$.

It is clear that $\{\Xi_m, m = 0, 1, \dots\}$ is an irreducible and aperiodic discrete time Markov chain. Therefore, the one-step transition probabilities of Ξ_m are given by

$$P_{h,n} = \begin{cases} \sum_{c=1}^{n+1} \left[\{1 - P_Y(\vartheta^*(\lambda))\} \frac{\sum_{k=1}^c b_c^{\odot(k)} q_k}{(1 - q_0)} + P_Y(\vartheta^*(\lambda)) b_c \right] l_{n-c+1} & \text{if } h = 0, n = 0, 1, \dots, \\ l_{n-h+1} & \text{if } n = h - 1, h, h + 1, \dots, h \geq 1, \\ 0 & \text{if } h \geq 1, n = 0, 1, \dots, h - 2 \end{cases} \quad (1)$$

where $l_n = \sum_{k=0}^n b_n^{\odot(k)} \omega_k = \int_0^\infty \sum_{k=0}^n e^{-\lambda x} \frac{(\lambda x)^k}{k!} b_n^{\odot(k)} dF_S(x)$; $n = 0, 1, 2, \dots$ is the probability that several batches with total of n customers arrive during a service time S .

Now, it is presumed that the Markov chain $\{\Xi_m, m = 0, 1, \dots\}$ is positive recurrent to ensure that the following limiting probabilities

$$\pi_n = \lim_{m \rightarrow \infty} P[\Xi_m = n], n = 0, 1, \dots$$

exist. Using Equation (1), we can write the steady state equation as

$$\pi_n = \pi_0 \sum_{c=1}^{n+1} \left[\{1 - P_Y(\vartheta^*(\lambda))\} \frac{\sum_{k=1}^c b_c^{\odot(k)} q_k}{(1 - q_0)} + P_Y(\vartheta^*(\lambda)) b_c \right] l_{n-c+1} + \sum_{c=1}^{n+1} \pi_c l_{n-c+1}, n = 0, 1, \dots \quad (2)$$

Now, the probability generating functions for $|z| \leq 1$ is defined as

$$P_L(z) = \sum_{n=0}^\infty z^n l_n, \text{ and } P_\Pi(z) = \sum_{n=0}^\infty z^n \pi_n.$$

Clearly, $P_L(z)$ can be written as

$$\begin{aligned} P_L(z) &= \sum_{n=0}^\infty z^n \sum_{k=0}^n b_n^{\odot(k)} \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^k}{k!} dF_S(x) \\ &= \sum_{k=1}^\infty \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^k}{k!} \sum_{n=1}^\infty z^n b_n^{\odot(k)} dF_S(x) + \int_0^\infty e^{-\lambda x} dF_S(x) \\ &= \sum_{k=1}^\infty \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^k}{k!} (P_X(z))^k dF_S(x) + \int_0^\infty e^{-\lambda x} dF_S(x), \\ &= \int_0^\infty e^{-\lambda x} dF_S(x) [e^{\lambda P_X(z)x} - 1] + \int_0^\infty e^{-\lambda x} dF_S(x) \\ &= \int_0^\infty e^{-[\lambda x - \lambda P_X(z)x]} dF_S(x) \\ &= S^*(\lambda(1 - P_X(z))) \end{aligned}$$

where $(P_X(z))^k = \sum_{n=1}^\infty z^n b_n^{\odot(k)}$.

Also, $\frac{d}{dz}P_L(z)|_{z=1} = \lambda E(X)E(S)$ is the average number of batch arrivals during the service time.

Now, multiplying Equation (2) by z^n and then taking summation over $n = 0, 1, \dots$, we get

$$P_{\Pi}(z) = \frac{\pi_0 \left[1 - \left\{ (1 - P_Y(\vartheta^*(\lambda))) \left\{ \frac{(P_V(z) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))P_X(z)}{1 - P_Y(\vartheta^*(\lambda))} \right\} \right\} \right] P_L(z)}{(P_L(z) - z)} \quad (3)$$

where $P_V(z) = \sum_{n=0}^{\infty} z^n \sum_{k=0}^n b_n^{(k)} \varrho_k = \vartheta^*(\lambda(1 - P_X(z)))$ and $\vartheta^*(\lambda)$ is defined on page 5.

Clearly, $\frac{d}{dz}P_V(z)|_{z=1} = \lambda E(X)E(\vartheta)$ is the average number of batch arrivals during the random vacation time.

Now, applying the normalizing condition $P_{\Pi}(1) = 1$, we obtain

$$\pi_0 = \frac{(1 - \lambda E(X)E(S))(1 - \vartheta^*(\lambda))(1 - P_Y(\vartheta^*(\lambda)))^{-1}}{E(X) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (4)$$

Utilizing Equation (4) in (3), we obtain

$$P_{\Pi}(z) = \frac{(1 - \lambda E(X)E(S))(1 - \vartheta^*(\lambda)) \left[\frac{1}{(1 - P_Y(\vartheta^*(\lambda)))} - \left\{ \frac{(P_V(z) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))P_X(z)}{1 - P_Y(\vartheta^*(\lambda))} \right\} \right] P_L(z)}{E(X) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] (P_L(z) - z)} \quad (5)$$

which is the PGF of queue size distribution at service completion epoch.

Let $E(\Pi)$ be the mean queue size at service completion epoch, then by differentiating Equation (5) both sides with respect to z and then taking limit $z \rightarrow 1$, we obtain

$$E(\Pi) = \frac{2\lambda E(X)E(S) - (\lambda E(X))^2(2(E(S))^2 - E(S^2)) + \lambda E(X^2)E(S)}{2(1 - \lambda E(X)E(S))} \\ + \frac{(\lambda E(X))^2 E(\vartheta^2) + P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))E(X^2) + E(X^2)\lambda E(\vartheta)}{2E(X) \left[\lambda E(X) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (6)$$

Remark 1. It can be observed from Equation (4) that $\lambda E(X)E(S) < 1$, which is a necessary and sufficient condition for the stable solution of the model. It is basically the stability condition of the model. Since $\lambda E(X)E(S) < 1$, therefore the equation $P_L(z) - z = 0$ never vanishes inside the region $|z| \leq 1$, which can be proved by using Rouché's Theorem (see Takacs (1962)).

4.2. Steady state system size probabilities

To obtain an explicit expression for steady state system size probabilities, we follow the method of Markov regenerative process, which gives much quicker result than the more popular method of supplementary variables techniques.

Let $h \in \{0, 1, \dots\}$ denote the total number of customers in the arriving batches to the system (including the one being served if any) and $s \in \{0, 1, 2\}$ indicate whether the system is in random vacation period (which consists of maximum Y number of random vacations) or dormant period or busy period at time t . Thus, the state of the system can be described

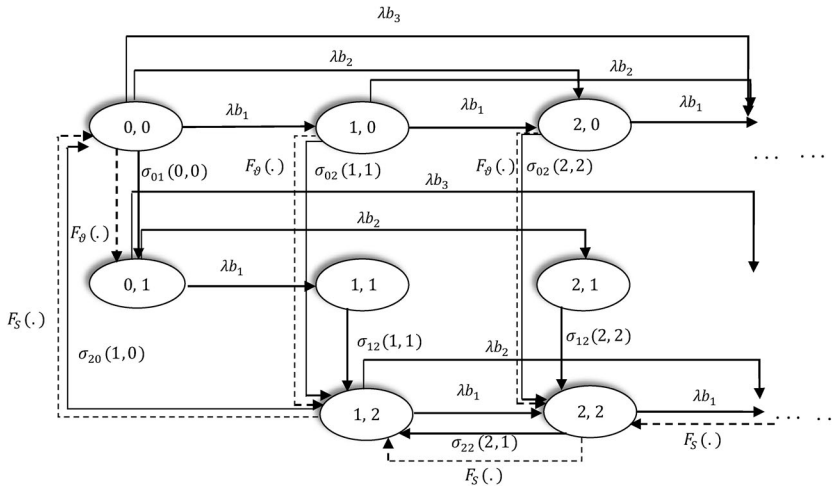


Figure 3. State transition diagram for $M^X/G/1$ system with random vacation policy.

by a 2-tuple or ordered pair $(h, s) \in \{0, 1, \dots\} \times \{0, 1, 2\} \equiv \Omega$. Consider the system state process $\{U(t), t \geq 0\}$, where $U(t)$ denote the total number of customers in the arriving batches to the system at time t . The process $\{U(t), t \geq 0\}$ is not a continuous time Markov chain (CTMC) since the sojourn time in each state is not exponentially distributed. Moreover, $U(t)$ is not even a semi-Markov process since between two completion epochs (i.e., between two sequential regeneration epochs) system state can change due to batch arrivals in random vacation period or dormant period or service completion. Define $\{\tau_n, n = 0, 1, \dots\}$ as the sequence of epochs in which either service is terminated or random vacation is terminated or dormant period is terminated and let $\zeta_n \in \Omega$ be the state of the system at these epochs. Here, the elements of Ω are called regeneration epochs and they occur at service completion epoch, random vacation completion epoch, dormant period completion epoch. Hence, we can claim that $\{\zeta_n, n = 0, 1, \dots\}$ is a DTMC with one-step transition probability matrix $K(\infty)$. By observing stochastic process $\{U(t), t \geq 0\}$ at these epochs, we see that the sequence $\{\zeta_n, \tau_n\}$ is an embedded Markov renewal sequence and $\{U(t), t \geq 0\}$ is a MRGP (see Kulkarni (1995)).

We now proceed to establish the local kernel $E(t)$ and global kernel $K(t)$ of the MRGP. To simplify the analysis, we first draw the state transition diagram and it is shown in Figure 3. In Figure 3, we have not considered the state $0, 2$ because when the server is in a busy state, there is at least one batch for service and if no batch is waiting in a queue, the system moves to random vacation period. In the state transition diagram, we apply an oval symbol with ordered pair (h, s) (as stated above) to show the state and apply a directed line to show a state transition. There are two types of state transitions as follows.

- Transition with a generally distributed sojourn times (corresponding to the service time and random vacation time) and is denoted by $F_S(\cdot)$ and $F_\theta(\cdot)$ with dotted line respectively.
- Transition with exponentially distributed sojourn time (corresponding to inter-arrival time) with transition rate $\sigma_{ij}(h, n), 0 \leq i, j \leq 2, h = 0, 1, \dots, n = 0, 1, \dots$ and is represented by the solid line. The possible transitions and corresponding rates are given in the following Table 1

Table 1. Table for transition and corresponding rates.

From	To	Transition rate
(0, 0)	(0, 1)	σ_{01} (0, 0)
(h, 0)	(n, 2)	σ_{02} (h, n) for h = n = 1, 2, ...
(1, 1)	(1, 2)	σ_{12} (1, 1)
(1, 2)	(0, 0)	σ_{20} (1, 0)
(h, 2)	(n, 2)	σ_{22} (h, n) for h = 2, 3, ..., n = h - 1.

Now, to obtain the system state probabilities of MRGP, we find the closed form expression for $K(t)$ and $E(t)$ as follows.

4.2.1. Closed form expression for $K(t)$

$K(t)$ is a global kernel matrix. The interpretation for the elements of the global kernel matrix $K(t)$ is as follows: $K_{(h,s);(n,s')}(t)$ denotes the probability that the system will be in state (n, s') , $s' \in \{0, 1, 2\}$, $n = 0, 1, \dots$ at the time of the next regeneration epoch which occur on or before time t , given that it was in state (h, s) just after the previous regeneration epoch. The evolution of MRGP between two sequential Markov regeneration epochs can be explained through a CTMC infinitesimal generator matrix which involves an exponential transition. Hence, by removing general transition $F_S(\cdot)$ and $F_\theta(\cdot)$ in Figure 3 we can construct the CTMC infinitesimal generator matrix with the ordered pair $\{(0,0), (0,1), (1,0), (1,1), (1,2), (2,0), (2,1), (2,2), \dots\}$ as follows:

$$Q = \begin{matrix} & \begin{matrix} (0,0) & (0,1) & (1,0) & (1,1) & (1,2) & (2,0) & (2,1) & (2,2) & \dots \end{matrix} \\ \begin{matrix} (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \\ (1,2) \\ (2,0) \\ (2,1) \\ (2,2) \\ \vdots \end{matrix} & \left[\begin{array}{ccccccccc|cccc} q_0^0 & \sigma_{01}(0,0) & \lambda b_1 & 0 & 0 & \lambda b_2 & 0 & 0 & \dots \\ 0 & q_1^0 & 0 & \lambda b_1 & 0 & 0 & \lambda b_2 & 0 & \dots \\ \hline 0 & 0 & q_0^1 & 0 & \sigma_{02}(1,1) & \lambda b_1 & 0 & 0 & \dots \\ 0 & 0 & 0 & q_1^1 & \sigma_{12}(1,1) & 0 & 0 & 0 & \dots \\ \hline \sigma_{20}(1,0) & 0 & 0 & 0 & q_2^1 & 0 & 0 & \lambda b_1 & \dots \\ \hline 0 & 0 & 0 & 0 & 0 & q_0^2 & 0 & \sigma_{02}(2,2) & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & q_1^2 & \sigma_{12}(2,2) & \dots \\ \hline \sigma_{20}(2,0) & 0 & 0 & 0 & \sigma_{22}(2,1) & 0 & 0 & q_2^2 & \dots \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \right] \end{matrix}$$

where $q_0^0 = -(\sigma_{01}(0,0) + \lambda \sum_{k=1}^{\infty} b_k)$, $q_1^0 = -\lambda \sum_{k=1}^{\infty} b_k$, $q_0^1 = -(\sigma_{02}(1,1) + \lambda \sum_{k=1}^{\infty} b_k)$, $q_1^1 = -\sigma_{12}(1,1)$, $q_2^1 = -(\sigma_{20}(1,0) + \lambda \sum_{k=1}^{\infty} b_k)$, $q_0^2 = -(\sigma_{02}(2,2) + \lambda \sum_{k=1}^{\infty} b_k)$, $q_1^2 = -\sigma_{12}(2,2)$, $q_2^2 = -(\sigma_{20}(2,0) + \sigma_{22}(2,1) + \lambda \sum_{k=1}^{\infty} b_k)$ and λb_k is the batch arrival rate.

Let $N(t) \in \{0, 1, \dots\}$ be the total number of customers in the arriving batches to the system during time interval $(0, t]$ and $D(t) \in \{0, 1, 2\}$ be the state of the system. Then we define

$$P_{ij}(h; n, t) = P\{N(t) = n, D(t) = j | N(0) = h, D(0) = i\}; n = h = 0, 1, \dots \quad (7)$$

$P_{ij}(h; n, t)$ is the probability that the subordinate CTMC will be in state (n, j) during time t given that it was initially in state (h, i) . Let $\mathbf{P}(t)$ be the corresponding probability matrix.

Then

$$\frac{d}{dt} \mathbf{P}(t) = \mathbf{P}(t) Q, P_{ii}(h; h, 0) = 1, i \in \{0, 1, 2\} \quad (8)$$

where $P_{ii}(h; h, 0)$ is the probability that the subordinate CTMC will be in the same state with h customers initially.

By solving Equation (8), we obtain $P_{ij}(h; n, t)$ and from the properties of transition probabilities, we get $\sigma_{ij}(h, n) = \lim_{t \rightarrow 0} \frac{P_{ij}(h; n, t)}{t} < \infty, \forall i \neq j \neq h \neq n$.

