RAIRO-Oper. Res. 53 (2019) 415–443
https://doi.org/10.1051/ro/2017029

RAIRO Operations Research
www.rairo-ro.org

ANALYSIS OF M/G/1 RETRIAL QUEUES WITH SECOND OPTIONAL SERVICE AND CUSTOMER BALKING UNDER TWO TYPES OF BERNOULLI VACATION SCHEDULE

S. Pavai Madheswari[1], B. Krishna Kumar[2] and P. Suganthi[1,*]

Abstract. An *M/G/1* retrial queueing system with two phases of service of which the second phase is optional and the server operating under Bernoulli vacation schedule is investigated. Further, the customer is allowed to balk upon arrival if he finds the server unavailable to serve his request immediately. The joint generating functions of orbit size and server status are derived using supplementary variable technique. Some important performance measures like the orbit size, the system size, the server utilisation and the probability that the system is empty are found. Stochastic decomposition law is established when there is no balking permitted. Some existing results are derived as special cases of our model under study. Interestingly, these performance measures are compared for various vacation schedules namely exhaustive service, 1-limited service, Bernoulli vacation and modified Bernoulli vacation schedules. Extensive numerical analysis is carried out to exhibit the effect of the system parameters on the performance measures.

Mathematics Subject Classification. 60K25, 90B22, 68M20.

Received October 6, 2016. Accepted April 21, 2017.

1. Introduction

In view of the network complexity and increasing number of customers, the customer behaviour and the retrial phenomenon may have a significant impact on the computer network performance. The special feature of the retrial queue is that an arriving customer who finds the server busy upon arrival may join the virtual group of blocked customers, called orbit, and retry for service after a random amount of time. Each blocked customer generates a stream of repeated requests independently of the rest of the customers in the retrial group. In the classical retrial policy, the intervals between successive repeated attempts are exponentially distributed with intensity $n\theta$, where θ is a rate of retrial and $n(n > 0)$ is number of customers in the orbit (see Yang and Templeton [37] and Falin [13]). However, there are certain situations in which the intervals between the successive retrials from the orbit are independent of the number of customers in it. For example when only the customer that is at the head of the orbit is allowed to conduct retrials or, more realistically, when the

Keywords. Retrial queue, two phase service, balking, Bernoulli vacation, modified Bernoulli vacation, supplementary variable, stochastic decomposition.

[1] Department of Mathematics, R.M.K. Engineering College, Chennai, India.

[2] Department of Mathematics, Anna University, Chennai, India.

*Corresponding author: psi.sh@rmkec.ac.in

Article published by EDP Sciences

server looks for customers from the orbit. This situation occurs in a variety of contexts: customers that do not find an idle server at a call center may leave their contact details and wait to be called back later; a person that answers emails for a commercial website begins immediately to answer a newly arrived email if he is idle, otherwise he needs some time to retrieve a waiting email from his list. This retrial policy is known as the constant retrial policy. The constant retrial policy was introduced by Fayolle [14] in a Markovian framework where both service and retrial (seeking) times are exponential. Choi *et al.* [8] considered this model where seeking times are generally distributed, and Martin and Artalejo [30] considered the model where the service times are generally distributed. Finally, Gomez−Corral [17] considered the single server retrial queue with the general retrial policy and general service and seeking times. Retrial queues are widely and successfully used as mathematical models of computer network systems, telephone switching systems and wireless network systems. We refer the readers to the books of Falin and Templeton [12] and Artalejo and Gomez−Corral [6] for a detailed study on the fundamental concepts of retrial queues. Recent bibliographies on retrial queues can be found in Artalejo ([1,2]). Artalejo and Falin [3] have made a comparative analysis on standard and retrial queueing systems.

In the past two decades a notable amount of research has been carried out in the queueing systems in which the server provides to each customer two phases of service in succession. The motivation for these types of models comes from some computer networks and telecommunication systems, where messages are processed in two stages by a single server of which the second stage may be optional. These queueing systems are characterized by the feature that all arrivals demand the first phase of service called essential the service and the second phase of service is optional which is also provided by the same server. An $M/G/1$ retrial queue with feedback and starting failure was studied by Krishna Kumar *et al.* [25]. Krishna Kumar *et al.* [24] have discussed the busy period of an $M/G/1$ retrial queue with two phases of service. Madan [28], Medhi [31] and Wang [33] have studied this kind of queueing models, among others.

Queueing systems with server vacations have been studied extensively in the past. This type of queueing models occur in many real life situations where the server is used for other secondary jobs. Applications arise naturally in call centers with multitask employees, maintenance activities *etc.* Comprehensive surveys on vacation queueing models can be found in Doshi [10], Takagi [32] and Ke *et al.* [19]. A wide class of policies governing the vacation mechanism have been discussed in the literature. Most of the analysis for retrial queues concerns the single vacation with exhaustive service schedule (Artalejo [4]). Keilson and Servi [21] introduced Bernoulli vacation schedule: If the queue is empty after a service completion, the server goes for vacation. On the otherhand if the queue is not empty then service begins with a specific probability $1 - a$ or a vacation period begins with probability a. At the end of a vacation period service begins if a customer is present in the queue. Otherwise, the server waits for the first customer to arrive. Several papers have recently appeared in the queueing literature in which the concept of general retrial times has been considered along with Bernoulli vacation schedule. However, in the retrial queueing systems, the server may not be aware of the status of the customers in the system as in the case of Keilson and Servi's model (see Krishna Kumar and Arivudainambi [25], Krishna Kumar and Pavai Madheswari [23], Wenhui [36] and Wang and Li [34]. Hence, it is necessary to modifiy the Bernoulli schedule introduced by them in the retrial context. It is reasonable to assume that even if there are no customers in the system, the server will wait for next customer to arrive with probability $1-a$ and chooses to go on vacation with probability a. This special kind of Bernoulli vacation is called as modified Bernoulli vacation schedule in queueing literature. Wang and Li [34] have studied a similar model with breakdown and repair.

As said earlier, the increase in data traffic, network complexity and the number of customers, considering the customer behaviour is also becoming essential. Wu *et al.* [35] and Ke and Chang [20] have dealt with an $M/G/1$ queue with balking. Gilbert [16] has also discussed retrials and balks. Most of the previous studies associated to retrial queues have a common assumption that the system has a single server who provides only one kind of service. Although some aspects have been discussed separately on queueing systems with repeated attempts, Bernoulli vacation schedule and two phases of service, it is seldom found any work that combines these features where the customer may balk the system when he finds the server is busy or on vacation. Also, there is no work found which compares the various vacation policies for a single queueing system. Hence to fill up the gap, in this paper an attempt is made to analyse an $M/G/1$ retrial queueing system where the server provides two phases of service of which the second phase is optional and the customer is allowed to balk when the server is busy or on vacation under the Bernoulli vacation schedule. Interestingly, the various vacation policies, namely, single vacation with exhaustive service, with 1-limited service, Bernoulli vacation and modified Bernoulli vacation schedule are compared numerically.

Apart from the theoretical importance, our retrial queueing model under study has many potential real time applications. For instance, cloud computing has been an emerging technology for provisioning computing resources and providing infrastructure through web applications. Cloud computing is a new cost-efficient computing paradigm in which information and computer resources have been accessed from web browser by users. In case of virtualization and resource time sharing, cloud serves a large user base with different types of user needs using a single set of physical resources. User is responsible to pay only for the used resources and services by Service Level Agreement (SLA), without any knowledge of how a service provider deals with underlying server-end machines. The service provider is required to execute service requests from a user by maintaining quality of service (QoS) requirements. If a new customer does not find any free server after connecting to the cloud service, then the system automatically redirects the request towards a waiting queue. At that moment, if the waiting queue is also fully occupied by other customers, then the newly arriving customer has to retry for service after certain time period (Ch. Banerjee *et al.* [7]).

The cloud users and service provider correspond to our customers and server respectively. Here, the service provider may undertake some maintenance activity which may be considered as server vacation. When the service provider is busy or engaged in maintenance activity, the user may be virtually waiting to try again later or may quit which correspond to our orbit and balking concept. The same user also may require some additional services which can be thought of as second optional service we talk about. Our retrial queueing model under consideration will be suitable to analyse the performance of the cloud computing system.

Another interesting application is the Internet of Things (IoT), which is a novel paradigm that is rapidly gaining ground in the scenario of modern wireless telecommunications. The basic idea of this concept is the pervasive presence around us of variety of things or objects such as Radio-Frequency Identification (RFID) tags, sensors, actuators, mobile phones *etc.*, which through unique addressing schemes, are able to interact, with each other and cooperate with their neighbours to reach common goals. IoT represents the next evolution of the Internet, taking a huge ability to gather, analyze and distribute data that we can turn into information known and ultimately, wisdom in this context. Internet of things is deployed with many type of sensors, each of which is an information source, and different type of sensors capture different content and format of information. Data obtained from the sensor is real time and the sensor collects the environment information at a certain frequency and keeps updating the data. IoT creates huge amount of data which have to be stored in different locations, preferably in cloud servers or distributed data bases. The process of retrieving the data for further analysis can be modeled using retrial queues with second optional service and server vacations.

The rest of the paper is organized as follows. In Section 2, we analyse the system under Bernoulli vacation schedule. The same system under modified Bernoulli vacation is studied in Section 3. Finally in Section 4, extensive numerical analysis is carried out to exhibit the effect of the system parameters on the performance measures and the effect of different vacation policies are also showcased.