Now, according to the definition of global kernel $K(t)$ in Kulkarni (1995), we want to find

$$K_{(h,s);(n,s')}(t) = P\{\zeta_1 = (n, s'), \tau_1 \leq t | \zeta_0 = (h, s), 0 \leq s \leq 2, s' \in \{0, 1, 2\}\} \quad (9)$$

where τ_1 is the interval between two sequential regeneration epochs.

Now, we define the elements of global kernel for each transition as follows: (See the Appendix)

In summary, the global kernel $K_{(h,s);(n,s')}(t)$ of MRGP (see Equation (9)) can be given as

$$K_{(h,s);(n,s')}(t) = \begin{cases} \int_0^t P_{01}(0; 0, x) dF_\vartheta(x), \\ \int_0^t P_{02}(h; n, x) dF_\vartheta(x), n = h = 1, 2, \dots, \\ P_{12}(h; n, t), n = h = 1, 2, \dots, \\ \int_0^t P_{20}(1; 0, x) dF_S(x), \\ \int_0^t P_{22}(h; n, x) dF_S(x), n = 1, 2, \dots, h = 1, 2, \dots, \\ 0, \text{ otherwise} \end{cases}$$

4.2.2. Closed form expression for $E(t)$

The local kernel matrix $E(t)$ describes the local behavior of the process between two sequential regeneration epochs. The description for the elements of local kernel matrix $E(t)$ is as follows: $E_{(h,s);(n,s')}(t)$ denotes the probability that the system will be in state (n, s') at time t and the next regeneration epoch occurs after time t given that the system was in state (h, s) . Hence, we can write the elements of $E(t)$ as

$$E_{(h,s);(n,s')}(t) = P\{U(t) = (n, s'), \tau_1 \geq t | \zeta_0 = (h, s)\} \quad (10)$$

where τ_1 is the time period between two sequential regeneration epochs.

Following the similar procedure in obtaining the global kernel $K(t)$, we find $E(t)$ as

$$E_{(h,s);(n,s')}(t) = \begin{cases} P_{02}(h; n, t)(1 - F_\vartheta(t)), n = h = 1, 2, \dots, \\ P_{12}(h; n, t), n = h = 1, 2, \dots, \\ P_{22}(h; n, t)(1 - F_S(t)), n = 1, 2, \dots, h = 1, 2, \dots, \\ 0, \text{ otherwise} \end{cases}$$

4.2.3. System state probabilities

To obtain the system state probabilities, we define a new variable $\alpha_{ss'}(h; n)$ which represents the average amount of time the MRGP spends in state s' with n total customers in

the arriving batches between two sequential regeneration epochs, given that it was in state s with h total customers in the arriving batches after last regeneration:

$$\alpha_{ss'}(h; n) = \int_0^\infty E_{(h,s);(n,s')}(t) dt$$

It is easy to show that for $s=0$, $s'=2$ and $n=h=1, 2, \dots$

$$\alpha_{02}(h; n) = E(X) \sum_{k=0}^n b_n^{\odot(k)} [1 - P_Y(\vartheta^*(\lambda))] \{1 - \vartheta^*(\lambda)\}^{-1} \int_0^\infty \frac{e^{-\lambda t} (\lambda t)^k}{k!} (1 - F_\vartheta(t)) dt - \frac{E(X) [P_Y(\vartheta^*(\lambda)) - b_n P_Y(\vartheta^*(\lambda))]}{\lambda} \quad (11)$$

For $s=1$, $s'=2$ and $h=n=1, 2, \dots$

$$\alpha_{12}(h; n) = \frac{b_n}{\lambda} \quad (12)$$

and for $s=2$, $s'=2$, $n=1, 2, \dots$ and $h=1, 2, \dots$, we have

$$\alpha_{22}(h; n) = E(X) g_{n-h+1}; n \geq \max(0, h-1) \quad (13)$$

where $g_n = \sum_{k=0}^n \int_0^\infty e^{-\lambda t} \left[\frac{(\lambda t)^k}{k!} \right] b_n^{\odot(k)} (1 - F_S(t)) dt$; $n=0, 1, \dots$

Since the system state is basically based on the value of h while $s \in \{0, 1, 2\}$, therefore the system state probabilities are defined as

$$\phi(n, s') = \frac{\sum_{h=0}^\infty \pi_h \alpha_{ss'}(h; n)}{\sum_{h=0}^\infty \pi_h \beta_h} \quad (14)$$

where π_h 's are the solution of $\pi = \pi K(\infty)$ and $\sum_{h=0}^\infty \pi_h = 1$. Here $\beta_h = \sum_{v \in \{0, 1, 2\}} \alpha_{vs'}(h; n)$; $n=0, 1, \dots$ and $\sum_{h=0}^\infty \pi_h \beta_h = \frac{1}{\lambda}$.

Since, the arrival process in batches is Poisson, therefore, by obeying Burke's theorem (see Cooper (1981), page 187–188) it is achieved that the system state probabilities $\{\phi(n, s'); (n, s') \in \Omega\}$ are positive and exist under the similar condition of limiting probabilities $\{\pi_n; n=0, 1, \dots\}$ of the Markov chain $\{\Xi_m, m=0, 1, \dots\}$, i.e., if and only if $\lambda E(X) E(S) < 1$.

Now, we consider the case of stationary system size distribution when the server state is in random vacation, i.e., $s'=0$, a probability argument with Equation (14) gives rise to

$$\phi(n, 0) = \lambda E(X) \frac{\pi_0}{\{1 - \vartheta^*(\lambda)\}} [1 - P_Y(\vartheta^*(\lambda))] \sum_{k=0}^n b_n^{\odot(k)} \int_0^\infty \frac{e^{-\lambda t} (\lambda t)^k}{k!} (1 - F_\vartheta(t)) dt - E(X) [\pi_0 (P_Y(\vartheta^*(\lambda)) - b_n P_Y(\vartheta^*(\lambda)))] ; n=0, 1, \dots \quad (15)$$

Next, we consider the case of stationary system size distribution when the server state is in dormant period, i.e., $s'=1$, a probability argument with Equation (14) gives rise to

$$\phi(n, 1) = \pi_0 b_n P_Y(\vartheta^*(\lambda)), n=0, 1, \dots \quad (16)$$

At the end we consider the case of stationary system size distribution when the server state is in busy period, i.e., $s'=2$, a probability argument with Equation (14) gives rise to

$$\phi(n, 2) = \lambda E(X) \left[\pi_0 \sum_{c=1}^{n+1} \left[\left\{ 1 - P_Y(\vartheta^*(\lambda)) \right\} \frac{\sum_{k=1}^c b_c^{(k)} q_k}{(1 - q_0)} + P_Y(\vartheta^*(\lambda)) b_c \right] g_{n-c+1} + \sum_{c=1}^{n+1} \pi_c g_{n-c+1} \right] \quad (17)$$

Now, we find the PGF of the above Equations (15)–(17), which are given in the following theorem.

Theorem 1. Under stability condition $\lambda E(X)E(S) < 1$, the PGF $P_0(z)$ of system size distribution when the server state is in random vacation period (i.e., $s' = 0$) is found as

$$P_0(z) = \frac{(1 - \lambda E(X)E(S))(1 - P_Y(\vartheta^*(\lambda))) \left[(1 - P_V(z)) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{(1 - P_Y(\vartheta^*(\lambda)))} (1 + (P_X(z))^2 - 2P_X(z)) \right]}{(1 - P_X(z)) \left[\lambda E(\vartheta) + \frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (18)$$

The PGF $P_1(z)$ of system size distribution when the server state is in dormant period (i.e., $s' = 1$) is found as

$$P_1(z) = \frac{(1 - \lambda E(X)E(S))(1 - \vartheta^*(\lambda))P_X(z)P_Y(\vartheta^*(\lambda))}{E(X)(1 - P_Y(\vartheta^*(\lambda))) \left[\lambda E(\vartheta) + \frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (19)$$

and the PGF $P_2(z)$ of system size distribution when the server is busy (i.e., $s' = 2$) is found as

$$P_2(z) = \frac{(1 - \lambda E(X)E(S))(1 - \vartheta^*(\lambda)) \left[\frac{1}{(1 - P_Y(\vartheta^*(\lambda)))} - \left\{ \frac{(P_V(z) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))P_X(z)}{1 - P_Y(\vartheta^*(\lambda))} \right\} \right] (1 - P_L(z))}{(1 - P_Y(\vartheta^*(\lambda))) [P_L(z) + zP_X(z) - P_X(z)P_L(z) - z] \left[\lambda E(\vartheta) + \frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (20)$$

Proof. Multiplying Equations (15)–(17) by z^n in both sides and summing over $n = 0, 1, \dots$ and then utilizing the established result from renewal theory

$$\int_0^\infty e^{-\theta x} (1 - F(x)) dx = \frac{1 - f^*(\theta)}{\theta}$$

where $F(\cdot)$ represents the CDF of an r.v and $f^*(\cdot)$ its LST, we obtain the results (18)–(20) respectively.

Next, from the PGF of Theorem 1, we find the system state probabilities, where the server can be in random vacation period(RV), in a dormant period (D), in a busy period (S) or in an idle period (I), which are summarized in Corollary 1.

Corollary 1. For $\lambda E(X)E(S) < 1$, the probability that the server is in random vacation period during a random vacation P_{RV} , the probability that the server is in dormant period P_D , the probability that the server is in busy period P_S and in idle period P_I are obtained as

$$P_{RV} = \frac{(1 - \lambda E(X)E(S))\lambda E(\vartheta)}{\left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (21)$$

$$P_D = \frac{(1 - \lambda E(X)E(S))(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{(1 - P_Y(\vartheta^*(\lambda))) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (22)$$

$$P_S = \lambda E(X)E(S) \quad (23)$$

$$P_I = 1 - \lambda E(X)E(S) \quad (24)$$

Proof. From Equations (18)–(20) by taking limit as $z \rightarrow 1$ and after calculation, we obtain the results (21)–(23), i.e., $\lim_{z \rightarrow 1} P_0(z) = P_{RV}$, $\lim_{z \rightarrow 1} P_1(z) = P_D$ and $\lim_{z \rightarrow 1} P_2(z) = P_S$. By adding (21) and (22), Equation (24) is obtained.

Lastly, we find the PGF, mean and second moment of system size, which are given in Corollary 2.

Corollary 2. Let N denote the system size in steady state. Then, under the stability condition $\lambda E(X)E(S) < 1$, the corresponding PGF is given by

$$P_N(z) = \frac{(1 - \lambda E(X)E(S))(1 - z)P_L(z)}{\{P_L(z) - z\}} \times \frac{\frac{(1 - \vartheta^*(\lambda))}{(1 - P_Y(\vartheta^*(\lambda)))} \left[1 - \left\{ (1 - P_Y(\vartheta^*(\lambda))) \left\{ \frac{(P_V(z) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))P_X(z)}{1 - P_Y(\vartheta^*(\lambda))} \right\} \right\} \right]}{\left[\lambda E(\vartheta) + \frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] (1 - P_X(z))} \quad (25)$$

The mean of system size, $E(N)$ is given by

$$E(N) = \lambda E(X)E(S) + \frac{(\lambda E(X))^2 E(S^2) + \lambda E(X^2)E(S)}{2(1 - \lambda E(X)E(S))} + \frac{\frac{2P_Y(\vartheta^*(\lambda))}{(1 - P_Y(\vartheta^*(\lambda)))} (1 - \vartheta^*(\lambda))E(X^2) + \lambda (\lambda E(X))^2 E(\vartheta^2) + 2E(X^2)E(\vartheta)}{2E(X) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (26)$$

and the second moment of system size, $E(N^2)$ is given by

$$E(N^2) = \lambda E(X)E(S) + \frac{E(X)(E(X) + 3E(X^3)) - 3(E(X^2))^2}{2(E(X))^2} + \frac{(\lambda E(X))^2 (\lambda E(X)E(S^3) + E(S^2))}{2(1 - \lambda E(X)E(S))} + \frac{\lambda E(S)(E(X^3) + E(X^2))}{2(1 - \lambda E(X)E(S))} + \frac{(1 - \lambda E(X)E(S))\lambda^2 E(X)E(X^2)E(S) - ((\lambda E(X))^2 E(S^2) + \lambda E(X^2)E(S))^2}{2(1 - \lambda E(X)E(S))^2} + \frac{\lambda^2 (E(X))^3 (\lambda E(X)E(\vartheta^3) + E(\vartheta^2))}{2(E(X))^2 \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (27)$$

Proof. From Equations (15)–(17), we get the probability of total system size in steady state as follows

$$\xi_n = \phi(n, 0) + E(X)\phi(n, 1) + (1 - \delta_{n,0})\phi(n, 2); \text{ for } n = 0, 1, \dots \quad (28)$$

where $\xi_n = P\{N = n\}$ and

$$\delta_{m,n} = \begin{cases} 1, & \text{if } m = n \\ 0, & \text{if } m \neq n \end{cases}$$

denotes Kronecker's delta function. Then multiplying Equation (28) by z^n and taking summation over $n = 0, 1, \dots$, we obtain

$$P_N(z) = P_0(z) + E(X)P_1(z) + zP_2(z) \quad (29)$$

which represents the connection between the 4 PGFs namely, $P_0(z), P_1(z), P_2(z)$ and $P_N(z)$. After simplifying Equation (29), we obtain the result (25). The result (25), i.e., $P_N(z)$ can also be written as

$$P_N(z) = \frac{E(X)(1-z)}{\{1 - P_X(z)\}} P_\Pi(z) \quad (30)$$

which represents the relation between the PGF of system size and PGF of queue size distribution at service completion epoch.

The result (26) follows directly from Equation (25) on differentiation with respect to z and assigning $z = 1$. Again, calculating the second differentiation of Equation (25) with respect to z and assigning $z = 1$ and then adding the result with (25), after simplification we obtain the formula (27).

Remark 2. It is essential to note that the stochastic decomposition property (see Fuhrmann and Cooper (1985)) can be exhibited for the underlying queueing system.

We can write the Equation (25) as

$$P_N(z) = P_o(z)P_Y(z)$$

where $P_o(z) = \frac{(1-\lambda E(X)E(S)(1-z)P_L(z))}{(P_L(z)-z)}$, the PGF of system size distribution of an ordinary $M^X/G/1$ queue and

$$P_Y(z) = \frac{1 - (1 - P_Y(\vartheta^*(\lambda))) \left[\frac{(P_Y(z) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))P_X(z)}{1 - P_Y(\vartheta^*(\lambda))} \right]}{(1 - P_Y(\vartheta^*(\lambda))) \left[\frac{\lambda E(\vartheta)E(X)}{(1 - \vartheta^*(\lambda))} + \frac{P_Y(\vartheta^*(\lambda))E(X)}{1 - P_Y(\vartheta^*(\lambda))} \right]} (1 - z)$$

is the PGF of additional system size distribution when the server is idle.

Thus, $P_N(z) = P_o(z)P_Y(z)$ confirms that the stochastic decomposition property of Fuhrmann and Cooper (1985) is also valid for the underlying queueing system under the stability condition.

5. System performance measures

In this section, we derive some essential system performance measures namely distribution of a number of batches served during the busy period, waiting time distribution and unfinished work.

5.1. Distribution of a number of batches served during the busy period

Here, we derive the PGF and mean number of batches served during the busy period in this model under stability condition $\lambda E(X)E(S) < 1$, which are given in the following theorem.

Theorem 2. Let ψ denote the number of batches served during the busy period. Then, under stability condition $\lambda E(X)E(S) < 1$, the corresponding PGF $P_\psi(z)$ and mean $E(\psi)$ are determined as

$$P_\psi(z) = (1 - P_Y(\vartheta^*(\lambda))) \left\{ \frac{(P_V(P(z)) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))P_X(P(z))}{1 - P_Y(\vartheta^*(\lambda))} \right\} \quad (31)$$

and

$$E(\psi) = \frac{E(X)P_Y(\vartheta^*(\lambda))}{1 - \lambda E(X)E(S)} \left[1 + \frac{\lambda E(\vartheta)(1 - P_Y(\vartheta^*(\lambda)))}{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))} \right] \quad (32)$$

Proof. The idle period is terminated if and only if at least one batch arrives either during a random vacation period or dormant period. Therefore, the probability that several batches with total of n customers arrives in the system at the idle period completion epoch is

$$a_n = \{1 - P_Y(\vartheta^*(\lambda))\} \frac{\sum_{k=1}^n b_n^{\odot(k)} q_k}{(1 - q_0)} + P_Y(\vartheta^*(\lambda)) b_n; n = 1, 2, \dots$$

Now, if there are several batches with total of n customers in the system at the idle period completion epoch, then utilizing the argument of Takagi (1991), we have

$$\begin{aligned} P_\psi(z) &= \sum_{n=1}^{\infty} a_n (P(z))^n \\ &= (1 - P_Y(\vartheta^*(\lambda))) \sum_{n=1}^{\infty} \frac{\sum_{k=1}^n b_n^{\odot(k)} q_k}{(1 - q_0)} (P(z))^n + P_Y(\vartheta^*(\lambda)) \sum_{n=1}^{\infty} b_n (P(z))^n \\ &= \frac{(1 - P_Y(\vartheta^*(\lambda)))}{1 - \vartheta^*(\lambda)} \left[\sum_{n=0}^{\infty} (P(z))^n \sum_{k=0}^n b_n^{\odot(k)} \int_0^{\infty} e^{-\lambda x} \frac{(\lambda x)^k}{k!} dF_{\vartheta}(x) - (P(z))^0 b_n^{\odot(0)} \int_0^{\infty} e^{-\lambda x} dF_{\vartheta}(x) \right] \\ &\quad + P_Y(\vartheta^*(\lambda)) \sum_{n=1}^{\infty} b_n (P(z))^n \\ &= (1 - P_Y(\vartheta^*(\lambda))) \left\{ \frac{(P_V(P(z)) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))P_X(P(z))}{1 - P_Y(\vartheta^*(\lambda))} \right\} \end{aligned} \quad (33)$$

where $P(z)$ follows the following function equation under stability

$$P(z) = P_X(z) S^*(\lambda - \lambda P_X(P(z)))$$

Then differentiating Equation (33) with respect to z and assigning $z=1$, we obtain the result (32).

5.2. Waiting time distribution

To obtain the waiting time distribution of a queue, first of all, we derive the waiting time of the first customer in an arriving batch which is denoted by W_1 . Next, we derive the waiting time of a random customer in a batch which is denoted by W .

By considering a batch with a single customer, its (the batch's) service time is the total service time of customer constituting the batch. The LST of this service time is $P_X(S^*(\theta))$. In this case the batch will have as its batch size $P_X(z) = z$ and the mean arrival rate of the batch is λ . Using the information and results obtained by Chaudhry and Templeton (1983) [see Chapter 3, page 147–150] and from Equation (5), we have

$$P_{\Pi}(z) = \frac{(1 - \lambda E(S))(1 - \vartheta^*(\lambda)) \left[\frac{1}{(1 - P_Y(\vartheta^*(\lambda)))} - \left\{ \frac{(\vartheta^*(\lambda - \lambda z) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda))z}{1 - P_Y(\vartheta^*(\lambda))} \right\} \right] P_X(S^*(\lambda - \lambda z))}{\left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] (P_X(S^*(\lambda - \lambda z)) - z)} \quad (34)$$

Utilizing above explanation and result (34), we find the results for waiting time distribution in the following theorem:

Theorem 3. *If $\lambda E(X)E(S) < 1$, then the LST of waiting time distribution of the first customer in an arriving batch is $W_1^*(\theta)$, the LST of waiting time distribution of a random customer in a batch is $W^*(\theta)$, the mean waiting time is $E(W)$ and the mean sojourn time is $E(W_s)$, of the underlying queueing system. These are obtained as*

$$\begin{aligned} W_1^*(\theta) &= \frac{(1 - \lambda E(X)E(S))[(1 - \vartheta^*(\lambda))^2 - (1 - P_Y(\vartheta^*(\lambda)))(1 - \vartheta^*(\lambda))(\vartheta^*(\theta) - \vartheta^*(\lambda))]}{(1 - P_Y(\vartheta^*(\lambda)))(1 - \vartheta^*(\lambda)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] \left[P_X(S^*(\theta)) - 1 + \frac{\theta}{\lambda} \right]} \\ &\quad + \frac{(1 - \lambda E(X)E(S)) \left[\left(\frac{\theta}{\lambda} - 1 \right) P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))^2 \right]}{(1 - P_Y(\vartheta^*(\lambda)))(1 - \vartheta^*(\lambda)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] \left[P_X(S^*(\theta)) - 1 + \frac{\theta}{\lambda} \right]} \quad (35) \\ W^*(\theta) &= \frac{(1 - \lambda E(X)E(S))(1 - \vartheta^*(\lambda))^2 \left[\sum_{k=1}^{\infty} (S^*(\theta))^{k-1} - \sum_{k=1}^{\infty} \sum_{n=1}^{k-1} b_n (S^*(\theta))^{k-1} \right]}{E(X)(1 - P_Y(\vartheta^*(\lambda)))(1 - \vartheta^*(\lambda)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] \left[P_X(S^*(\theta)) - 1 + \frac{\theta}{\lambda} \right]} \\ &\quad - \frac{(1 - \lambda E(X)E(S))((1 - \vartheta^*(\lambda))(\vartheta^*(\theta) - \vartheta^*(\lambda)) \sum_{k=1}^{\infty} (S^*(\theta))^{k-1} \left[1 - \sum_{n=1}^{k-1} b_n \right])}{E(X)(1 - P_Y(\vartheta^*(\lambda)))(1 - \vartheta^*(\lambda)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] \left[P_X(S^*(\theta)) - 1 + \frac{\theta}{\lambda} \right]} \quad (36) \\ &\quad + \frac{(1 - \lambda E(X)E(S))P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))(\vartheta^*(\theta) - \vartheta^*(\lambda)) \sum_{k=1}^{\infty} (S^*(\theta))^{k-1} \left[1 - \sum_{n=1}^{k-1} b_n \right]}{E(X)(1 - P_Y(\vartheta^*(\lambda)))(1 - \vartheta^*(\lambda)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] \left[P_X(S^*(\theta)) - 1 + \frac{\theta}{\lambda} \right]} \\ &\quad + \frac{(1 - \lambda E(X)E(S)) \left(\frac{\theta}{\lambda} - 1 \right) P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda)) \sum_{k=1}^{\infty} (S^*(\theta))^{k-1} \left[1 - \sum_{n=1}^{k-1} b_n \right]}{E(X)(1 - P_Y(\vartheta^*(\lambda)))(1 - \vartheta^*(\lambda)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] \left[P_X(S^*(\theta)) - 1 + \frac{\theta}{\lambda} \right]} \end{aligned}$$

$$\begin{aligned}
E(W) = & \frac{(\lambda(E(X))^2 E(S^2) + E(X^2)E(S)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]}{2E(X)(1 - \lambda E(X)E(S)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \\
& + \frac{\lambda E(\vartheta^2)E(X)(1 - \lambda E(X)E(S))}{2E(X)(1 - \lambda E(X)E(S)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \\
& + \frac{\lambda E(\vartheta^2)E(X)(1 - \lambda E(X)E(S))}{2E(X)(1 - \lambda E(X)E(S)) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]}
\end{aligned} \tag{37}$$

$$\begin{aligned}
E(W_s) = & \frac{[2E(X)E(S) + E(X^2)E(S) + \lambda(E(X))^2(E(S^2) - 2(E(S))^2)]}{2E(X)(1 - \lambda E(X)E(S))} \\
& + \frac{\lambda E(\vartheta^2)E(X)}{2E(X) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]}
\end{aligned} \tag{38}$$

Proof. Consider that the waiting time of each batch does not depend on the arrival process. Then the arrival time of batches which are left in the system are those, where arrival stays in the system (i.e., in queue and in service). It follows from Equation (34) that (see Fuhrmann and Cooper (1985))

$$P_{\Pi}(z) = W_1^*(\lambda - \lambda z)P_X(S^*(\lambda - \lambda z)) \tag{39}$$

Now, assigning $\theta = \lambda - \lambda z$ in (39) and utilizing Equation (34), we obtain the result (35).

Suppose, the arrival customers are selected randomly in a batch. Then, the probability of a random customer being at k th position in its arrival batch is

$$\varsigma_k = \frac{1 - \sum_{n=1}^{k-1} b_n}{E(X)}; k = 1, 2, \dots$$

Now, as the waiting time of this k^{th} customer in a batch is the sum of the waiting time of the first customer in an arriving batch and the service times of the first $k - 1$ customers of the arriving batch. Hence it follows (since we are considering the random customer) that the LST of waiting time distribution of a random customer in a batch is given by (see Chaudhry and Templeton (1983))

$$W^*(\theta) = W_1^*(\theta) \times \sum_{k=1}^{\infty} \varsigma_k (S^*(\theta))^{k-1} \tag{40}$$

Now, exploiting (35) in (40), we get Equation (36). The result (37) is obtained directly from Equation (36) on differentiating both sides with respect to θ and putting $\theta = 0$. Adding $E(S)$ with (37), we obtain the result (38).

Corollary 3. The second moment of waiting time $E(W^2)$ and sojourn time $E(W_s^2)$ are given by

$$\begin{aligned}
 E(W^2) = & \frac{\lambda(E(X))^2(E(S^3) + (1 - \lambda E(X)E(S))E(S^2))}{2E(X)(1 - \lambda E(X)E(S))^2} \\
 & + \frac{E(S)(E(X^3) - 2\lambda(E(X^2))^2E(S) + (1 - \lambda E(X)E(S))E(X^2))}{2E(X)(1 - \lambda E(X)E(S))^2} \\
 & + \frac{\lambda E(X)(E(S^2)E(X^2) + (E(S))^2E(X^3)) + \lambda^2(E(X))^3(E(S)E(S^3) - (E(S^2))^2)}{2E(X)(1 - \lambda E(X)E(S))^2} \quad (41) \\
 & + \frac{E(X^2)E(S^2)(E(X) - E(X^2))}{2(E(X))^2(1 - \lambda E(X)E(S))^2} + \frac{\lambda(E(\vartheta^3) + E(\vartheta^2))}{2 \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]}
 \end{aligned}$$

and

$$\begin{aligned}
 E(W_s^2) = & E(X^2) + \frac{\lambda(E(X))^2E(S^3) + \lambda E(X)(E(S^2)E(X^2) + E(S)E(X^3))}{2E(X)(1 - \lambda E(X)E(S))^2} \\
 & + \frac{E(S)(E(X^3) - 2\lambda(E(X^2))^2E(S) + \lambda^2(E(S))^3(E(S)E(S^3) - (E(S^2))^2))}{2E(X)(1 - \lambda E(X)E(S))^2} \\
 & + \frac{[E(X)(1 - \lambda E(X)E(S))(2E(X)E(S) + E(X^2)E(S) + \lambda(E(X))^2(E(S^2) - 2(E(S))^2))]}{2(E(X))^2(1 - \lambda E(X)E(S))^2} \\
 & + \frac{E(X^2)E(S^2)(E(X) - E(X^2))}{2(E(X))^2(1 - \lambda E(X)E(S))^2} + \frac{\lambda(E(\vartheta^3) + E(\vartheta^2))}{2 \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (42)
 \end{aligned}$$

5.3. Unfinished work

In this subsection, the main objectives are to determine the LST of the distribution function of the unfinished work and the mean unfinished work. The following theorem summarizes the results:

Theorem 4. Let ϖ be the r.v representing the length of an unfinished work and $\varpi^*(\theta)$ be its corresponding LST. For $\lambda E(X)E(S) < 1$,

$$\varpi^*(\theta) = \frac{\theta(1 - \lambda E(X)E(S)) \left[(1 - P_V(S^*(\theta))) + \frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} (1 - P_X(S^*(\theta))) \right]}{\left[\lambda E(\vartheta) + \frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right] (\theta - \lambda + \lambda P_X(S^*(\theta))) E(X)(1 - S^*(\theta))} \quad (43)$$

and

$$L_{\varpi} = E(\varpi) = \frac{\lambda^2 (E(\vartheta))^2 E(X) E(\vartheta_r)}{\left[\lambda E(\vartheta) + \frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} + \frac{\lambda E(X) E(S) E(S_r) + E(X_r)}{(1 - \lambda E(X) E(S))} \quad (44)$$

where $E(S_r) = \frac{E(S^2)}{2E(S)}$ and $E(\vartheta_r) = \frac{E(\vartheta^2)}{2E(\vartheta)}$ represents the mean residual service time and vacation time respectively. $E(X_r) = \frac{E(X^2)}{2E(X)}$ represent mean residual batch size.

Proof. Let χ be the r.v describing the length of an unfinished work of an idle period and $\chi^*(\theta)$ be its LST, then by utilizing the standard argument of queueing theory (see Takagi (1991)), we can write

$$\begin{aligned} \chi^*(\theta) &= P_Y(S^*(\theta)) \\ &= \frac{1 - (1 - P_Y(\vartheta^*(\lambda))) \left[\frac{(P_Y(S^*(\theta)) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{P_Y(\vartheta^*(\lambda)) P_X(S^*(\theta))}{1 - P_Y(\vartheta^*(\lambda))} \right]}{(1 - P_Y(\vartheta^*(\lambda))) \left[\frac{\lambda E(\vartheta) E(X)}{(1 - \vartheta^*(\lambda))} + \frac{P_Y(\vartheta^*(\lambda)) E(X)}{1 - P_Y(\vartheta^*(\lambda))} \right] (1 - S^*(\theta))} \end{aligned} \quad (45)$$

Since $\varpi^*(\theta)$ is the LST of the distribution function of the unfinished work, therefore by utilizing the decomposition result of Leung (1992), we can write

$$\varpi^*(\theta) = \chi^*(\theta) \varpi_{M^X/G/1}^*(\theta) \quad (46)$$

where $\varpi_{M^X/G/1}^*(\theta) = \frac{\theta(1 - \lambda E(X) E(S))}{\theta - \lambda + \lambda P_X(S^*(\theta))}$ is the LST of waiting time distribution of the first customer in a batch and it belongs to an ordinary $M^X/G/1$ queue. Thus, utilizing (45) in (46), the result (43) is obtained.