2. System under Bernoulli vacation schedule

2.1. Model description

We consider a single server retrial queueing system with two phases of service, first phase service (FPS) which is essential for all the customers and second phase service (SPS) which is optional as depicted in Figure 1. Customers arrive to the system according to a Poisson process with rate λ. When the arriving customer finds the server free, its service starts immediately. On the other hand if the server is busy or on vacation, the arriving customer leaves the service area and joins the orbit with probability $1 - b$ or leaves the system permanently (balks) with probability b. The customers in the orbit are persistent in the sense that they keep making retrials until they receive their requested service. We consider the retrial queue with FCFS orbit where the retrial

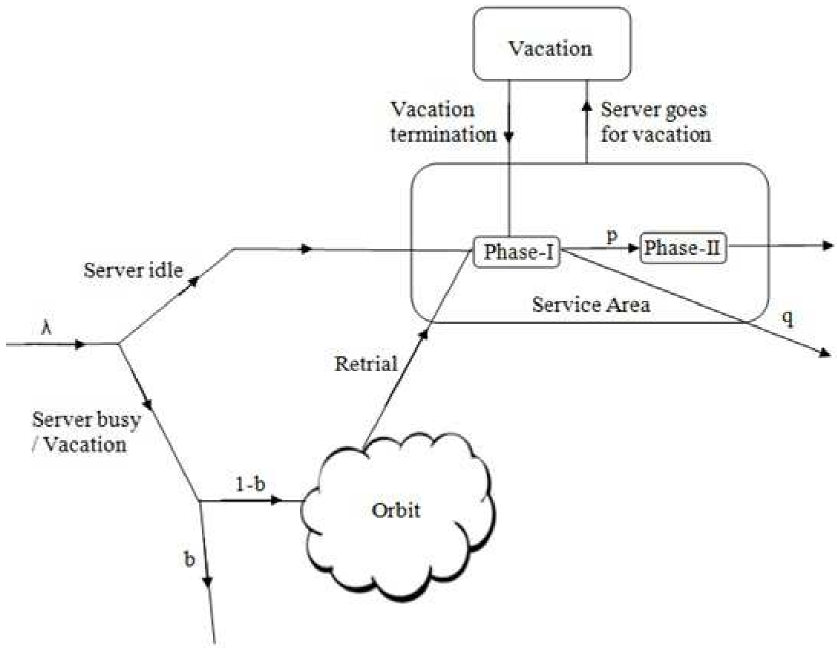


FIGURE 1. The model.

time of the customers in the orbit is generally distributed with distribution function $R(x)$ and Laplace Stieltjes transform (LST) $R^*(\theta)$. The conditional completion rate of retrial time is $\theta(x)\mathrm{d}x = \frac{dR(x)}{1-R(x)}$.

The server provides a preliminary first phase of regular service (FPS) denoted by S_1 to all arriving customers. As soon as the FPS of a customer is completed, the customer may leave the system with probability $q\ (= 1-p)$ or may be provided with a second phase of optional service (SPS) denoted by S_2 with probability $p\ (0 \leq p \leq 1)$. The service times follow general laws with probability distribution functions $S_i(x),\ i = 1, 2$, LST $S_i^*(\theta)$ and conditional completion rates $\mu_i(x)\mathrm{d}x = \frac{dS_i(x)}{1-S_i(x)},\ i = 1, 2$.

It is assumed that the server goes for vacation according to Bernoulli vacation schedule which is characterized by the feature that if the orbit is empty after a service completion then the server begins a vacation period. On the other hand, if the orbit is not empty, the server continues to serve the customer waiting in the orbit with probability $1-a$ or may go for a vacation with probability $a\ (0 \leq a \leq 1)$. The server remains in idle time, waiting for the customer to serve either from the orbit or a new primary arrival. The vacation time V follows general distribution with distribution functions $V(x)$, LST $V^*(\theta)$ and conditional completion rate $\nu(x)\mathrm{d}x = \frac{dV(x)}{1-V(x)}$.

The state of the system at time t can be described by the Markov process $\{N(t);\ t \geq 0\} = \{(C(t), X(t), \xi(t));\ t \geq 0\}$, where $C(t)$ denotes the server state $0, 1, 2$ or 3 according to the server is idle, busy with FPS, busy with SPS or on vacation and $X(t)$ corresponds to the number of customers in orbit at time t. If $C(t) = 0$ and $X(t) > 0$, then $\xi(t)$ represents elapsed retrial time, if $C(t) = 1$ (2) and $X(t) \geq 0$, then $\xi(t)$ represents elapsed service time in FPS (SPS) and if $C(t) = 3$ and $X(t) \geq 0$, then $\xi(t)$ represents elapsed vacation time.

2.2. Ergodicity condition

We first obtain the necessary and sufficient condition for the system to be stable. To this end, in the following theorem, we establish the ergodicity of the embedded Markov chain at departure/vacation completion epochs. Let $\{t_n;\ n \in N\}$ be the sequence of epochs at which either a service completion occurs or a vacation period ends. The sequence of random vectors $X_n = (C(t+), X(t+))$ forms a Markov chain, which is the embedded Markov chain for our queueing system with state space $S = \{0, 1, 2, 3\} \times \{0, 1, 2, 3, \ldots\} - (0, 0)$.

Theorem 2.1. *Let X_n be the orbit length at the time of either nth customer's departure or vacation completion epoch, $n \geq 1$. Then $\{X_n;\ n \geq 1\}$ is ergodic if and only if $\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] < R^*(\lambda)$.*

Proof. It is not difficult to see that $\{X_n;\ n \geq 1\}$ is an irreducible and aperiodic Markov chain. To prove ergodicity, we shall use Foster's criterion: An irreducible and aperiodic Markov chain is ergodic, if there exists a non-negative function $f(j),\ j \in N$ and $\epsilon > 0$ such that the mean drift $\psi_j = E[f(X_{n+1}) - f(X_n)/X_n = j]$ is finite for all $j \in N$ and $\psi_j \leq -\epsilon$ for all $j \in N$, except perhaps a finite number j.

In our case, we consider the function $f(j) = j$. Then we have

$$\psi_j = \begin{cases} \lambda(1-b)[E(S_1) + pE(S_2) + E(V)], & \text{for} \quad j = 0 \\ \lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] - R^*(\lambda), & \text{for} \quad j = 1, 2, 3, \ldots. \end{cases}$$

Clearly, the inequality $\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] < R^*(\lambda)$ is a sufficient condition for ergodicity. The same inequality is also necessary for ergodicity. As noted in Sennott *et al.* [26], we can guarantee non ergocity of the Markov chain $\{X_n;\ n \geq 1\}$, if it satisfies Kaplan's condition, namely $\psi_j < \infty$ for all $j \geq 0$ and there exists $j_0 \in N$ such that $\psi_j \geq 0$ for $j \geq j_0$. Notice that, in our case, Kaplan's condition is satisfied because $r_{ij} = 0$ for $j < i - 1$ and $i > 0$, where $R = (r_{ij})$ is the one step transition probability matrix of $\{X_n;\ n \geq 1\}$. Then, $\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] \geq R^*(\lambda)$ implies the non-ergodicity of the Markov chain. □

Remark 2.2. Since the arrival stream is a Poisson process, it can be shown from Burke's theorem (see Cooper [9], pp. 187–188) that the stady state probabilities of $\{(C(t), X(t));\ t \geq 0\}$ exist and are positive if and only if $\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] < R^*(\lambda)$.

From the mean drift $\psi_j = \lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] - R^*(\lambda)$, for $j \geq 1$, we have the reasonable conclusion that the term $\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)]$ has three components: new arrivals during the first phase service of the server $\lambda(1-b)E(S_1)$, during the second phase service of the server $\lambda(1-b)pE(S_2)$ and during vacation $\lambda(1-b)aE(V)$. Further, $R^*(\lambda)$ is the expected number of orbiting customers who enter service successfully, given that the previous service time leaves j customers in the orbit. For stability, we require that new customers arrive during a service and vacation time more slowly than orbiting customers seeking service, at the commencement of service. That is, $\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] < R^*(\lambda)$.

2.3. Steady state distribution

In this section, we study the stationary distribution for the system under consideration. For the Markov process $\{N(t);\ t \geq 0\}$, we define the probability $P_0(t) = P\{C(t) = 0, X(t) = 0\}$ and the probability densities

$$P_n(x,t)\mathrm{d}x = P\{C(t) = 0, \quad X(t) = n, \quad x \leq \xi(t) < x + \mathrm{d}x\}, \quad \text{for} \quad t \geq 0,\ x \geq 0 \quad \text{and} \quad n \geq 1,$$

$$Q_{i,n}(x,t)\mathrm{d}x = P\{C(t) = i, \quad X(t) = n, \quad x \leq \xi(t) < x + \mathrm{d}x\}, \quad \text{for} \quad t \geq 0, \quad x \geq 0,\ n \geq 0 \quad \text{and} \quad i = 1, 2,$$

and

$$V_n(x,t)\mathrm{d}x = P\{C(t) = 3, \quad X(t) = n, \quad x \leq \xi(t) < x + \mathrm{d}x\}, \quad \text{for } t \geq 0,\ x \geq 0 \text{ and } n \geq 0,$$

We assume that the condition $\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] < R^*(\lambda)$ is fulfilled, so that the limiting probability $P_0 = \lim_{t\to\infty} P_0(t)$, and the limiting densities $P_n(x) = \lim_{t\to\infty} P_n(x,t)$, for $x \geq 0$ and $n \geq 1$, $Q_{i,n}(x) = \lim_{t\to\infty} Q_{i,n}(x,t)$, for $x \geq 0$, $i = 1, 2$, $n \geq 0$ and $V_n(x) = \lim_{t\to\infty} V_n(x,t)$ for $x \geq 0, n \geq 0$ exist.