Again, by differentiating Equation (43) both sides with respect to θ and assigning $\theta = 0$, formula (44) is obtained.

Remark 3. It is necessary to note that the result (43) can be rearranged for the underlying queueing model and for that we need to define some other general descriptors which are listed below:

Let T_{RV} denote the length of a random vacation period during a regeneration cycle. Then the average residence time say $E(T_{RV})$, of the server in the random vacation period during a regeneration cycle is given by

$$E(T_{RV}) = \frac{E(X) P_Y(\vartheta^*(\lambda)) E(\vartheta)}{\left[\frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (47)$$

Consider T_D as the length of a dormant period during a regeneration cycle. Then the average residence time say $E(T_D)$, of the server in the dormant period during a regeneration cycle is given by

$$E(T_D) = \frac{E(X) P_Y(\vartheta^*(\lambda))}{\lambda} \quad (48)$$

Similarly, consider T_I as the length of an idle period during a regeneration cycle. Then the average residence time say $E(T_I)$, of the server in the idle period during a regeneration cycle is given by

$$E(T_I) = E(T_{RV}) + E(T_D) = \frac{E(X) \left[\lambda E(\vartheta) P_Y(\vartheta^*(\lambda)) + \frac{P_Y(\vartheta^*(\lambda))^2 (1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]}{\lambda \left[\frac{P_Y(\vartheta^*(\lambda)) (1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (49)$$

Now, by utilizing the theory of regenerative process, we get the following conditional time average probabilities $\bar{P}_{RV}$ and $\bar{P}_d$ as follows:

$$\begin{aligned} \bar{P}_{RV} &= P(\text{server is in random vacation} / \text{server is idle}) \\ &= \frac{E(T_{RV})}{E(T_I)} \\ &= \frac{\lambda E(\vartheta)}{\left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda)) P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \end{aligned} \quad (50)$$

and

$$\begin{aligned} \bar{P}_d &= P(\text{server is in dormant period} / \text{server is idle}) \\ &= \frac{E(T_D)}{E(T_I)} \\ &= \frac{(1 - \vartheta^*(\lambda)) P_Y(\vartheta^*(\lambda))}{(1 - P_Y(\vartheta^*(\lambda))) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda)) P_Y(\vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \end{aligned} \quad (51)$$

Using (50) and (51), we can rearrange the Equation (43) as

$$\varpi^*(\theta) = \frac{\bar{P}_d(\theta \lambda E(\vartheta) - \lambda^2 E(X) E(\vartheta) E(S))(1 - P_X(S^*(\theta))) + \bar{P}_{RV}(\theta - \theta \lambda E(X) E(S))(1 - P_V(S^*(\theta)))}{\lambda E(\vartheta) E(X) (1 - S^*(\theta)) (\theta - \lambda + \lambda P_X(S^*(\theta)))} \quad (52)$$

Now, by assigning $\bar{P}_d = 0$ and $\bar{P}_{RV} = 1$ in Equation (52), the result justifies the Equation (3.7) of Takagi (1991). Similarly, by putting $\bar{P}_d = 1$ and $\bar{P}_{RV} = 0$ in Equation (52), the result justifies with Takagi (1991).

6. Some special cases

Here, we discuss momentarily some particular cases of the model, which is consistent with the existing literature.

Case 1: If we are considering point mass distribution at Y , then $P_Y(\vartheta^*(\lambda)) = (\vartheta^*(\lambda))^Y$ and therefore Equation (5) yields

$$P_{\Pi}(z) = \frac{(1 - \lambda E(X)E(S))(1 - \vartheta^*(\lambda)) \left[\frac{1}{(1 - (\vartheta^*(\lambda))^Y)} - \left\{ \frac{(P_V(z) - \vartheta^*(\lambda))}{1 - \vartheta^*(\lambda)} + \frac{(\vartheta^*(\lambda))^Y P_X(z)}{1 - (\vartheta^*(\lambda))^Y} \right\} \right] P_L(z)}{E(X) \left[\lambda E(\vartheta) + \frac{(1 - \vartheta^*(\lambda))(\vartheta^*(\lambda))^Y}{1 - (\vartheta^*(\lambda))^Y} \right] (P_L(z) - z)} \quad (53)$$

Equations (53) represent the PGF of queue size distribution at a service completion epoch of the $M^X/G/1$ with atmost Y vacation. Also we note that Equation (53) is consistent with Ke and Chu (2006)

Case 2: Suppose Y follows geometric distribution, then the probability mass function of the r.v Y is

$$y_i = p(1 - \delta_{i,J})\bar{p}^{i-1} + \delta_{i,J}\bar{p}^{J-1}, i = 1, 2, \dots, J \quad (54)$$

then the Equations (5) and (38) yield

$$P_{\Pi}(z) = \frac{(1 - \lambda E(X)E(S)) \left(\frac{(1 - (\bar{p}\vartheta^*(\lambda))^J)(P_V(z) - 1)}{\vartheta^*(\lambda)[(\bar{p}\vartheta^*(\lambda))^{J-1}(1 - \bar{p}\vartheta^*(\lambda)) + p(1 - (\bar{p}\vartheta^*(\lambda))^{J-1})]} - 1 + P_X(z) \right) P_L(z)}{E(S) \left(1 + \frac{\lambda E(\vartheta)(1 - (\bar{p}\vartheta^*(\lambda))^J)}{\vartheta^*(\lambda)[(\bar{p}\vartheta^*(\lambda))^{J-1}(1 - \bar{p}\vartheta^*(\lambda)) + p(1 - (\bar{p}\vartheta^*(\lambda))^{J-1})]} \right) (z - P_L(z))} \quad (55)$$

and

$$E(W_s) = E(S) + \frac{E(X)E(S) + \lambda(E(X))^2 E(S^2)}{2E(X)(1 - \lambda E(X)E(S))} + \frac{\lambda(1 - (\bar{p}\vartheta^*(\lambda))^J)E(\vartheta^2)}{2(\lambda E(\vartheta)(1 - (\bar{p}\vartheta^*(\lambda))^J) + \vartheta^*(\lambda)[(\bar{p}\vartheta^*(\lambda))^{J-1}(1 - \bar{p}\vartheta^*(\lambda)) + p(1 - (\bar{p}\vartheta^*(\lambda))^{J-1})])} \quad (56)$$

Equations (55) and (56) represents the PGF of queue size distribution at a service completion epoch and average sojourn time of the $M^X/G/1$ system having randomized vacation policy with at most J vacations respectively. Note that for J tends to ∞ and $p = 0$, the above Equations (55) and (56) are in accordance with the results obtained by Choudhury (2002b). Also, we note that the Equations (55) and (56) are consistent with Ke, Huang, and Pearn (2010).

7. Cost optimal policy

In this section, we develop a long run average cost function for $M^X/G/1$ queue in which we can compare different random vacation policies. To compare it, we build a homologous cost function that has been widely used in various practical works (see Lee and Srinivasan (1989), Tadj and Choudhury (2005)). For this purpose, we need to find out an explicit expression for $E(S_c)$ defined as average length of the busy cycle which is given in the following theorem.

Theorem 5. Under stability condition $\lambda E(X)E(S) < 1$,

$$E(S_c) = \frac{\lambda P_Y(\vartheta^*(\lambda))E(\vartheta) + \frac{P_Y(\vartheta^*(\lambda))^2(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))}}{\lambda(1 - \lambda E(X)E(S)) \left[\frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (57)$$

Proof. Let S_c denote the length of a busy cycle of a system which consists of a server's random vacation period (T_{RV}), dormant period (T_D) and a busy period (T_S) during a regeneration cycle, i.e., the cycle length S_c can be written as $S_c = T_{RV} + T_D + T_S$.

Now, we consider $T_{(1)}$ as the length of a busy period of the server started with a batch of one customer. Then, under stability condition $\lambda E(X)E(S) < 1$, the LST of $T_{(1)}$ in terms of Takacs functional equation (see Takagi (1991)) is given by

$$T_{(1)}^*(\theta) = S^*(\lambda + \theta - \lambda P_X(T_{(1)}^*(\theta))) \quad (58)$$

Differentiating the result of Equation (58) with respect to θ and taking $\theta = 0$, the average length of a busy period initiated by a batch of one customer is found as

$$E(T_{(1)}) = \frac{\lambda E(X)E(S)}{\lambda(1 - \lambda E(X)E(S))} \quad (59)$$

Now, we are going to find out the average residence time say $E(T_S)$, during which the server is in busy period during a regeneration cycle. Clearly if a batch of minimum one customer arrives during the random vacation period (with probability $1 - P_Y(\vartheta^*(\lambda))$), a busy period with a batch of one customer is started. In this case, the average residence time of the server in the busy period during a regeneration cycle (type 1 cycle) is

$$\begin{aligned} E(T_S^1) &= \frac{E(X)P_Y(\vartheta^*(\lambda))E(\vartheta)}{\left[\frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} E(T_{(1)}) \\ &= \frac{E(X)P_Y(\vartheta^*(\lambda))E(\vartheta)}{\left[\frac{P_Y(\vartheta^*(\lambda))(1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \frac{\lambda E(X)E(S)}{\lambda(1 - \lambda E(X)E(S))} \end{aligned} \quad (60)$$

Now, we presume the case that the random vacation period has expired without any batch and the server enters into a dormant period. In this case, if a batch of minimum one customer arrives during the dormant period, a busy period is initiated. Then, the average residence time of the server in the busy period during a regeneration cycle (type 2 cycle) is

$$E(T_S^2) = \frac{E(X)P_Y(\vartheta^*(\lambda))}{\lambda} E(T_{(1)}) = \frac{(E(X))^2 P_Y(\vartheta^*(\lambda))E(S)}{\lambda(1 - \lambda E(X)E(S))} \quad (61)$$

Therefore, the total average residence time of the server in busy period during a regeneration cycle is given by

$$E(T_S) = E(T_S^1) + E(T_S^2) \quad (62)$$

Finally, adding Equations (47) (48), and (62), we obtain

$$E(S_c) = E(T_{RV}) + E(T_D) + E(T_S) = \frac{\lambda E(\vartheta) P_Y(\vartheta^*(\lambda)) + \frac{P_Y(\vartheta^*(\lambda))^2 (1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))}}{\lambda (1 - \lambda E(X) E(S)) \left[\frac{P_Y(\vartheta^*(\lambda)) (1 - \vartheta^*(\lambda))}{1 - P_Y(\vartheta^*(\lambda))} \right]} \quad (63)$$

which represents the average length of a busy cycle of the system.

7.1. Average cost function

Here, the purpose is to establish an optimum random vacation policy in order to minimize the average cost function which is stated below. Let us define:

C_{su} = Setup cost per unit time for each busy cycle in the system. Thus, in long run the average setup cost is $\frac{C_{su}}{E(S_c)}$.

C_{hl} = Holding cost per unit time for each batch available in the system. Therefore, in long run the average holding cost is $C_{hl}E(\Pi)$

C_{RV} = Per unit time cost while the server is in random vacation. Therefore, in long run the average cost owing to server is in random vacation is $C_{RV}P_{RV}$.

C_D = Per unit time cost while the server is in dormant. Thus, in long run the average cost owing to server is in dormant is C_DP_D .

C_S = Per unit time cost for holding the server busy. Therefore, in long run the average cost owing to server being busy is $C_S\lambda E(X)E(S)$.

Using the above cost parameter and renewal reward theory (see Ross (1983)), in long run the average cost per unit time is given by

$$T(RVP) = C_{su} \frac{1}{E(S_c)} + C_{hl}E(\Pi) + C_{RV}P_{RV} + C_DP_D + C_S\lambda E(X)E(S) \quad (64)$$

We determine the optimal value of the control parameter by minimizing the average cost function obtained in Equation (64). To find the optimum random vacation policy, i.e., where the cost is minimum, we have to select a locally optimal vacation policy RVP^* so that in each selection it satisfies the following inequality

$$T(RVP^1) \geq T(RVP^2) \geq T(RVP^*)$$

where $T(RVP^*)$ depict the expected cost per unit time of locally optimal random vacation policy and $T(RVP^1)$ and $T(RVP^2)$ represents the expected cost per unit time of other two random vacation policies with different probability distributions of the discrete r.v Y . The above inequality holds because locally optimal random vacation policy (RVP^*) will always give the minimum expected cost per unit time and hence, the expected cost per unit time of locally optimal random vacation policy ($T(RVP^*)$) is less than or equal to the other two policies. The relation between the other two policies $T(RVP^1)$ and $T(RVP^2)$ is first policy $T(RVP^1)$ will be greater than or equal to the other one $T(RVP^2)$.

In order to show the applicability of cost optimal policy in the problem, we present six numerical examples in Example 1 in numerical section for three different service time distributions and different cost parameters.

In the numerical examples, we consider IEEE 802.16e standard as describe in [section 2](#). For the purpose of numerical examples, We presume that the sleep time follows an exponential distribution with PDF $f_{\vartheta}(x) = \nu e^{-\nu x}, x > 0$.

8. Numerical illustration

In this section, we present the numerical results for the performance measures that we have obtained in the previous section for IEEE 802.16e standard which is described in [Section 2](#). Here, basically we are going to show the impacts of system parameters on the previously obtained key performance measures in the form of graphs. In the previous sections, we have obtained the probability P_{RV} of a server is in random vacation period (which is a sleep period in IEEE802.16e standard), the probability P_D of a server is in dormant period (Warm-up period), the mean waiting time $E(W)$, mean number $E(\psi)$ of batches served during the busy period (service period) and mean unfinished work L_{ϖ} .

Here, the packet arrival rate in arriving batch is λ . For the purpose of numerical illustration, we consider that the batch size has a geometric distribution starting with minimum one customer, then using the parameter b we have the PMF $P\{X_i = k\} = b(1 - b)^{k-1}; k = 1, 2, \dots, 0 < b < 1$, PGF $P_X(z) = \frac{bz}{1 - (1-b)z}$, mean $E(X) = \frac{1}{b}$ and second moment $E(X^2) = \frac{2-b}{b^2}$ respectively.

Next, we assume that the vacation time which is considered as a sleep period in IEEE802.16e standard follows an exponential distribution with parameter ν , then we have the PDF $f_{\vartheta}(x) = \nu e^{-\nu x}, x > 0$, LST $\vartheta^*(\theta) = \frac{\nu}{\nu + \theta}$, mean $E(\vartheta) = \frac{1}{\nu}$ and second moment $E(\vartheta^2) = \frac{2}{\nu^2}$ respectively.

Again, we presume that the service time conforms to:

- 3-stage Erlang distribution (denoted by 3-stage Erlang) with parameter μ , then we have the PDF $f_S(x) = \frac{1}{2}\mu^3 x^2 e^{-\mu x}, x > 0$, LST $S^*(\theta) = \frac{3\mu}{(3\mu + \theta)^2}$, mean $E(S) = \frac{3}{\mu}$ and second moment $E(S^2) = \frac{3(3+1)}{\mu^2}$ respectively.
- Exponential distribution (denoted by Exp) with parameter μ , then we have the PDF $f_S(x) = \mu e^{-\mu x}, x > 0$, LST $S^*(\theta) = \frac{\mu}{\mu + \theta}$, mean $E(S) = \frac{1}{\mu}$ and second moment $E(S^2) = \frac{2}{\mu^2}$ respectively.
- Hyper-Exponential distribution (denoted by Hyper-Exp) with parameter μ , then we have the PDF $f_S(x) = 0.5\mu e^{-\mu x} + 0.5\mu^2 e^{-\mu^2 x}, x > 0$, LST $S^*(\theta) = \frac{0.5\mu}{\mu + \theta} + \frac{0.5\mu^2}{\mu^2 + \theta}$, mean $E(S) = \frac{0.5(\mu+1)}{\mu^2}$ and second moment $E(S^2) = \frac{\mu+1}{\mu^2}$ respectively.

Now, we consider that the maximum number of sleep Y follows:

- Degenerate distribution (denoted by Deg) at point j , then we have the PMF $P\{Y = j\} = 1$, PGF $P_Y(z) = \frac{1}{1-z}$, mean $E(X) = j$ and second moment $E(X^2) = j^2$ respectively.
- Geometric distribution (denoted by Geo) with parameter p , then we have the PMF $P\{Y = j\} = p(1 - p)^{j-1}; j = 1, 2, \dots, 0 < p < 1$, PGF $P_Y(z) = \frac{pz}{1 - (1-p)z}$, mean $E(Y) = \frac{1}{p}$ and second moment $E(Y^2) = \frac{2-p}{p^2}$ respectively.

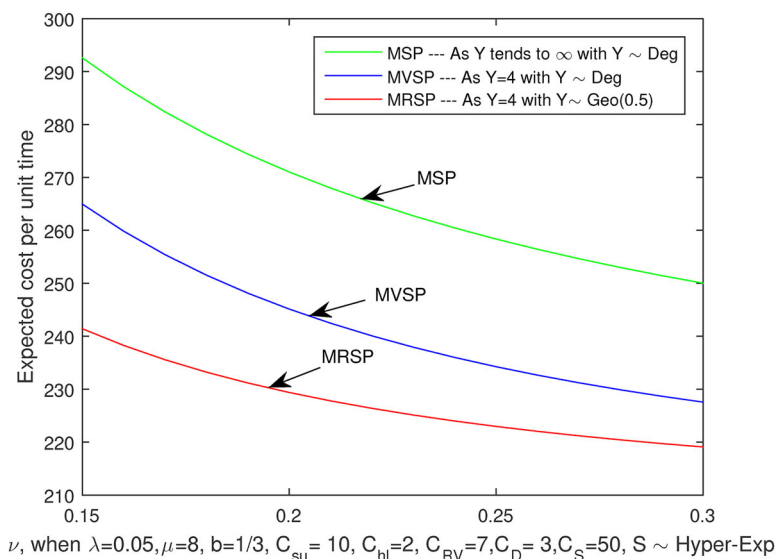


Figure 4. Impact of ν on expected cost per unit time.

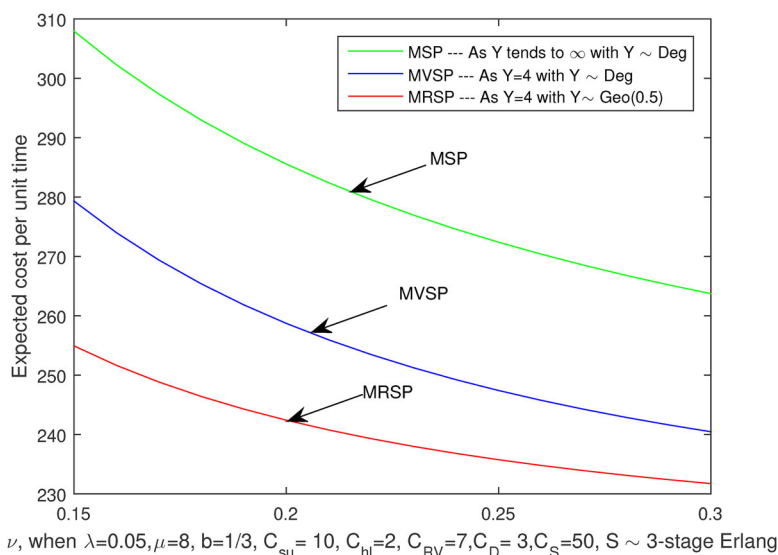


Figure 5. Impact of ν on expected cost per unit time.

Now, we are going to display some examples of IEEE802.16e standard with different parameters i.e., for different parametric values, the IEEE 802.16e standard apply in different systems for example WIMAX, WIFI, 3 G etc.

Example 1. Considering the cost parameters as $C_{su} = 10, C_{hl} = 2, C_{RV} = 7, C_D = 3, C_S = 50$, model parameters as $\lambda = 0.05$ and $\mu = 8$, we displays the impact of sleep (vacation) rate ν on expected cost per unit time (64) for three sleep (vacation) policies as well as for three service time distributions in Figures 4–6. Figures 7–9 also show the effect of ν on expected cost per unit time (64) for three sleep policies as well as for three service

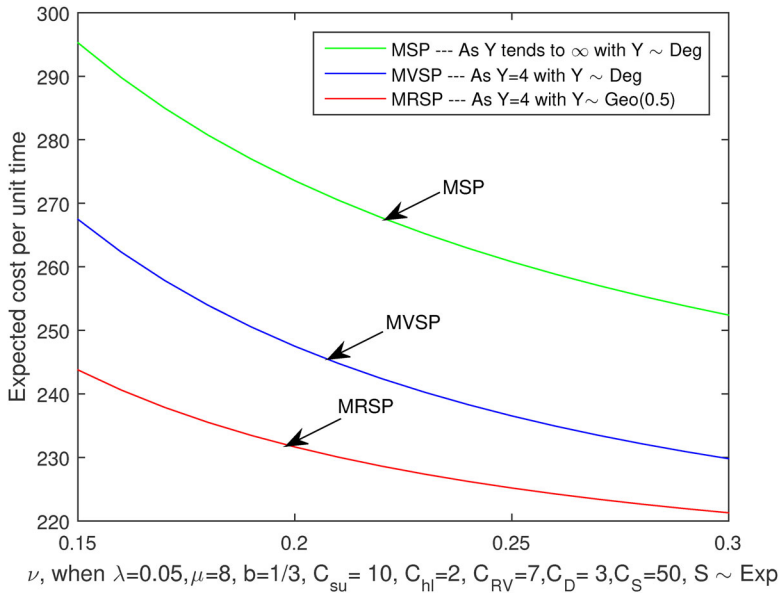


Figure 6. Impact of ν on expected cost per unit time.

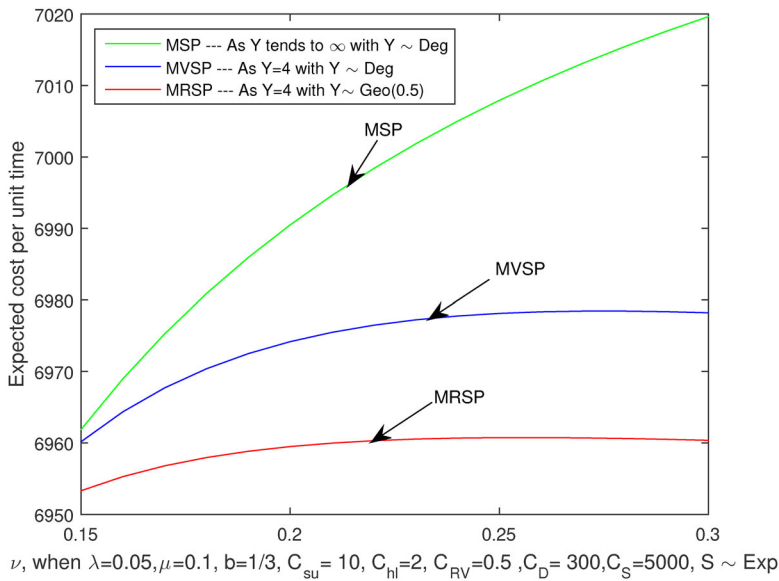


Figure 7. Impact of ν on expected cost per unit time.

time distributions but here $\mu = 0.1, C_{RV} = 0.5, C_D = 300, C_S = 5000$. Other parameters are same as in Figures 4–6.

In Figures 4–6 as well as in Figures 7–9, we compare the expected cost per unit time of three sleep (vacation) policies: multiple sleep policy (MSP), multiple variant sleep policy (MVSP) and multiple randomized sleep policy (MRSP) for three service time distributions namely, 3-stage Erlang, exponential and hyper-exponential. MSP is a policy where the server takes multiple numbers of consecutive sleeps until it finds one batch

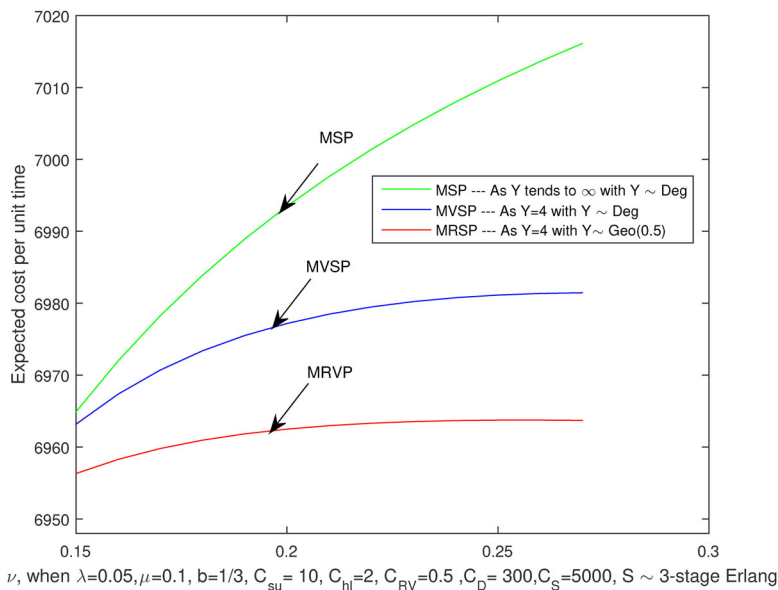


Figure 8. Impact of ν on expected cost per unit time.

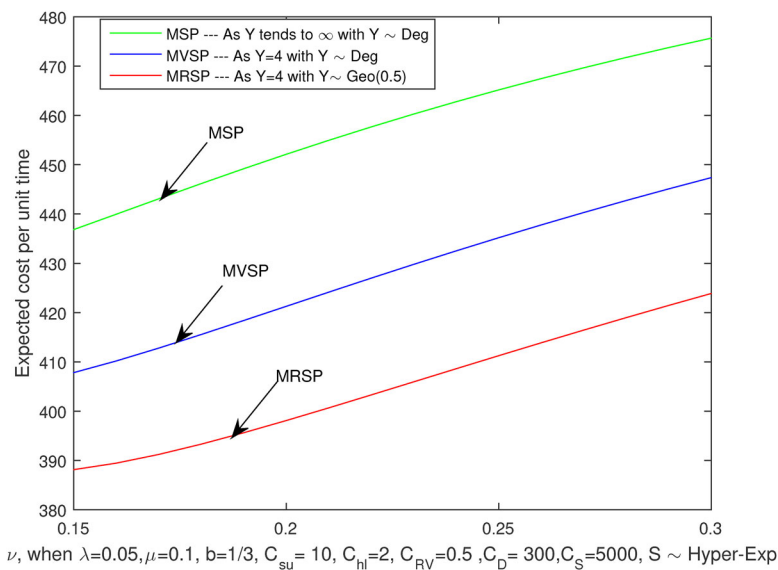


Figure 9. Impact of ν on expected cost per unit time.

waiting in a queue at a sleep period completion epoch. MVSP is a policy where the server takes a fixed number of multiple sleeps until it finds at least one batch waiting in a queue at a sleep period completion epoch. If no batch is found after completing a fixed number of multiple sleeps, the server goes for the warm-up state and waits for the first arrival. MRSP is described as follows: After the system becomes empty, the server immediately goes for a sleep. If there is at least one batch waiting in the queue upon returning from a sleep, the server is immediately activated for service. Otherwise, if no

batch is waiting in a queue at a sleep period completion epoch, the server either remains idle in the system with probability p or leaves for another sleep with probability $1 - p$. This pattern continues until a fixed number of multiple sleeps are completed. If the system is still empty after completing a fixed number of multiple sleeps, the server goes for the warm-up state and waits for the first arrival. The effect of ν on expected cost per unit time are also shown in Figures 4–6, when $C_{su} = 10, C_{hl} = 2, C_{RV} = 7, C_D = 3, C_S = 50$, model parameters $\lambda = 0.05$ and $\mu = 8$, and in Figures 7–9, when $C_{su} = 10, C_{hl} = 2, C_{RV} = 0.5, C_D = 300, C_S = 5000$, model parameters $\lambda = 0.05$ and $\mu = 0.1$. We summarize some findings:

1. When Y tends to ∞ and Y follows a degenerate distribution and $C_D = 3, C_{RV} = 7, C_S = 50$ and $\mu = 8$ in Figures 4–6, we can see that expected cost decreases as the sleep rate ν increases whereas for $C_D = 300, C_{RV} = 0.5, C_S = 5000$ and $\mu = 0.1$, in Figures 7–9, we notice that as the sleep rate ν increases, expected cost also increases before attains stability, this is because the model becomes an $M/G/1$ system with multiple sleep policy.
2. If $Y = 4$, i.e., a finite value and Y follows a degenerate distribution, the model becomes the $M/G/1$ queue with multiple variant sleep policy and for $C_D = 3, C_{RV} = 7, C_S = 50$ and $\mu = 8$, in Figures 4–6, we see that the expected cost per unit time decreases as ν increases whereas for $C_D = 300, C_{RV} = 0.5, C_S = 5000$ and $\mu = 0.1$, in Figures 7–9, we observe that as ν increases expected cost also increases before it reaches stability.
3. When $Y = 4$ (a finite value) and Y follows a geometric distribution with parameter $p = 0.5$ and $C_D = 3, C_{RV} = 7, C_S = 50$ and $\mu = 8$, in Figures 4–6, we can find that expected cost decreases as the values of ν increases whereas for $C_D = 300, C_{RV} = 0.5, C_S = 5000$ and $\mu = 0.1$, in Figures 7–9, we observe that as ν increases, expected cost also increases, this is because that the model becomes an $M/G/1$ system with randomized sleep policy.

It is clear that when $C_D = 3, C_S = 50$ i.e., comparatively small and $\mu = 8$ and $C_{RV} = 7$, i.e., comparatively large, expected cost decreases as the values of ν increases, on the other hand, when $C_D = 300, C_S = 5000$ i.e., comparatively large, and $C_{RV} = 0.5$ and $\mu = 0.1$ i.e., comparatively small, expected cost increases as the values of ν increases.

By comparing expected cost per unit time in terms of three vacation policies yields, $MSP > MVSP > MRSP$, i.e., $MRSP$ is better than the two other policies. We also observe that the expected cost per unit time is larger for 3-stage Erlang service time distribution as compared to exponential and hyper-exponential distribution.

Therefore, if we want to minimize the expected cost per unit time, we can consider multiple randomized sleep (vacation) policy as a optimal policy, which shows minimum expected cost per unit time compared to the expected cost of other two policies.

From a practical point of view, some other useful measures are the probability P_{RV} of a server when it is in sleep period, the Probability P_D of a server when it is in warm-up period, mean number $E(\psi)$ of batches served during the awake period and mean unfinished work L_m respectively.