Using supplementary variable technique, we obtain the system of equations that govern the dynamics of the system behaviour under steady state as:

$$\lambda P_0 = \int_0^\infty V_0(x)\nu(x)\mathrm{d}x, \tag{2.1}$$

$$\frac{\mathrm{d}P_n(x)}{\mathrm{d}x} + (\lambda + \theta(x))P_n(x) = 0, \quad n \geq 1, \tag{2.2}$$

$$\frac{\mathrm{d}Q_{i,0}(x)}{\mathrm{d}x} + (\lambda(1-b) + \mu_i(x))Q_{i,0}(x) = 0, \quad i = 1, 2, \tag{2.3}$$

$$\frac{\mathrm{d}Q_{i,n}(x)}{\mathrm{d}x} + (\lambda(1-b) + \mu_i(x))Q_{i,n}(x) = \lambda(1-b)Q_{i,n-1}(x), \quad i = 1, 2, \quad n \geq 1, \tag{2.4}$$

$$\frac{\mathrm{d}V_0(x)}{\mathrm{d}x} + (\lambda(1-b) + \nu(x))V_0(x) = 0, \tag{2.5}$$

$$\frac{\mathrm{d}V_n(x)}{\mathrm{d}x} + (\lambda(1-b) + \nu(x))V_n(x) = \lambda(1-b)V_{n-1}(x), \quad n \geq 1. \tag{2.6}$$

The steady state boundary conditions are

$$\begin{aligned} P_n(0) = &\int_0^\infty V_n(x)\nu(x)\mathrm{d}x + (1-a)q\int_0^\infty Q_{1,n}(x)\mu_1(x)\mathrm{d}x \\ &+ (1-a)\int_0^\infty Q_{2,n}(x)\mu_2(x)\mathrm{d}x, \quad n \geq 1, \end{aligned} \tag{2.7}$$

$$Q_{1,0}(0) = \int_0^\infty P_1(x)\theta(x)\mathrm{d}x + \lambda P_0, \tag{2.8}$$

$$Q_{1,n}(0) = \int_0^\infty P_{n+1}(x)\theta(x)\mathrm{d}x + \lambda\int_0^\infty P_n(x)\mathrm{d}x, \ n \geq 1, \tag{2.9}$$

$$Q_{2,n}(0) = p\int_0^\infty Q_{1,n}(x)\mu_1(x)\mathrm{d}x, \ n \geq 0, \tag{2.10}$$

$$V_0(0) = q\int_0^\infty Q_{1,0}(x)\mu_1(x)\mathrm{d}x + \int_0^\infty Q_{2,0}(x)\mu_2(x)\mathrm{d}x, \tag{2.11}$$

$$V_n(0) = aq\int_0^\infty Q_{1,n}(x)\mu_1(x)\mathrm{d}x + a\int_0^\infty Q_{2,n}(x)\mu_2(x)\mathrm{d}x, \quad n \geq 1. \tag{2.12}$$

The normalising condition is

$$P_0 + \sum_{n=1}^{\infty}\int_0^\infty P_n(x)\mathrm{d}x + \sum_{n=0}^{\infty}\sum_{i=1}^{2}\int_0^\infty Q_{i,n}(x)\mathrm{d}x + \sum_{n=0}^{\infty}\int_0^\infty V_n(x)\mathrm{d}x = 1. \tag{2.13}$$

We define the following probability generating functions for solving the equations (2.1)−(2.12):

$$P(x,z)=\sum_{n=1}^{\infty}P_n(x)z^n;\quad Q_i(x,z)=\sum_{n=0}^{\infty}Q_{i,n}(x)z^n; i=1,2\quad\text{and}\quad V(x,z)=\sum_{n=0}^{\infty}V_n(x)z^n.$$

The joint steady state distribution of the system under different server states are discussed in the following theorem.

Theorem 2.3. *If* $\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]<R^*(\lambda)$, *then the steady state distributions of* $\{N(t);\ t\geq 0\}$ *are obtained as*

$$P(x,z)=\frac{\begin{array}{r}\lambda P_0 z\Big\{V^*\big(\lambda(1-b)\big)\Big[1-S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)\Big]+(1-a)\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\Big\}\\ \times e^{-\lambda x}\Big(1-R(x)\Big)\end{array}}{\begin{array}{r}V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{array}},\tag{2.14}$$

$$Q_1(x,z)=\frac{\begin{array}{r}\lambda P_0\Big\{(1-a)\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ +(1-z)R^*(\lambda)V^*\big(\lambda(1-b)\big)\Big\}\times e^{-\lambda(1-b)(1-z)x}\Big(1-S_1(x)\Big)\end{array}}{\begin{array}{r}V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{array}},\tag{2.15}$$

$$Q_2(x,z)=\frac{\begin{array}{r}p\lambda P_0\Big\{(1-a)\Big(1-V^*(\lambda(1-b)(1-z))\Big)\big[z+(1-z)R^*(\lambda)\big]\\ +(1-z)R^*(\lambda)V^*\big(\lambda(1-b)\big)\Big\}S_1^*\big(\lambda(1-b)(1-z)\big)\times e^{-\lambda(1-b)(1-z)x}\Big(1-S_2(x)\Big)\end{array}}{\begin{array}{r}V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{array}},\tag{2.16}$$

and

$$V(x,z)=\frac{\begin{array}{r}\lambda P_0\Big\{\Big\{\big[z+(1-z)R^*(\lambda)\big](1-a)+a(1-z)R^*(\lambda)V^*\big(\lambda(1-b)\big)\Big\}S_1^*\big(\lambda(1-b)(1-z)\big)\\ \times\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)-(1-a)z\Big\}\times e^{-\lambda(1-b)(1-z)x}\Big(1-V(x)\Big)\end{array}}{\begin{array}{r}V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{array}}.\tag{2.17}$$

where

$$S_i^*(\lambda(1-b)(1-z))=\int_0^{\infty}e^{-\lambda(1-b)(1-z)x}\mu_i(x)[1-S_i(x)]\mathrm{d}x,\quad i=1,2,$$
$$V^*(\lambda(1-b)(1-z))=\int_0^{\infty}e^{-\lambda(1-b)(1-z)x}\nu(x)[1-V(x)]\mathrm{d}x,$$
$$R^*(\lambda)=\int_0^{\infty}e^{-\lambda x}\theta(x)[1-R(x)]\mathrm{d}x.$$

Proof. Multiplying (2.2)−(2.6) by z^n and summing over all n, we get

$$\frac{\partial P(x,z)}{\partial x} + [\lambda + \theta(x))]P(x,z) = 0, \tag{2.18}$$

$$\frac{\partial Q_i(x,z)}{\partial x} + [\lambda(1-b)(1-z) + \mu_i(x)]Q_i(x,z) = 0,\ i = 1,2, \tag{2.19}$$

and

$$\frac{\partial V(x,z)}{\partial x} + [\lambda(1-b)(1-z) + \nu(x)]V(x,z) = 0. \tag{2.20}$$

From (2.7)−(2.12), we obtain

$$P(0,z) = \int^{\infty} V(x,z)\nu(x)\mathrm{d}x + (1-a)q\int_0^{\infty} Q_1(x,z)\mu_1(x)\mathrm{d}x + (1-a)\int_0^{\infty} Q_2(x,z)\mu_2(x)\mathrm{d}x - \lambda P_0, \tag{2.21}$$

$$Q_1(0,z) = \lambda P_0 + \lambda\int_0^{\infty} P(x,z)\mathrm{d}x + \frac{1}{z}\int_0^{\infty} P(x,z)\theta(x)\mathrm{d}x, \tag{2.22}$$

$$Q_2(0,z) = p\int_0^{\infty} Q_1(x,z)\mu_1(x)\mathrm{d}x, \tag{2.23}$$

$$V(0,z) = aq\int_0^{\infty} Q_1(x,z)\mu_1(x)\mathrm{d}x + a\int_0^{\infty} Q_2(x,z)\mu_2(x)\mathrm{d}x. \tag{2.24}$$

Solving (2.18)−(2.20), we get

$$P(x,z) = P(0,z)\mathrm{e}^{-\lambda x}[1 - R(x)], \tag{2.25}$$

$$Q_i(x,z) = Q_i(0,z)\mathrm{e}^{-\lambda(1-b)(1-z)x}[1 - S_i(x)], \tag{2.26}$$

$$V(x,z) = V(0,z)\mathrm{e}^{-\lambda(1-b)(1-z)x}[1 - V(x)]. \tag{2.27}$$

Solving (2.5), we get

$$V_0(x) = V_0(0)\mathrm{e}^{-\lambda(1-b)x - \int_0^x \nu(u)du}.$$

Using the above in (2.1), we get

$$V_0(0) = \frac{\lambda P_0}{V^*(\lambda(1-b))}. \tag{2.28}$$

After some algebraic manipulations, we obtain the required results (2.14)−(2.17). □

For the limiting probability generating functions $P(x,z),\ Q_i(x,z),\ i=1,2$ and $V(x,z)$, we define the partial probability generating functions as

$$\begin{aligned} P(z) &= \int_0^\infty P(x,z)\mathrm{d}x,\\ Q_i(z) &= \int_0^\infty Q_i(x,z)\mathrm{d}x,\ i=1,2\\ \text{and}\quad V(z) &= \int_0^\infty V(x,z)\mathrm{d}x. \end{aligned}$$

Here, $P(z)$ is the probability generating function of the orbit size when the server is idle, $Q_i(z)$ is the probability generating function of the orbit size when the server is busy serving phase i service, $i=1,2$ and $V(z)$ is the probability generating function when the server is on vacation.

Define the probability generating function of the number of customers in the system as $K(z)=P_0+P(z)+zQ_1(z)+zQ_2(z)+V(z)$ and the probability generating function of the number of customers in the orbit as $H(z)=P_0+P(z)+Q_1(z)+Q_2(z)+V(z)$, where P_0 is the probability that the server is idle in the system. The following theorem gives the main results of our model under consideration.