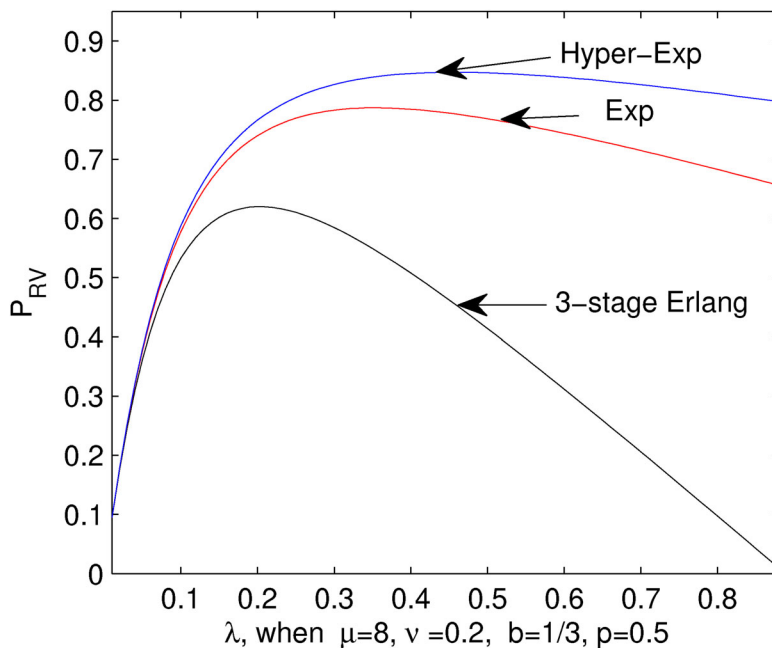


Figure 10. Impact of λ on P_{RV} for different service time distribution.

Example 2. Considering the sleep time as exponential distribution, while batch size and Y follows a geometric distribution and parametric values as $\mu = 8$, $\nu = 0.2$, $b = \frac{1}{3}$, $p = 0.5$, we exhibit the behavior of arrival rate λ on P_{RV} , P_D , $E(\psi)$ and L_w due to three service time distributions namely, 3-stage Erlang, exponential and hyper-exponential in Figures 10–13. In the following subsections, we summarize some findings:

Impact of packet arrival rate λ on probability that the server goes for sleep period P_{RV} for different service time distributions

In Figure 10, we observe that

1. While arrival rate λ increases the probability that the server goes for sleep period P_{RV} increases initially and then decrease before it reaches stability, i.e., while the packet arrival rate λ is comparatively smaller the probability that the server goes for sleep period P_{RV} increases whereas the arrival rate λ becomes larger the probability that the server goes for sleep period P_{RV} decreases and probability of going for awake period increases.
2. The hyper-exponential distribution provides highest system state probability P_{RV} , 3-stage Erlang provides lowest system state probability P_{RV} and exponential distribution remains in between hyper-exponential and 3-stage Erlang distribution. It is clear from the fact that a higher coefficient of variation (or variance) of distribution gives the higher P_{RV} values.

Impact of packet arrival rate λ on probability that the server goes for warm-up period P_D for different service time distributions

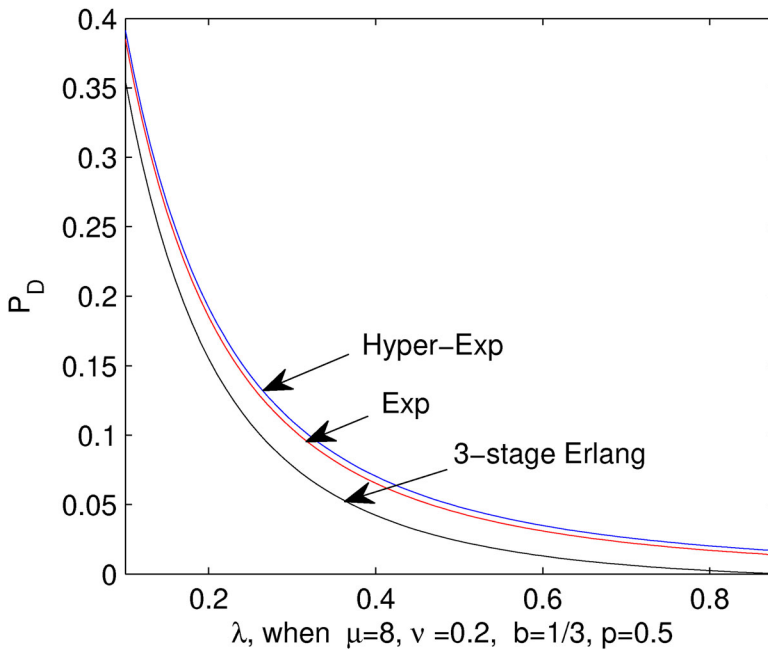


Figure 11. Impact of λ on P_D for different service time distribution.

From Figure 11, we see that

1. The system state probability P_D decreases with the increasing value of arrival rate λ as it should be i.e., as the arrival rate λ increases, the probability that the server goes for warm-up period (as it is a waiting period) decreases and the probability that server goes for awake period increases.
2. The comparison of three service time distributions namely, hyper-exponential, exponential and 3-stage Erlang for the system state probability P_D is similar as in Figure 7.

Impact of packet arrival rate λ on mean number $E(\psi)$ of batches served during the awake period and mean unfinished work L_m for different service time distributions

In Figures 12 and 13, we observe that

1. The mean $E(\psi)$ and L_m increases while the arrival rate λ increases i.e., as the values of arrival rate λ increases, the mean number of batches served during awake period and mean unfinished work increases i.e., as the arrival rate increases, mean number of batches also increases which increases the mean number of batches served during awake period and mean unfinished work.
2. For the fixed value of λ , the mean number of batches $E(\psi)$ served during awake period and mean unfinished work L_m are larger for 3-stage Erlang distribution compared to the other two distributions, i.e., exponential and hyper-exponential. This means that when the service time follows 3-stage Erlang distribution, the server can serve maximum number of batches during awake period as compared to other two distribution (since 3-stage Erlang distribution have lower coefficient of variation (C.V) and lower C.V value gives more precise estimate).

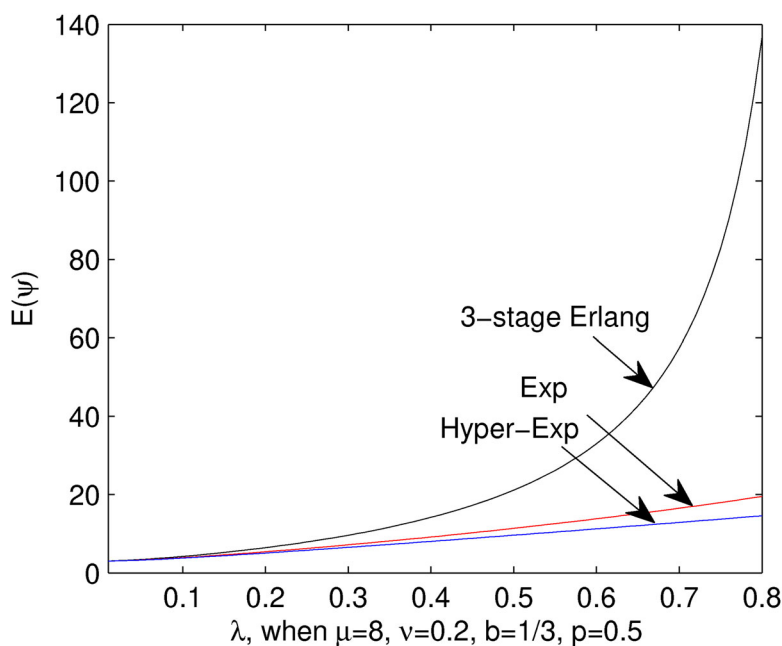


Figure 12. Impact of λ on $E(\psi)$ for different service time distribution.

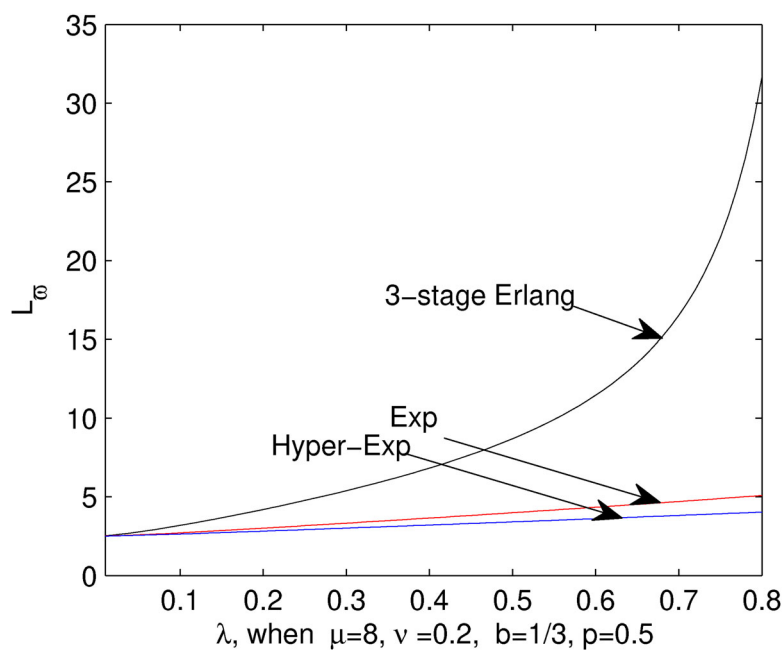


Figure 13. Impact of λ on L_w for different service time distribution.

Example 3. Presuming the model parameters as $\lambda = 0.05$, $\nu = 0.2$, $b = \frac{1}{3}$, $p = 0.5$, we display the impact of service rate μ on P_{RV} , P_D , $E(\psi)$ and L_w for varied service time distributions namely, 3-stage Erlang, exponential and hyper-exponential in Figures 14–17.

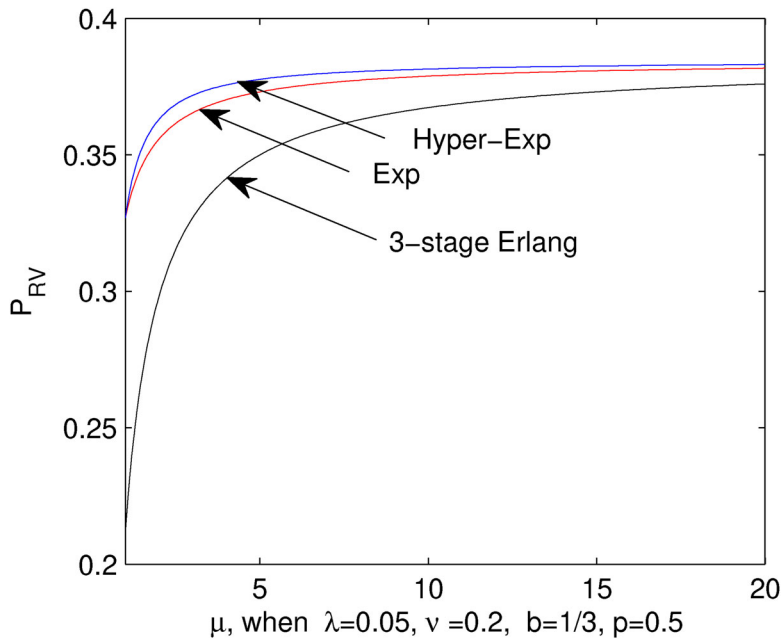


Figure 14. Impact of μ on P_{RV} for different service time distribution.

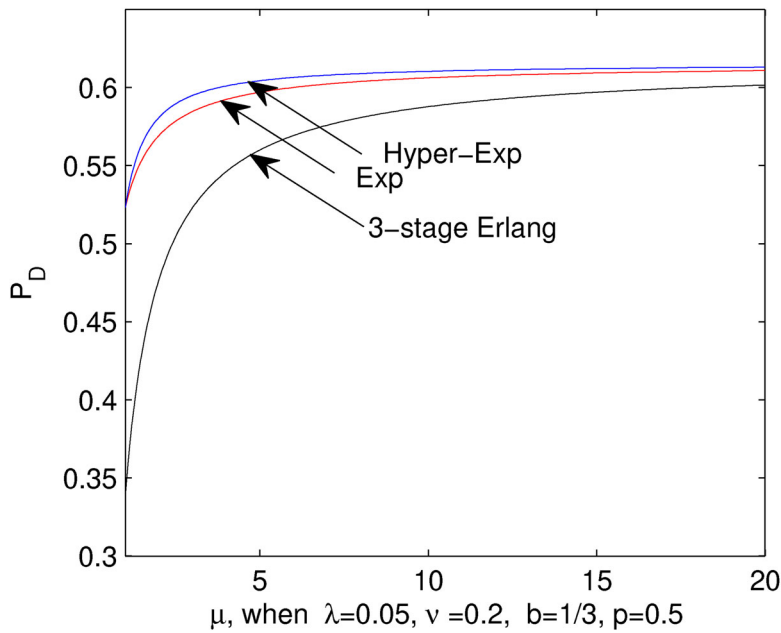


Figure 15. Impact of μ on P_D for different service time distribution.

Here, we consider that the sleep time follows an exponential distribution, batch size and Y follows a geometric distribution. In the following subsections, we summarize some findings:

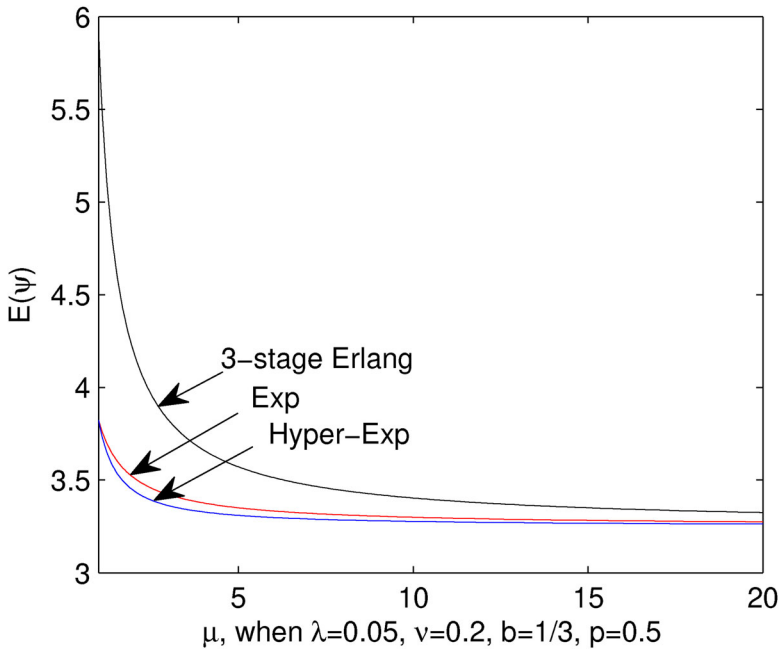


Figure 16. Impact of μ on $E(\psi)$ for different service time distribution.

Impact of service rate μ on probability that server goes for the sleep period P_{RV} and probability that server goes for the warm-up period P_D for different service time distributions

While the other model parameters are fixed, Figures 14 and 15 show

1. The system state probabilities P_{RV} and P_D both increase with the increasing value of service rate μ i.e., as the service rate (the rate at which service has been given) increases, the probability that the server completing the service of all packets in awake period increases which leads to increase the probability that the server goes for sleep period and warm-up period.
2. Hyper-exponential distribution provides highest system state probabilities P_{RV} and P_D whereas 3-stage Erlang provides the lowest and exponential stays in between hyper-exponential and 3-stage Erlang.

Impact of service rate μ on mean number $E(\psi)$ of batches served during the awake period and mean unfinished work L_w for different service time distributions

From Figures 16 and 17, we observe that

1. The averages $E(\psi)$ and L_w both decrease for the increasing value of μ before it arrives stability, which is completely opposite from Figures 11 and 12. This is because, as the service rate in the awake period increases the average service time for the batches decreases which decreases the mean number of batches served during awake period and mean unfinished work.
2. The averages $E(\psi)$ and L_w are larger for 3-stage Erlang service time distribution as compared to exponential and hyper-exponential service time distribution.

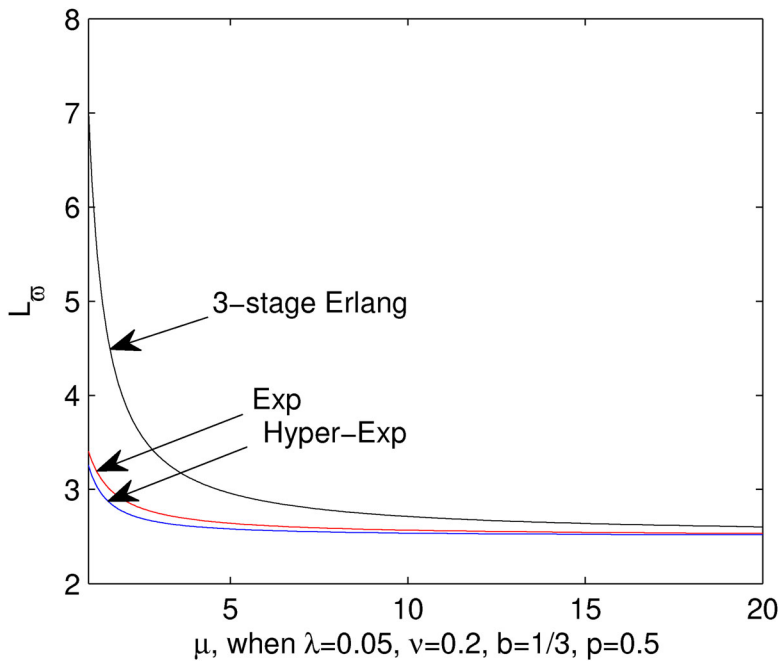


Figure 17. Impact of μ on L_m for different service time distribution.

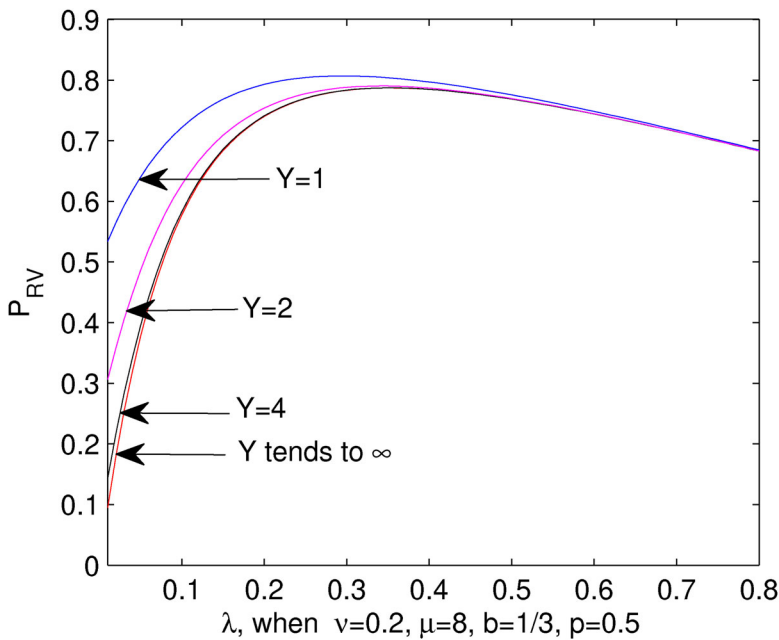


Figure 18. Impact of λ on P_{RV} for different Y values.

Example 4. Presuming the sleep time and service time as exponential distribution, while batch size and Y follows a geometric distribution and model parameters as $\mu=8$, $\nu=0.2$, $b=\frac{1}{3}$ and $p=0.5$, we display the impact of packet arrival rate λ on P_{RV} , P_D , L_m and $E(W)$ in Figures 18–21. In the following subsections, we summarize some investigations:

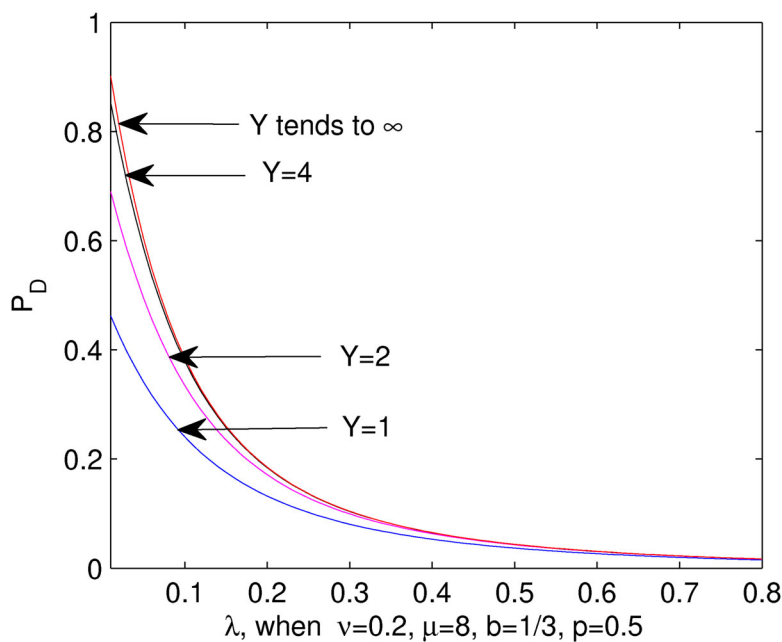


Figure 19. Impact of λ on P_D for different Y values.

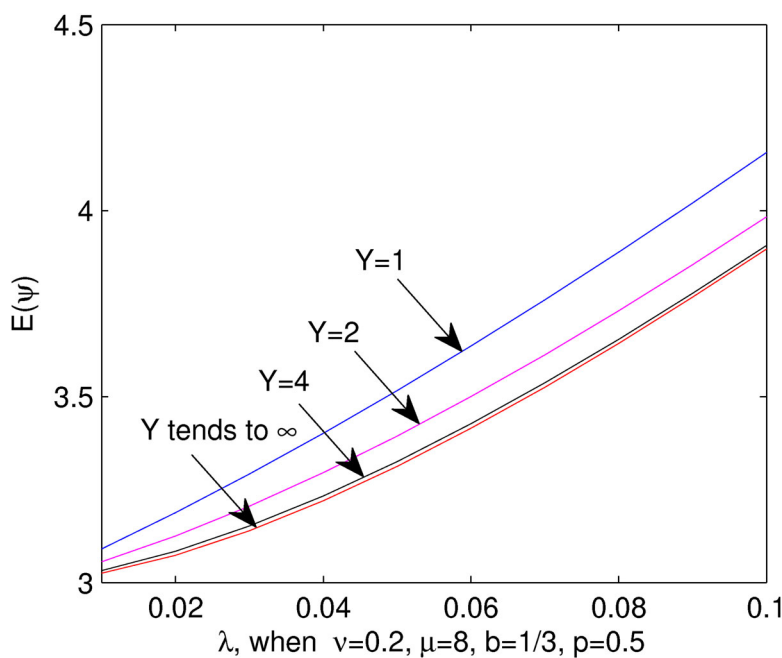


Figure 20. Impact of λ on $E(\psi)$ for different Y values.

Impact of packet arrival rate λ on probability that the server is in sleep period P_{RV} for different values of maximum number of sleep Y

From Figure 18, we observe that

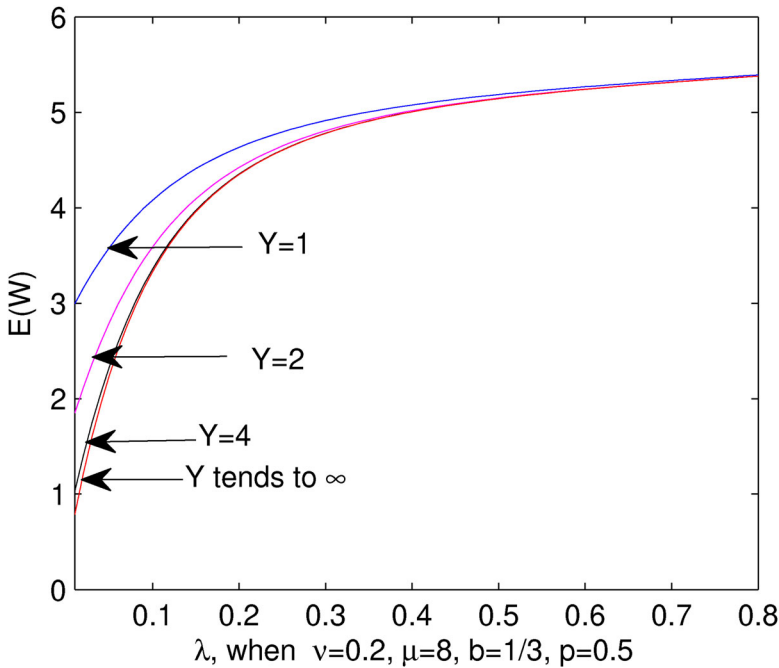


Figure 21. Impact of λ on $E(W)$ for different Y values.

1. The system state probability P_{RV} increases sharply in the beginning while arrival rate λ increases and then decreases slowly for different values of Y . A possible explanation for this situation is as follows: for any value of Y , while the arrival rate λ is comparatively small then it is obvious that the probability that the server goes for sleep period P_{RV} increases rapidly whereas it decreases slowly while the arrival rate λ becomes high.
2. The probability that the server goes for sleep period P_{RV} decreases with the increasing value of Y , this is because as the maximum number of sleep Y increases the probability of that Y^{th} sleep decreases (as Y follows a geometric distribution), which leads to decrease in probability of the server is in sleep period P_{RV} .
3. For a large value of λ , the difference between the values of P_{RV} is small as Y increases.

Impact of packet arrival rate λ on probability that the server is in warm-up period P_D for different maximum number of sleep Y

From Figure 19, we see that

1. For a fixed value Y , the probability that the server is in warm-up period P_D decreases in the increasing value of λ . This is because as the packet arrival rate λ increases the probability of the server goes for awake period increases which decreases the probability that the server is in warm-up period.
2. For a fixed value of λ , the probability that server goes for warm-up P_D increases in the increasing value of Y . This is because as the maximum number of sleeps

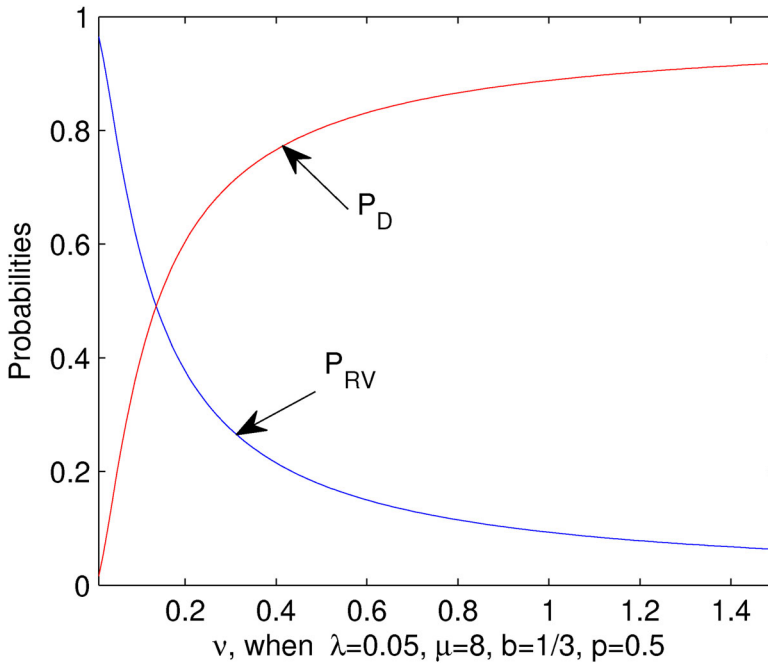


Figure 22. Impact of ν on P_{RV} and P_D .

Y increases, the probability of that Y^{th} sleep decreases (as Y follows geometric distribution), which leads to increase in probability that the server goes for warm-up period.

The figures are useful for finding the optimal value of Y . It displays that the values of P_{RV} and P_D for $Y=4$ and Y tends to ∞ are very close, therefore, it may be rewarding if we select the value of Y between 4 and ∞ .

Impact of packet arrival rate λ on mean number $E(\psi)$ of batches served during the awake period and mean waiting time $E(W)$ for different values of maximum number of sleep Y

From Figures 20 and 21, we observe that

1. For a fixed value Y , $E(\psi)$ and $E(W)$ both increase for increasing λ as it should be. This means as the packet arrival rate increases but service rate does not change (i.e., only arrival rate increases), then it is obvious that the mean waiting time of packets and mean number of batches served during awake period in system increases.
2. For a fixed value of λ , $E(\psi)$ and $E(W)$ both decrease with an increasing value of Y . The reason is that as maximum number of sleeps Y increases, the probability of that Y^{th} sleep decreases (as Y follows a geometric distribution), i.e., the probability that the server takes consecutive vacations decreases and the probability that the customers are served through a service rate μ increases, which gives rise to decrease in $E(\psi)$ and $E(W)$.

Example 5. Considering the model parameters as $\lambda = 0.05$, $\mu = 8$, $b = \frac{1}{3}$, $p = 0.5$, we displays the impact of sleep rate ν on P_{RV} and P_D in Figure 22. Here, we presume that the service time, sleep time, batch size and Y follows the similar distribution as in Figures 18–21.

Impact of sleep rate ν on probability that the server is in sleep period P_{RV} and probability that the server is in warm-up period P_D

Figure 22 reveals that

1. As the values of sleep rate ν increases, the probability that the server is sleep period P_{RV} decreases and the probability that the server is in warm-up period P_D increases. This is in view of the fact that while sleep rate increases the average sleep time decreases and a probability of packets in the arriving batches served in the system increases, which gives rise to the decrease in system state probability P_{RV} and increase in the system state probability P_D .
2. Two curves are intersected at one point and this intersection point is the optimum point which gives us the value of ν at which probability P_{RV} is minimized and probability P_D is maximized.

Thus, as shown in the analysis, we can conclude that the probability associated with sleep (vacation) has a significant role. Also, batch arrival plays an important role in system performance measures and cannot be disregarded when we analyze a continuous time system because by presuming batch arrival in the system, it turns into a very complex system. Thus, the system designers must carefully observe how customers arrive in the system and what distribution it follows in designing their particular system.