Theorem 2.4. *If* $\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]<R^*(\lambda)$, *then the partial probability generating functions are given as*

$$P(z)=\frac{\begin{aligned}&P_0 z\Big(1-R^*(\lambda)\Big)\Big\{V^*\big(\lambda(1-b)\big)\Big[1-S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\\ &\quad\times\Big(aV^*\big(\lambda(1-b)(1-z)\big)+(1-a)\Big)\Big]+(1-a)\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\Big\}\end{aligned}}{\begin{aligned}&V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ &\quad\times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{aligned}},\tag{2.29}$$

$$Q_1(z)=\frac{\begin{aligned}&P_0\Big(1-S_1^*\big(\lambda(1-b)(1-z)\big)\Big)\Big\{(1-a)\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ &\quad+(1-z)R^*(\lambda)V^*\big(\lambda(1-b)\big)\Big\}\end{aligned}}{\begin{aligned}&(1-b)(1-z)V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\\ &\quad\times\big[z+(1-z)R^*(\lambda)\big]\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{aligned}},\tag{2.30}$$

$$Q_2(z)=\frac{\begin{aligned}&pP_0S_1^*\big(\lambda(1-b)(1-z)\big)\Big(1-S_2^*\big(\lambda(1-b)(1-z)\big)\Big\{(1-a)\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\\ &\quad\times\big[z+(1-z)R^*(\lambda)\big]+(1-z)R^*(\lambda)V^*\big(\lambda(1-b)\big)\Big\}\end{aligned}}{\begin{aligned}&(1-b)(1-z)V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\\ &\quad\times\big[z+(1-z)R^*(\lambda)\big]\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{aligned}},\tag{2.31}$$

$$V(z)=\frac{\begin{array}{r}P_0\big[1-V^*\big(\lambda(1-b)(1-z)\big)\big]\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\\ \times\Big(\big[z+(1-z)R^*(\lambda)\big](1-a)+a(1-z)R^*(\lambda)V^*\big(\lambda(1-b)\big)\Big)-(1-a)z\Big\}\end{array}}{\begin{array}{r}(1-b)(1-z)V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\\ \times\big[z+(1-z)R^*(\lambda)\big]\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{array}}, \tag{2.32}$$

and the probability generating functions of the number of customers in the system $K(z)$ and in the orbit $H(z)$ are given as

$$K(z)=\frac{\begin{array}{r}P_0\Bigg\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\Big[\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\\ \times\Big\{(1-a)\big[z+(1-z)R^*(\lambda)\big]+aR^*(\lambda)V^*\big(\lambda(1-b)\big)\Big\}+V^*\big(\lambda(1-b)\big)R^*(\lambda)\\ \times\Big((1-b)\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big)\Big]+z\Big[bV^*\big(\lambda(1-b)\big)R^*(\lambda)\\ -b(1-a)\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\Big(1-R^*(\lambda)\Big)\Big]\Bigg\}\end{array}}{\begin{array}{r}(1-b)V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{array}}, \tag{2.33}$$

and

$$H(z)=\frac{\begin{array}{r}P_0\Bigg\{V^*\big(\lambda(1-b)\big)R^*(\lambda)\Big(1-(1-b)z\Big)+(1-a)\Big(1-V^*\big(\lambda(1-b)(1-z)\big)\Big)\Big[(1-b)z\\ \times\Big(1-R^*(\lambda)\Big)+R^*(\lambda)\Big]-bS_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)(1-z)\big)\Big)\\ \times V^*\big(\lambda(1-b)\big)R^*(\lambda)\Big(aV^*\big(\lambda(1-b)(1-z)\big)+(1-a)\Big)\Bigg\}\end{array}}{\begin{array}{r}(1-b)V^*(\lambda(1-b))\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*\big(\lambda(1-b)\big)\Big)\big[z+(1-z)R^*(\lambda)\big]\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)-z\Big\}\end{array}}, \tag{2.34}$$

where

$$P_0=\frac{V^*(\lambda(1-b))\Big\{R^*(\lambda)-\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]\Big\}}{V^*(\lambda(1-b))R^*(\lambda)\Big(1+b\lambda[E(S_1)+pE(S_2)+aE(V)]\Big)+\lambda(1-a)E(V)[1-b(1-R^*(\lambda))]}\cdot \tag{2.35}$$

Proof. Intergrating equations (2.14)−(2.17) with respect to x from 0 to ∞, we obtain the results (2.29)−(2.32). Using (2.29)−(2.32), we get, after considerable algebric manipulations, the probability generating function of the number of customers in the system $K(z)$ and that in the orbit $H(z)$ as in equation (2.33) and (2.34). Finally, the unknown probability P_0 is determined using the normalising condition $P_0+P(1)+Q_1(1)+Q_2(1)+V(1)=1$. By setting $z=1$ in $K(z)$ and applying L- Hospital's rule we get P_0 as in equation (2.35). □

2.4. Performance measures

In this section, we present some interesting performance measures of the system considered under steady state. Let U be the steady state probability that the server is busy (server utilization), I be the steady state probablitity that the server is idle during the retrial time or on vacation, L_s be the mean number of customers in the system and L_q be the mean number of customers in the orbit.

Using the partial probability generating fuctions derived in Theorem 2.4, we obtain

$$U = Q_1(1) + Q_2(1) = \frac{P_0\lambda\Big\{R^*(\lambda)V^*(\lambda(1-b)) + \lambda(1-a)(1-b)E(V)\Big\}[E(S_1)+pE(S_2)]}{V^*(\lambda(1-b))\Big\{R^*(\lambda) - \lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]\Big\}},$$

$$I = P_0 + P(1) + V(1) = \frac{\begin{array}{r} P_0\Big\{R^*(\lambda)V^*(\lambda(1-b))\Big[1+\lambda\Big(abE(V)-(1-b)[E(S_1)+pE(S_2)]\Big)\Big] \\ +(1-a)\lambda\Big[((1-b)+R^*(\lambda))E(V) - \lambda(1-b)[E(S_1)+pE(S_2)]\Big]\Big\} \end{array}}{V^*(\lambda(1-b))\Big\{R^*(\lambda)-\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]\Big\}}.$$

The mean number, L_s, of customers in the system under steady state condition is obtained as

$$\begin{aligned} L_s &= K'(1) \\ &= \frac{\begin{array}{r} \lambda^2(1-b)\Big\{R^*(\lambda)V^*(\lambda(1-b))\Big[E(S_1^2)+pE(S_2^2)+aE(V^2)+2pE(S_1)E(S_2)+2aE(V)E(S_1) \\ +2apE(V)E(S_2)\Big]\Big\} + \lambda^2(1-b)\Big\{(1-a)[2E(V)[E(S_1)+pE(S_2)] - b(1-R^*(\lambda))E(V^2)]\Big\} \\ +2\lambda\Big\{(1-a)(1-b)(1-R^*(\lambda))E(V) + R^*(\lambda)V^*(\lambda(1-b))[E(S_1)+pE(S_2)]\Big\} \end{array}}{\begin{array}{r} 2(1-b)V^*(\lambda(1-b))\Big\{V^*(\lambda(1-b))R^*(\lambda)\Big(1+\lambda b[E(S_1)+pE(S_2)+aE(V)]\Big) \\ +\lambda(1-a)(1-b)E(V)(1-R^*(\lambda))\Big\} \end{array}} \\ &\quad - \frac{\begin{array}{r} \lambda^2(1-b)^2\Big\{E(S_1^2)+pE(S_2^2)+aE(V^2)+2pE(S_1)E(S_2)+2aE(S_1)E(V)+2apE(S_2)E(V)\Big\} \\ +2\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)](1-R^*(\lambda)) \end{array}}{2\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)] - R^*(\lambda)}. \end{aligned}$$

The mean number, L_q, of customers in the orbit under steady state condition is given as

$$\begin{aligned} L_q &= H'(1) \\ &= \frac{\begin{array}{r} \lambda\Big\{b(1-b)\lambda R^*(\lambda)V^*(\lambda(1-b))\Big[E(S_1^2)+pE(S_2^2)+aE(V^2)+2pE(S_1)E(S_2)+2aE(S_1)E(V) \\ +2apE(S_2)E(V)\Big] + (1-a)(1-b)\Big[\lambda((1-b)(1-R^*(\lambda))+R^*(\lambda))E(V^2)+2(1-R^*(\lambda))E(V)\Big]\Big\} \end{array}}{\begin{array}{r} 2(1-b)V^*(\lambda(1-b))\Big\{V^*(\lambda(1-b))R^*(\lambda)\Big(1+\lambda b[E(S_1)+pE(S_2)+aE(V)]\Big) \\ +\lambda(1-a)[1-b(1-R^*(\lambda))]E(V)\Big\} \end{array}} \\ &\quad - \frac{\begin{array}{r} \lambda\Big\{\lambda(1-b)\Big[E(S_1^2)+pE(S_2^2)+aE(V^2)+2pE(S_1)E(S_2)+2aE(S_1)E(V)+2apE(S_2)E(V)\Big] \\ +2[E(S_1)+pE(S_2)+aE(V)](1-R^*(\lambda))\Big\} \end{array}}{2V^*(\lambda(1-b))\Big\{\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]-R^*(\lambda)\Big\}}. \end{aligned}$$

Further, the probability, P_{EO}, of orbit being empty is defined by

$$P_{EO} = P_0 + Q_{1,0} + Q_{2,0} + V_0,$$

where $Q_{i,0}$, $i = 1, 2$, are the probabilities that the orbit is empty while the server is busy with FPS and SPS respectively, V_0 is the probability that the orbit is empty while the server is on vacation and P_0 is the probability of an empty system. We observe that

$$\begin{aligned}
Q_{1,0} &= \frac{P_0(1 - S_1^*(\lambda(1-b)))}{(1-b)V^*(\lambda(1-b))S_1^*(\lambda(1-b))\Big(q + pS_2^*(\lambda(1-b))\Big)},\\
Q_{2,0} &= \frac{pP_0(1 - S_1^*(\lambda(1-b)))}{(1-b)V^*(\lambda(1-b))\Big(q + pS_2^*(\lambda(1-b))\Big)},\\
V_0 &= \frac{P_0(1 - V^*(\lambda(1-b)))}{(1-b)V^*(\lambda(1-b))},
\end{aligned}$$

and hence

$$P_{EO} = \frac{P_0\Big\{1 - bS_1^*(\lambda(1-b))\Big(q + pS_2^*(\lambda(1-b))\Big)V^*(\lambda(1-b))\Big\}}{(1-b)V^*(\lambda(1-b))S_1^*(\lambda(1-b))\Big(q + pS_2^*(\lambda(1-b))\Big)},$$

where P_0 is as given in (3.51). Let W_s be the average time a customer spends in the system under steady state. Due to Littles formula, we have,

$$W_s = \frac{L_s}{\lambda(1-b)}.$$

Another interesting performance measure in retrial context is the mean of the system busy period. The system busy period B is defined as the period that starts at an epoch when an arriving customer finds an empty system and ends at the next departure epoch at which the system is empty. The mean length of the system busy period, $E(B)$, of the model under investigation is obtained in a direct way. By using the theory of regenerative processes which leads to the limiting probability $P_0 = \lim_{t\to\infty} P\{(C(t), X(t)) = (0, 0)\}$, we have

$$P_0 = \frac{E(T_{00})}{\frac{1}{\lambda(1-b)} + E(B)},$$

where T_{00} is the amount of time in a regenerative cycle during which the system is in the state (0,0). Clearly, we have

$$E(T_{00}) = \frac{1}{\lambda(1-b)}.$$

Substituting for $E(T_{00})$ and rearranging the terms, we get

$$E(B) = \frac{1}{\lambda(1-b)(P_0^{-1} - 1)}.$$

Using (3.51) in the above, we obtain

$$E(B) = \frac{[(1-b) + bR^*(\lambda)]\Big\{V^*(\lambda(1-b))[E(S_1) + pE(S_2) + aE(V)] + (1-a)E(V)\Big\}}{(1-b)V^*(\lambda(1-b))\Big\{R^*(\lambda) - \lambda(1-b)[E(S_1) + pE(S_2) + aE(V)]\Big\}}.$$

2.5. Single vacation with exhaustive service

In this section, we discuss the $M/G/1$ retrial queue with second optional service and customer balking under single vacation with exhaustive service. By taking $a = 0$ in our system with Bernoulli vacation discussed in previous subsections we obtain the results of a special case where the vacation schedule is governed by exhaustive service *i.e.*, the server goes on vacation only when there are no customers waiting in the orbit. The probability generating function, $K(z)$, of the system size and the probability P_0 that the system is empty are obtained as

$$K(z) = \frac{\begin{array}{r} P_0\Big\{S_1^*(\lambda(1-b)(1-z))\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)\Big[\Big(1-V^*(\lambda(1-b)(1-z))\Big)[z+(1-z)R^*(\lambda)] \\ +V^*(\lambda(1-b))R^*(\lambda)[(1-b)-z]\Big] + bz\Big[V^*(\lambda(1-b))R^*(\lambda) - \Big(1-V^*(\lambda(1-b)(1-z))\Big)(1-R^*(\lambda))\Big]\Big\} \end{array}}{(1-b)V^*(\lambda(1-b))\Big\{[z+(1-z)R^*(\lambda)]S_1^*(\lambda(1-b)(1-z))(q+pS_2^*(\lambda(1-b)(1-z))) - z\Big\}}$$

and

$$P_0 = \frac{V^*(\lambda(1-b))\Big\{R^*(\lambda) - \lambda(1-b)[E(S_1)+pE(S_2)]\Big\}}{\lambda[1-b(1-R^*(\lambda))]E(V) + V^*(\lambda(1-b))R^*(\lambda)\Big(1+\lambda b[E(S_1)+pE(S_2)]\Big)}.$$

The mean number, L_s, of customers in the system under steady state is obtained as:

$$L_s = \frac{\begin{array}{r} \lambda^2(1-b)\Big\{R^*(\lambda)V^*(\lambda(1-b))\Big[E(S_1^2)+pE(S_2^2)+2pE(S_1)E(S_2)\Big] + 2E(V)[E(S_1)+pE(S_2)] \\ -b(1-R^*(\lambda))E(V^2))\Big\} + 2\lambda\Big\{(1-b)(1-R^*(\lambda))E(V) + R^*(\lambda)V^*(\lambda(1-b))[E(S_1)+pE(S_1)]\Big\} \end{array}}{2(1-b)V^*(\lambda(1-b))\Big\{V^*(\lambda(1-b))R^*(\lambda)\Big(1+\lambda b[E(S_1)+pE(S_2)]\Big) + \lambda(1-b)E(V)(1-R^*(\lambda))\Big\}}$$

$$- \frac{\lambda^2(1-b)^2\Big[E(S_1^2)+pE(S_2^2)+2pE(S_1)E(S-2)\Big] + 2\lambda(1-b)(1-R^*(\lambda))[E(S_1)+pE(S_2)]}{2\lambda(1-b)[E(S_1)+pE(S_2)] - R^*(\lambda)}.$$

The mean number, L_q, of customer waiting in the orbit under steady state is got as

$$Lq = \frac{\begin{array}{r} \lambda\Big\{b(1-b)\lambda R^*(\lambda)V^*(\lambda(1-b))[E(S_1^2)+pE(S_2^2)+2pE(S_1)E(S_2)] \\ +(1-b)\Big[\lambda\Big((1-b)(1-R^*(\lambda))+R^*(\lambda)\Big)E(V^2) + 2(1-R^*(\lambda))E(V)\Big]\Big\} \end{array}}{2(1-b)V^*(\lambda(1-b))\Big\{V^*(\lambda(1-b))R^*(\lambda)\Big(1+\lambda b[E(S_1)+pE(S_2)]\Big) + \lambda E(V)[1-b(1-R^*(\lambda))]\Big\}}$$

$$- \frac{\lambda\Big\{\lambda(1-b)[E(S_1^2)+pE(S_2^2)+2pE(S_1)E(S_2)] + 2[E(S_1)+pE(S_2)](1-R^*(\lambda))\Big\}}{2V^*(\lambda(1-b))\Big\{\lambda(1-b)[E(S_1)+pE(S_2)] - R^*\lambda\Big\}}.$$

It is noted that the expressions for $K(z)$, $H(z)$ and P_0 agree with Arivudainambi and Godhandaraman [5].

2.6. Single vacation with 1-limited service

We deduce the PGF for the system size and other performance measures of an $M/G/1$ retrial queue with second optional service and customer balking under signle vacation with 1- limited service in this section by taking $a = 1$ in our model.

$$K(z) = \frac{P_0\Big\{S_1^*(\lambda(1-b)(1-z))\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)\Big[\Big(1-V^*(\lambda(1-b)(1-z))\Big)R^*(\lambda)V^*(\lambda(1-b)) + V^*(\lambda(1-b))R^*(\lambda)[(1-b)V^*(\lambda(1-b)(1-z))-z]\Big] + zbV^*(\lambda(1-b))R^*(\lambda)\Big\}}{(1-b)V^*(\lambda(1-b))\Big\{S_1^*(\lambda(1-b)(1-z))\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)[z+(1-z)R^*(\lambda)] \times V^*(\lambda(1-b)(1-z))-z\Big\}},$$

where

$$P_0 = \frac{R^*(\lambda)-\lambda(1-b)[E(S_1)+pE(S_2)+E(V)]}{R^*(\lambda)\Big\{1+\lambda b[E(S_1)+pE(S_2)+E(V)]\Big\}}.$$

The mean system size L_s is derived as

$$L_s = \frac{\lambda^2(1-b)R^*(\lambda)V^*(\lambda(1-b))\Big\{E(S_1^2)+pE(S_2^2)+E(V^2)+2pE(S_1)E(S_2)+2E(V)E(S_1) + 2pE(V)E(S_2)\Big\} + 2\lambda R^*(\lambda)V^*(\lambda(1-b))[E(S_1)+pE(S_2)]}{2(1-b)V^*(\lambda(1-b))V^*(\lambda(1-b))R^*(\lambda)(1+\lambda b[E(S_1)+pE(S_2)+E(V)])}$$

$$-\frac{\lambda^2(1-b)^2\Big\{E(S_1^2)+pES_2^2+E(V^2)+2pE(S_1)E(S_2)+2E(S_1)E(V)+2pE(S_2)E(V)\Big\} + 2\lambda(1-b)\Big\{E(S_1)+pE(S_2)+aE(V)\Big\}(1-R^*(\lambda))}{2\lambda(1-b)[E(S_1)+pE(S_2)+E(V)]-R^*(\lambda)},$$

and mean orbit size L_q is obtained as

$$L_q = H^{'}(1)$$

$$= \frac{\lambda\Big\{b(1-b)\lambda R^*(\lambda)V^*(\lambda(1-b))\Big[E(S_1^2)+pE(S_2^2)+E(V^2)+2pE(S_1)E(S_2)+2E(S_1)E(V) + 2pE(S)E(V)\Big]\Big\}}{2(1-b)V^*(\lambda(1-b))\Big\{V^*(\lambda(1-b))R^*(\lambda)\big(1+\lambda b[E(S_1)+pE(S_2)+E(V)]\big)\Big\}}$$

$$-\frac{\lambda\Big\{\lambda(1-b)\Big[E(S_1^2)+pE(S_2^2)+E(V^2)+2pE(S_1)E(S_2)+2E(S_1)E(V)+2pE(S_2)E(V)\Big] + 2(E(S_1)+pE(S_2)+E(V))(1-R^*(\lambda))\Big\}}{2V^*(\lambda(1-b))\Big\{\lambda(1-b)[E(S_1)+pE(S_2)+E(V)]-R^*(\lambda)\Big\}}.$$

Krishna Kumar and Arivudainambi [22] have studied $M/G/1$ retrial queue with Bernoulli vacation in which if we take $q = 1$, we will get the results for 1-limited service vacation system.