9. Conclusion and future work

In this paper, we have analyzed an $M^X/G/1$ queueing system with random vacation policy. To the best of the knowledge, MRGP approach is applied first time to study the steady state system size probabilities for the $M^X/G/1$ system with random vacation policy. It has been believed that the proposed model may be of great interest in the design and operation of various digital systems such as telecommunication systems, manufacturing systems, production or inventory systems.

Managerial implication: The proposed model also discusses the embedded Markov chain technique to obtain steady state queue size distribution at service completion point and some other system performance measures that can provide the information about the efficiency of the model so that it can be used in an IEEE802.16e standard. As shown in the numerical analysis with different example, we can say that-

- As the sleep rate (i.e., the rate at which the MS takes sleep) increases the service rate $\mu = 8$, arrival rate $\lambda = 0.05$, batch size parameter $b = 1/3$, set-up cost (C_{su}) = 10, holding cost (C_{hl}) = 2, sleep period cost (C_{RV}) = 7, warm-up period cost (C_D) = 3, service cost (C_S) = 50 and service time follows exponential, hyper-exponential and 3-stage Erlang distribution, expected cost per unit time decreases and for $\lambda = 0.05, \mu = 0.1$, $C_{su} = 10$, $C_{hl} = 2$, $C_{RV} = 0.5$, $C_D = 300$, $C_S = 5000$ as sleep rate increases expected cost per unit time increases.

By comparing the Figures 4–9, we can conclude that if the service rate during the awake period is comparatively higher i.e., average service time is lesser, also if the cost for sleep period is higher and cost for warm-up period and awake period is lesser then the expected cost per unit time is comparatively less. So, the system engineer should consider the service rate as higher. Set-up cost, holding cost, arrival rate does not effect so much.

- while the packet arrival rate λ is comparatively smaller the probability of the server goes for sleep period P_{RV} increases and probability of the server goes for warm-up period P_D decreases whereas as the arrival rate λ becomes larger the probability of the server goes for sleep period P_{RV} decreases and probability of the server goes for warm-up period P_D increases and probability of going for awake period increases.
- If we consider the service time as hyper-exponential distribution, the probability that the server goes for sleep period increases as compared to exponential and 3-stage Erlang distribution. Therefore, when the server provides service through hyper-exponential distribution, then the service time variation will be greater and consequently average service time of packets will be less and the probability that server goes for sleep period will be increase.
- If the packet arrival rate λ increases but other model parameter does not change, for any maximum number of sleep Y the mean waiting time and mean unfinished work increases as it is obvious. On the other hand, as the maximum number of sleep Y increases the probability of that Y^{th} sleep decreases (as Y follows geometric distribution) then the probability that the packets in an arriving batch served through a service rate increases which decreases mean waiting time and mean unfinished work.

Thus we can conclude that probability associated with sleep i.e., considering maximum number of sleep as a random variable plays a significant role in the analysis. Also different service time distribution, different cost parameter, system parameter plays an important role in the analysis.

We hope this investigation will definitely help the system engineer and system developer to modify or build their system. They should carefully observe how system behaves as the parameter changes.

It would be interesting to investigate some other queueing systems, e.g., the proposed model can be extended to gated or limited service discipline or the corresponding batch service queue under random vacation policy.

Acknowledgement

Authors are thankful to the editor and two anonymous reviewers for their valuable suggestions and comments which helped improve the manuscript to great extent and gratefully acknowledges for the financial support received from the Department of Telecommunications (DoT), India.

ORCID

Selvamuthu Dharmaraja  <http://orcid.org/0000-0003-2892-0864>

References

- Abolnikov, L., and A. Dukhovny. 1991. Markov chains with transition delta matrix: Ergodicity condition, invariant probability measures and applications. *Journal of Applied Mathematics and Stochastic Analysis* 4 (4):333–55. doi:[10.1155/S1048953391000254](https://doi.org/10.1155/S1048953391000254).
- Alouf, S., E. Altman, and A. P. Azad. 2008. Analysis of an $M/G/1$ queue with repeated inhomogeneous vacations with application to IEEE 802.16e power saving mechanism. Paper presented at the Proc. QEST, Saint-Malo, France, September 14–17, 27–36.
- Azad, A. P., S. Alouf, E. Altman, V. Borker, and G. Paschos. 2009. Vacation policy optimization with application to IEEE 802.16e power saving mechanism. Paper presented at the Wireless Days (WD), 2nd IFIP, Paris, France, December 15–17, 1–7.
- Azad, A. P., S. Alouf, E. Altman, V. Borkar, and G. S. Paschos. 2011. Optimal control of sleep periods for wireless terminals. *IEEE Journal on Selected Areas in Communications* 29 (8): 1605–17. doi:[10.1109/JSAC.2011.110910](https://doi.org/10.1109/JSAC.2011.110910).
- Baba, Y. 1996. On the $M^X/G/1$ queue with vacation times. *Operations Research Letters* 5:93–8. doi:[10.1016/0167-6377\(86\)90110-0](https://doi.org/10.1016/0167-6377(86)90110-0).
- Bratiichuk, N. S. 2000. Exact formulas for $E^G/G/1/N$ – type queuing systems. *Ukrainian Mathematical Journal* 5:1182–94.
- Chaudhry, M. L., and J. G. C. Templeton. 1983. *A first course in bulk queues*. New York: John Wiley and Sons.
- Choudhury, G. 2002a. A batch arrival queue with a vacation time under single vacation policy. *Computers & Operations Research* 29:1941–55. doi:[10.1016/S0305-0548\(01\)00059-4](https://doi.org/10.1016/S0305-0548(01)00059-4).
- Choudhury, G. 2002b. Analysis of the $M^X/G/1$ queueing system with vacation times. *Sankhya: The Indian Journal of Statistics* 64:37–49.
- Cooper, R. B. 1981. *Introduction to queueing theory*. Amsterdam: Elsevier. doi:[10.1145/800175.809851](https://doi.org/10.1145/800175.809851).
- Doshi, B. T. 1986. Queueing systems with vacations a survey. *Queueing Systems* 1 (1):29–66. doi:[10.1007/BF01149327](https://doi.org/10.1007/BF01149327).
- Doshi, B. T. 1990. Single sever queues with vacations. In *Stochastic analysis of computer and communication systems*, ed. H. Takagi, 217–65. Amsterdam: North Holland.
- Fuhrmann, S. W., and R. B. Cooper. 1985. Stochastic decomposition in the $M/G/1$ queue with generalized Vacations. *Operations Research*. 33:1117–29. doi:[10.1287/opre.33.5.1117](https://doi.org/10.1287/opre.33.5.1117).
- Heyman, D. P. 1968. Optimal operating policies for $M/G/1$ queueing system. *Operations Research* 16:362–82. doi:[10.1287/opre.16.2.362](https://doi.org/10.1287/opre.16.2.362).
- Jeyakumar, S., and B. Senthilnathan. 2017. Modelling and analysis of a bulk service queueing model with multiple working vacations and server breakdown. *RAIRO - Operations Research* 51 (2):485–508. doi:[10.1051/ro/2016037](https://doi.org/10.1051/ro/2016037).
- Kella, O. 1989. The threshold policy in the $M/G/1$ queue with server vacations. *Naval Research Logistics* 36 (1):111–23. doi:[10.1002/1520-6750\(198902\)36:1<111::AID-NAV3220360109>3.0.CO;2-3](https://doi.org/10.1002/1520-6750(198902)36:1<111::AID-NAV3220360109>3.0.CO;2-3).
- Kulkarni, V. G. 1995. *Modelling and analysis of stochastic systems*. London: Chapman & Hall.
- Ke, J. C. 2007. Operating characteristic analysis on the $M^X/G/1$ system with a variant vacation policy and balking. *Applied Mathematical Modelling* 31:1321–37. doi:[10.1016/j.apm.2006.02.012](https://doi.org/10.1016/j.apm.2006.02.012).
- Ke, J. C., F. M. Chang, and T. H. Liu. 2019. $M/M/c$ balking retrial queue with vacation. *Quality Technology and Quantitative Management* 16:1–12.
- Ke, J. C., and Y. K. Chu. 2006. A modified vacation model $M^X/G/1$ system. *Applied Stochastic Models in Business and Industry* 22:1–16. doi:[10.1002/asmb.607](https://doi.org/10.1002/asmb.607).
- Ke, J. C., C. H. Wu, and Z. G. Zhang. 2010. Recent developments in vacation queueing models: A short survey. *International Journal of Operations Research* 7:3–8.
- Ke, J. C., K. B. Huang, and W. L. Pearn. 2010. The randomized vacation policy for a batch arrival queue. *Applied Mathematical Modelling* 34 (6):1524–38. doi:[10.1016/j.apm.2009.09.007](https://doi.org/10.1016/j.apm.2009.09.007).
- Ke, J.-C., H.-I. Huang, and Y.-K. Chu. 2010. Batch arrival queue with N policy and atmost J vacations. *Applied Mathematical Modelling* 34 (2):451–66. doi:[10.1016/j.apm.2009.06.003](https://doi.org/10.1016/j.apm.2009.06.003).

- Kempa, W. M. 2011. Analysis of departure process in batch arrival queue with multiple vacations and exhaustive service. *Communications in Statistics - Theory and Methods* 40 (16):2856–65. doi:[10.1080/03610926.2011.562767](https://doi.org/10.1080/03610926.2011.562767).
- Kempa, W. M. 2014. On transient queue-size distribution in the batch arrivals system with a single vacation policy. *Kybernetika* 50:126–41.
- Levy, Y., and U. Yechiali. 1975. Utilization of idle time in an $M/G/1$ queue with server vacations. *Management Science* 22:202–11. doi:[10.1287/mnsc.22.2.202](https://doi.org/10.1287/mnsc.22.2.202).
- Lee, H. S., and M. M. Srinivasan. 1989. Control policies for the queueing system. *Management Science* 35 (6):708–21. doi:[10.1287/mnsc.35.6.708](https://doi.org/10.1287/mnsc.35.6.708).
- Lee, H. W., S. S. Lee, J. O. Park, and K. C. Chae. 1994. Analysis of $M^X/G/1$ queue with N policy and multiple vacations. *Journal of Applied Probability* 31:467–96.
- Lee, S. S., H. W. Lee, S. H. Yoon, and K. C. Chae. 1995. Batch arrival queue with N policy and single vacation. *Computers & Operations Research* 22:173–89. doi:[10.1016/0305-0548\(94\)E0015-Y](https://doi.org/10.1016/0305-0548(94)E0015-Y).
- Leung, K. K. 1992. On the additional delay in an $M/G/1$ queue with generalized vacations and exhaustive service. *Operations Research* 40:272–83.
- Park, Y., and G. U. Hwang. 2007. Performance modeling and analysis of the sleep-mode in IEEE 802.16e WMAN. Paper presented at the Proc. IEEE Vehicular Technology Conference (VTC), Dublin, Ireland, April 22–25, 2801–6.
- Ross, S. M. 1983. *Stochastic processes*. New York: Wiley.
- Seo, J. B., S. Q. Lee, N. H. Park, H. W. Lee, and C. H. Cho. 2004. Performance analysis of sleep mode operation in IEEE802.16e. *Vehicular Technology Conference* 2:1169–73.
- Takacs, L. 1962. *Introduction to the theory of queues*. New York: Oxford University Press.
- Takagi, H. 1991. *Queueing analysis: Vacation and priority systems*. Vol. 1. Amsterdam: North Holland.
- Tadj, L., and G. Choudhury. 2005. Optimal design and control of queues. *TOP* 13 (2):359–412. doi:[10.1007/BF02579061](https://doi.org/10.1007/BF02579061).
- Tian, N., and Z. G. Zhang. 2006. *Vacation queueing models: Theory and applications*. New York: Springer.
- Yang, D. Y., M. F. Chang, and J. C. Ke. 2016. On an unreliable retrial queue with general repeated attempts and J optional vacations. *Applied Mathematical Modelling* 40 (4):3275–3288.
- Yang, B., Z. Hou, and J. Wu. 2018. Analysis of the equilibrium strategies in the $Geo/Geo/1$ queue with multiple working vacations. *Quality Technology and Quantitative Management* 15 (6): 663–85.
- Zhang, Z. G., R. G. Vickson, and M. J.A. van Eenige. 1997. Optimal two threshold policies in an $M/G/1$ queue with two vacati types. *Performance Evaluation* 29 (1):63–80. doi:[10.1016/S0166-5316\(96\)00005-3](https://doi.org/10.1016/S0166-5316(96)00005-3).

Appendix

Transition from random vacation period to dormant period: We presume that the general transition of server random vacation, i.e., τ_1 is distributed with CDF $F_\theta(\cdot)$. Server random vacation period (which consists of maximum Y number of random vacations) (i.e., $s = 0$) is enabled if and only if no batch is waiting in a queue, i.e., the system is empty. After completing a random vacation period if the server stays empty, it goes to a dormant period (i.e., $s = 0, s' = 1$). Hence, we have

$$K_{(0,0):(0,1)}(t) = \int_0^t P_{01}(0; 0, x) dF_\theta(x) \quad (A1)$$

Transition from random vacation period to service period: Here, we consider server random vacation as the transition between two sequential regeneration epochs, i.e., τ_1 is distributed with CDF $F_\theta(\cdot)$. As soon as the system becomes empty, the server goes to a random vacation period (i.e., $s = 0$). If any batch arrives during the random vacation period, the server directly goes to an awake period or service period (i.e., $s = 0, s' = 2$). Hence, we have

$$K_{(h,0);(n,2)}(t) = \int_0^t P_{02}(h; n, x) dF_\emptyset(x), n = h = 1, 2, \dots \quad (\text{A2})$$

Transition from dormant period to service period: Here, we presume the case of the server's dormant period, i.e., after completion of a random vacation period if the server finds no batch holding in a queue, it directly goes to a dormant period and waits for the first batch arrival. If any batch arrival occurs during the dormant period, it immediately goes to an awake period or service period (i.e., $s = 1, s' = 2$). Thus, we have

$$K_{(h,1);(n,2)}(t) = P_{12}(h; n, t), n = h = 1, 2, \dots \quad (\text{A3})$$

Transition from service period to random vacation period: Here, we suppose that the batch service is the transition between two sequential regeneration epochs, i.e., τ_1 is distributed with CDF $F_S(\cdot)$. When the server completes the service of all customers in the arriving batches and no batch is waiting in the queue, i.e., the leaving batch is the final batch (i.e., $s = 2, h = 1$), then the server goes to a random vacation period (i.e., $s' = 0, n = 0$). Therefore, we have

$$K_{(1,2);(0,0)}(t) = \int_0^t P_{20}(1; 0, x) dF_S(x) \quad (\text{A4})$$

Transition from service period to service period: Here, we presume the general transition of the system's service period, i.e., τ_1 is distributed with CDF $F_S(\cdot)$. Clearly, service period is available only when at least one batch is waiting in a queue. If the service termination of each batch does not empty the queue, the system stays in the service period (i.e., $s = 2, s' = 2$). Hence, we have

$$K_{(h,2);(n,2)}(t) = \int_0^t P_{22}(h; n, x) dF_S(x), n = 1, 2, \dots, h = 1, 2, \dots \quad (\text{A5})$$