2.7. Special cases

Case (i). Taking $p \to 0$, $b \to 0$, in our results, we get the M/G/1 retrial queue with Bernoulli vacation. Here, the probability generating function, $K(z)$, of the system size is obtained as

$$K(z) = \frac{P_0 S_1^*(\lambda(1-b)(1-z))\Big\{(1-V^*(\lambda(1-z)))(1-a)[z+(1-z)R^*(\lambda)] + (1-z)R^*(\lambda)V^*(\lambda)\Big\}}{V^*(\lambda)\Big\{[z+(1-z)R^*(\lambda)]S_1^*(\lambda(1-z))(aV^*(\lambda(1-z))+(1-a)) - z\Big\}},$$

where

$$P_0 = \frac{V^*(\lambda)\Big\{R^*(\lambda) - \lambda[E(S_1)+aE(V)]\Big\}}{\lambda(1-a)E(V) + R^*(\lambda)V^*(\lambda)}.$$

which are consistent with the results of Krishna Kumar and Arivudainambi [25].

Case (ii). Letting $a \to 1$, $b \to 0$, and $p \to 0$, the model reduces to M/G/1 retrial queue where the server takes single vacation with 1-limited service. In this case, we get the probability generating function, $K(z)$, of the system size is obtained as

$$K(z) = \frac{\Big\{R^*(\lambda) - \lambda[E(S_1)+E(V)]\Big\}S_1^*(\lambda(1-z))(1-z)}{\Big\{[z+(1-z)R^*(\lambda)]S_1^*(\lambda(1-z))V^*(\lambda(1-z)) - z\Big\}},$$

where

$$P_0 = \frac{\{R^*(\lambda) - \lambda[E(S_1)+E(V)]\}}{R^*(\lambda)}.$$

Case (iii). If $R^*(\lambda) \to 1$, $p \to 0$, $b \to 0$, our model reduces to an M/G/1 queue with Bernoulli vacation schedule. Thus, we get the probability generating function, $K(z)$, of the number of customers in the system (see Takagi [32]) as

$$K(z) = \frac{P_0 S_1^*(\lambda(1-z))\Big\{(1-a)(1-V^*(\lambda(1-z))) + (1-z)V^*(\lambda)\Big\}}{V^*(\lambda)\Big\{S_1^*(\lambda(1-z))\Big(aV^*(\lambda(1-z))+(1-a)\Big) - z\Big\}},$$

where

$$P_0 = \frac{V^*(\lambda)\Big\{1 - \lambda[E(S_1)+aE(V)]\Big\}}{\lambda(1-a)E(V) + V^*(\lambda)}.$$

Case (iv). Allowing $R^*(\lambda) \to 1$, $p \to 0$, $b \to 0$ and $a \to 0$, our model reduces to an M/G/1 queue with exhaustive vacation. In this case, the probability generating function, $K(z)$, of the number of customers in the system obtained as

$$K(z) = \frac{[1-\lambda E(S_1)]S_1^*(\lambda(1-z))[1-V^*(\lambda(1-z)) + (1-z)V^*(\lambda)]}{[S_1^*(\lambda(1-z)) - z][\lambda E(V) + V^*(\lambda)](1-z)}.$$

Case (v). Taking $R^*(\lambda) \to 1$, $p \to 0$, $b \to 0$ and $a \to 1$, our model reduces to M/G/1 queue with 1-limited service (see Takagi [32]). Here, the probability generating function, $K(z)$, of the number of customers in the system is given as

$$K(z) = \frac{[1-\lambda(E(S_1)+E(V))]S_1^*(\lambda(1-z))(1-z)}{S_1^*(\lambda(1-z))V^*(\lambda(1-z)) - z}.$$

2.8. Stochastic decomposition

Stochastic decomposition has been widely observed in queueing systems with server vacations (Fuhrmann and Cooper [15], Doshi [11] and Takagi [32]). A key result in these analysis is that the number of customers in the system in steady state at a random point of time is distributed as the sum of two independent random variables one of which is the number of customers in the corresponding standard queueing system without vacation at a random point of time and the other random variable may have different interpretation which is usually related to the system size given that the server is on vacation.

Stochastic decomposition has also been found to hold for some M/G/1 retrial queueing models (see Artalejo [1,2]). But, in our retrial system with server vacation where customer is permitted to balk, the decomposition law becomes inapplicable because of customer balking. Shanthikumar [27] has made a remark when discussing the stochastic decomposition in M/G/1 type queues that all cases of balking and reneging cannot be accommodated. Hur and Paik [18] have also said in their concluding remarks that they could not apply the decomposition property of generalized vacation because the customers arrival rates vary as the server status changes.

However, in the absence of balking, *i.e.*, when $b = 0$, we can establish the stochastic decomposition property in our model under investigation in an elegant way. Our retrial queue with second optional service and Bernoulli vacation can be thought of as an M/G/1 queue with generalized vacations in which the vacation begins at the end of each service times. Let $\Pi(z)$ be the probability generating function of the number of customers in the M/G/1 queueing system with second optional service in steady state at a random point of time, $\chi(z)$ be the probability generating function of the number of customers in the generalized vacation system at a random point of time given that the server is on vacation or idle and $K(z)$ be the probability generating function of the random variable being decomposed. Then the stochastic decomposition law can be expressed mathematically as

$$K(z) = \Pi(z) \times \chi(z). \tag{2.36}$$

Letting $b = 0$ in (2.33), we get the probability generating function of the number of customers in the system when balking is not permitted as

$$K(z) = \frac{\begin{array}{r} P_0 S_1^*(\lambda(1-z))\Big(q + pS_2^*(\lambda(1-z))\Big)\Big\{\Big(1 - V^*(\lambda(1-z))\Big)(1-a)[z + (1-z)R^*(\lambda)] \\ +(1-z)V^*(\lambda)R^*(\lambda)\Big\} \end{array}}{V^*(\lambda)\Big\{[z + (1-z)R^*(\lambda)]S_1^*(\lambda(1-z))\Big(q + pS_2^*(\lambda(1-z))\Big)\Big(aV^*(\lambda(1-z)) + (1-a)\Big) - z\Big\}}\cdot \tag{2.37}$$

We have the probability generating function of the number of customers in the M/G/1 queueing system with second optional service in steady state (Madan [29]) as

$$\Pi(z) = \frac{\Big(1 - \lambda[E(S_1) + pE(S_2)]\Big)(1-z)S_1^*(\lambda(1-z))\Big(q + pS_2^*(\lambda(1-z))\Big)}{S_1^*(\lambda(1-z))\Big(q + pS_2^*(\lambda(1-z))\Big) - z}\cdot \tag{2.38}$$

To obtain an expression for $\chi(z)$, we first define the generalized vacation in our context. We say that the server is on vacation if the server is on regular vacation or he is idle either due to retrials of customers from the orbit (if any) or due to no customer in the system. We have,

$$\chi(z) = \frac{P_0 + P(z) + V(z)}{P_0 + P(1) + V(1)}\cdot \tag{2.39}$$

Using (2.29), (2.32) and (2.35) in the above, by taking $b = 0$, we obtain

$$\chi(z) = \frac{P_0\Big\{\Big(1 - V^*(\lambda(1-z))\Big)(1-a)[z + (1-z)R^*(\lambda)] + (1-z)V^*(\lambda)R^*(\lambda)\Big\} \times \Big[S_1^*(\lambda(1-z))\Big(q + pS_2^*(\lambda(1-z))\Big) - z\Big]}{V^*(\lambda)\Big\{[z + (1-z)R^*(\lambda)]S_1^*(\lambda(1-z))\Big(q + pS_2^*(\lambda(1-z))\Big)\Big(aV^*(\lambda(1-z)) + (1-a)\Big) - z\Big\} \times \Big[1 - \lambda[E(S_1) + PE(S_2)]\Big](1-z)} \cdot \tag{2.40}$$

From equations (2.37), (2.38) and (2.40), the stochastic decompostion property given by (2.36) is easily verified.

3. System under modified Bernoulli vacation schedule

The Bernoulli vacation schedule, considered in previous sections, assumes that at the end of a service completion if there are no customers waiting in the orbit the server goes on vacation. This type of vacation schedule was discussed by Keilson and Servi [21] for GI/G/1 vacation systems where the server is aware of the status of the customers in the system. But in the retrial queueing systems since there is no waiting line and the customers are not physically waiting in the system the server is not aware of the status of the customers in the system. Hence the Bernoulli vacation schedule should be slightly modified in retrial queueing systems. Wang and Li [34] have assumed that even when there are no customers in the system, the server either waits for the next customer to arrive with probability $1 - a$ or chooses to go on vacation with probability a. Here we discuss this type of modified Bernoulli vacation schedule.

In the case of modified Bernoulli vacation schedule, the system of steady state equations can be obtained by using the supplementary variable technique as

$$\lambda P_0 = \int_0^\infty V_0(x)\nu(x)\mathrm{d}x + (1-a)q\int_0^\infty Q_{1,0}(x)\mu_1(x)\mathrm{d}x + (1-a)\int_0^\infty Q_{2,0}(x)\mu_2(x)\mathrm{d}x, \tag{3.1}$$

and

$$V_n(0) = aq\int_0^\infty Q_{1,n}(x)\mu_1(x)\mathrm{d}x + a\int_0^\infty Q_{2,n}(x)\mu_2(x)\mathrm{d}x, \qquad n \geq 0, \tag{3.2}$$

and other governing equations and boundary conditions are same as (2.2)−(2.11). Following the similar procedure, the partial probability generating functions are obtained as

$$P(z) = \frac{P_0 z\Big\{1 - S_1^*(\lambda(1-b)(1-z))\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big)\Big\}(1 - R^*(\lambda))}{\Big\{\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big)[z + (1-z)R^*(\lambda)] \times S_1^*(\lambda(1-b)(1-z)) - z\Big\}}, \tag{3.3}$$

$$Q_1(z) = \frac{P_0 R^*(\lambda)\Big(1 - S_1^*(\lambda(1-b)(1-z))\Big)}{(1-b)\Big\{\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big)[z + (1-z)R^*(\lambda)] \times S_1^*(\lambda(1-b)(1-z)) - z\Big\}}, \tag{3.4}$$

$$Q_2(z) = \frac{P_0 p R^*(\lambda) S_1^*(\lambda(1-b)(1-z)\Big(1 - S_2^*(\lambda(1-b)(1-z))\Big)}{(1-b)\Big\{\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big)[z + (1-z)R^*(\lambda)] \times S_1^*(\lambda(1-b)(1-z)) - z\Big\}}, \tag{3.5}$$

and

$$V(z) = \frac{P_0 a R^*(\lambda) S_1^*(\lambda(1-b)(1-z))\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(1 - V^*(\lambda(1-b)(1-z))\Big)}{(1-b)\Big\{\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big)[z + (1-z)R^*(\lambda)] \times S_1^*(\lambda(1-b)(1-z)) - z\Big\}}\cdot \tag{3.6}$$

The probabaility generating functions of the system and orbit size are found as

$$K(z) = \frac{P_0 R^*(\lambda)\Big\{\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)S_1^*(\lambda(1-b)(1-z))\Big[(1-b)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big) + a\Big(1 - V^*(\lambda(1-b)(1-z))\Big) - z\Big] + bz\Big\}}{(1-b)\Big\{\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big)[z + (1-z)R^*(\lambda)] \times S_1^*(\lambda(1-b)(1-z)) - z\Big\}}, \tag{3.7}$$

and

$$H(z) = \frac{P_0 R^*(\lambda)\Big\{S_1^*(\lambda(1-b)(1-z))\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big[(1-b)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big) + a\Big(1 - V^*(\lambda(1-b)(1-z))\Big) - 1\Big] - (1-b)z + 1\Big\}}{(1-b)\Big\{\Big(q + pS_2^*(\lambda(1-b)(1-z))\Big)\Big(aV^*(\lambda(1-b)(1-z)) + (1-a)\Big)[z + (1-z)R^*(\lambda)] \times S_1^*(\lambda(1-b)(1-z)) - z\Big\}}, \tag{3.8}$$

where P_0 is given by

$$P_0 = \frac{R^*(\lambda) - \lambda(1-b)[E(S_1) + pE(S_2) + aE(V)]}{R^*(\lambda)\Big\{\lambda b[E(S_1) + pE(S_2) + aE(V)] + 1\Big\}}\cdot \tag{3.9}$$

The corresponding mean system size L_s and mean orbit size L_q are derived as:

$$L_s = \frac{\lambda\Big\{\lambda(1-b)[b(1-R^*(\lambda)) - 1]\Big(E(S_1^2) + pE(S_2^2) + aE(V^2) + 2pE(S_1)E(S_2) + 2aE(S_1)E(V) + 2apE(S_2)E(V)\Big) + 2(1-b)[E(S_1) + pE(S_2) + aE(V)]\Big(\lambda[E(S_1) + pE(S_2)] - (1 - R^*(\lambda)) \times \Big[b\Big(E(S_1) + pE(S_2) + aE(V)\Big) + 1\Big]\Big)\Big\}}{2\Big\{b\lambda(E(S_1) + pE(S_2) + aE(V)) + 1\Big\}\Big\{\lambda(1-b)[E(S_1) + pE(S_2) + aE(V)] - R^*(\lambda)\Big\}}, \tag{3.10}$$

and

$$Lq = \frac{\begin{array}{c}\lambda^2(1-b)\Big[E(S_1^2)+pE(S_2^2)+aE(V^2)+2pE(S_1)E(S_2)+2aE(S_1)E(V)+2apE(S_2)E(V)\Big]\\ \times\Big(b(1-R^*(\lambda))-1\Big)+(1-R^*(\lambda))[E(S_1)+pE(S_2)+aE(V)]\Big[1-b\lambda[E(S_1)+pE(S_2)+aE(V)]\Big]\end{array}}{2\Big\{\lambda[E(S_1)+pE(S_2)+aE(V)][b\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]-bR^*(\lambda)+(1-b)]-R^*(\lambda)\Big\}}. \tag{3.11}$$

As in Section 2.4, here also we obtain the other system performance measures as

$$U = Q_1(1)+Q_2(1) = \frac{P_0R^*(\lambda)\lambda[E(S_1)+pE(S_2)]}{R^*(\lambda)-\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]},$$

$$I = P_0+P(1)+V(1) = \frac{P_0R^*(\lambda)(1-b)\Big\{\lambda(1-b)[E(S_2)+pE(S_2)]-\lambda baE(V)-1\Big\}}{R^*(\lambda)-\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]}.$$

$$Q_{1,0} = \frac{P_0(1-S_1^*(\lambda(1-b)))}{(1-b)\Big(aV^*(\lambda(1-b))+(1-a)\Big)S_1^*(\lambda(1-b))\Big(q+pS_2^*(\lambda(1-b))\Big)},$$

$$Q_{2,0} = \frac{pP_0\Big(1-S_1^*(\lambda(1-b))\Big)S_1^*(\lambda(1-b))}{(1-b)\Big(aV^*(\lambda(1-b))+(1-a))S_1^*(\lambda(1-b))\Big(q+pS_2^*(\lambda(1-b))\Big)},$$

$$V_0 = \frac{P_0\Big(q+pS_2^*(\lambda(1-b))\Big)\Big(1-V^*(\lambda(1-b))\Big)S_1^*(\lambda(1-b))}{(1-b)\Big(aV^*(\lambda(1-b))+(1-a)\Big)S_1^*(\lambda(1-b))\Big(q+pS_2^*(\lambda(1-b))\Big)},$$

and

$$P_{EO} = \frac{P_0\Big\{1-bS_1^*(\lambda(1-b))\Big(q+pS_2^*(\lambda(1-b))\Big)V^*(\lambda(1-b))\Big\}}{(1-b)\Big(aV^*(\lambda(1-b))+(1-a)\Big)S_1^*(\lambda(1-b))\Big(q+pS_2^*(\lambda(1-b))\Big)},$$

where P_0 is as given in (3.9). In the absence of balking, *i.e.*, when $b=0$, we can also nicely establish the stochastic decomposition property for this model as

$$K(z) = \Pi(z)\times\chi(z),$$

where

$$\Pi(z) = \frac{\Big[1-\lambda[E(S_1)+pE(S_2)]\Big](1-z)S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)}{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)-z},$$

and

$$\chi(z) = \frac{\begin{array}{r}\Big\{R^*(\lambda)-\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)]\Big\}\Big\{S_1^*\big(\lambda(1-b)(1-z)\big)\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)\\ \times\Big[1-b\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)\Big]-(1-b)z\Big\}\end{array}}{\begin{array}{r}(1-b)\Big\{1+\lambda abE(V)-\lambda(1-b)[E(S_1)+pE(S_2)]\Big\}\Big\{\Big(q+pS_2^*(\lambda(1-b)(1-z))\Big)\\ \times\Big(aV^*(\lambda(1-b)(1-z))+(1-a)\Big)[z+(1-z)R^*(\lambda)]S_1^*\big(\lambda(1-b)(1-z)\big)-z\Big\}\end{array}}.$$

4. Numerical illustrations

In this section, we present the numerical results that illustrate the qualitative behaviour of the key performance measures of the queueing system under investigation in the form of graphs. We study the effect of the system parameters arrival rate λ, retrial rate θ and balking probability b on the following performance measures:

- the probability P_0 that the system is empty
- the expected number, L_s, of customers in the system
- the server utilisation U
- the probability P_{EO} that the orbit is empty.

In Figures 2−12, the service times of FPS and SPS, retrial time from the orbit and vacation time are assumed to follow the exponential distribution with density functions $S_1(x) = \mu_1\ \mathrm{e}^{-\mu_1 x}$, $S_2(x) = \mu_2\ \mathrm{e}^{-\mu_2 x}$, $a(x) = \theta\ \mathrm{e}^{-\theta x}$ and $\nu(x) = \gamma\ \mathrm{e}^{-\gamma x}$ respectively. Further, the parameters are chosen as $\mu_1 = 10, \mu_2 = 15, \nu = 4, a = 0.5$ and $p = 0.6$ satisfying the stability condition $\lambda(1-b)[E(S_1)+pE(S_2)+aE(V)] < R^*(\lambda)$. In all the Figures 2−12, we consider four cases of vacation policies namely single exhaustive vacation (SEV), single vacation with 1- limited service (1-LS), Bernoulli vacation (BV) and modified Bernoulli vacation (MBV).

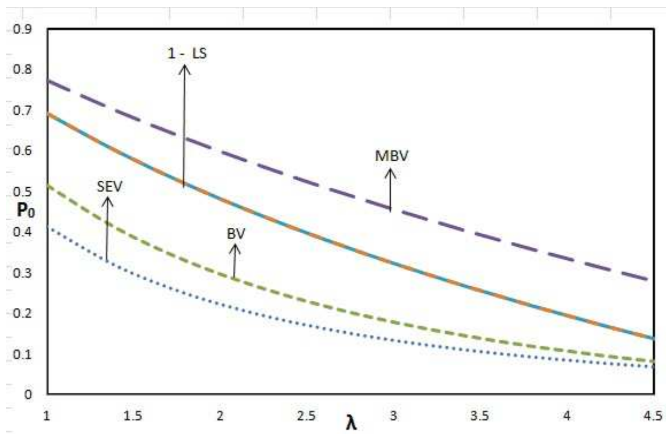


Figure 2. P_0 *versus* λ for $\theta = 0.5, b = 0.8$.

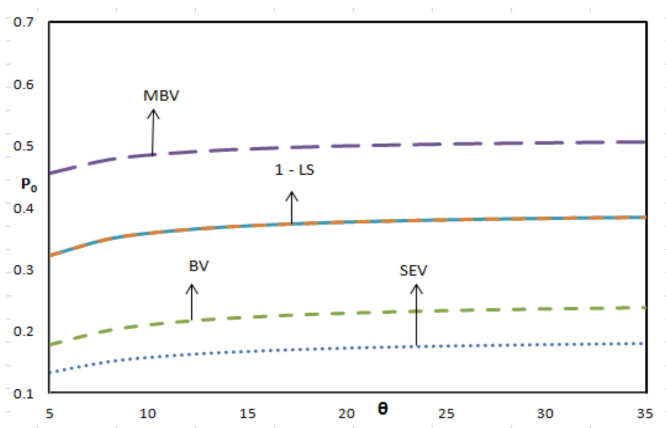


Figure 3. P_0 *versus* θ for $\lambda = 3, b = 0.8$.

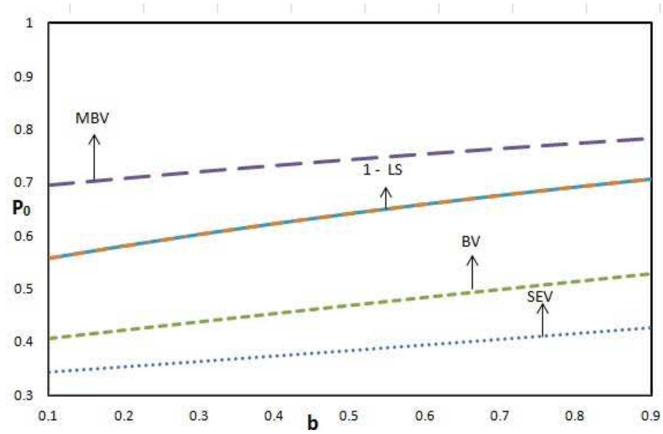


FIGURE 4. P_0 *versus* b for $\lambda = 1, \theta = 5$.

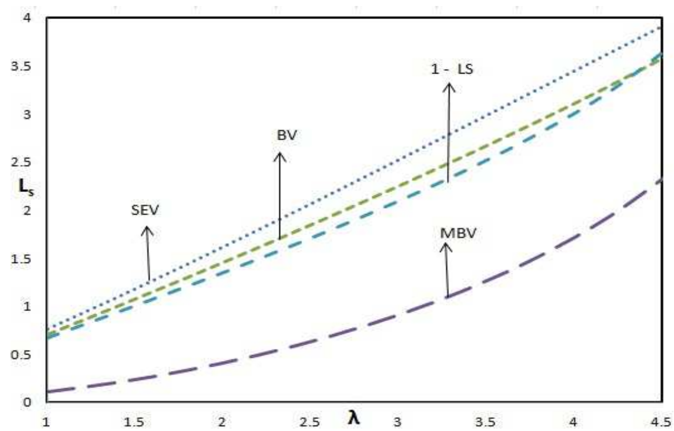


FIGURE 5. L_s *versus* λ for $\theta = 5, b = 0.8$

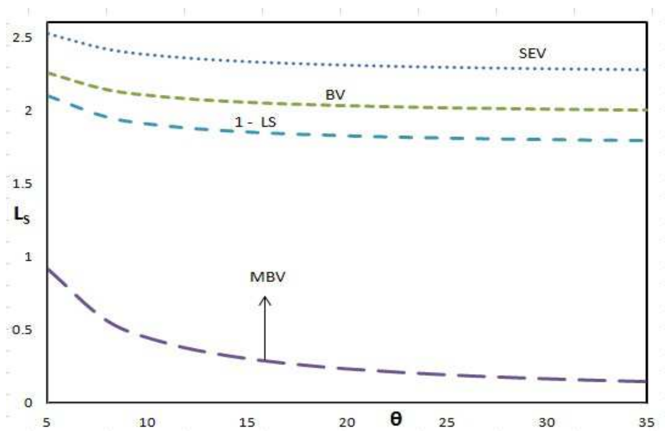


FIGURE 6. L_s *versus* θ for $\lambda = 3, b = 0.8$.

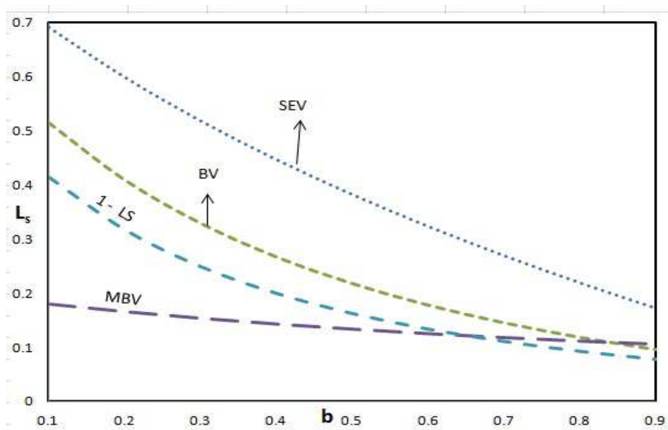


FIGURE 7. L_s *versus* b for $\lambda = 1, \theta = 5$.

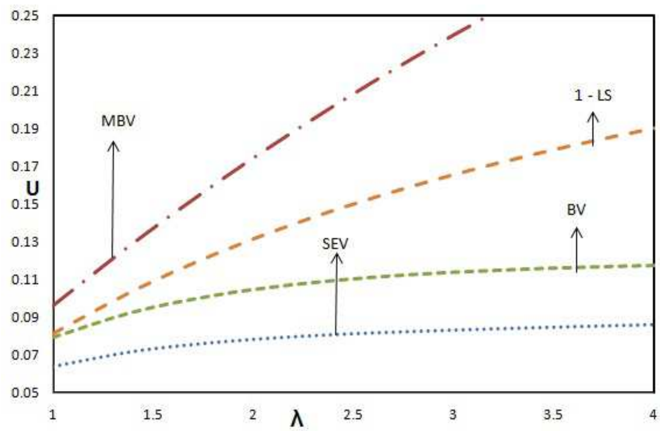


FIGURE 8. U *versus* λ for $\theta = 5, b = 0.8$.

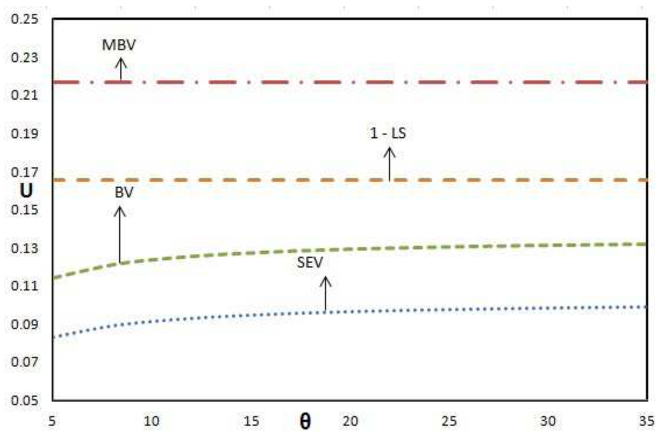


FIGURE 9. U *versus* θ for $\lambda = 3, b = 0.8$.

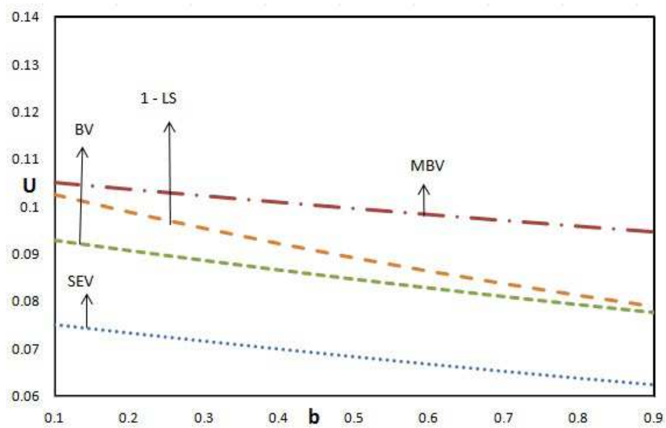


FIGURE 10. U *versus* b for $\lambda = 1, \theta = 5$.

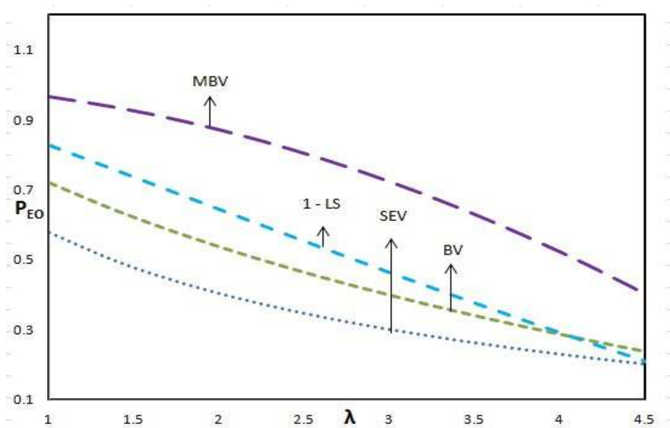


FIGURE 11. P_{EO} *versus* λ for $\theta = 5, b = 0.8$.

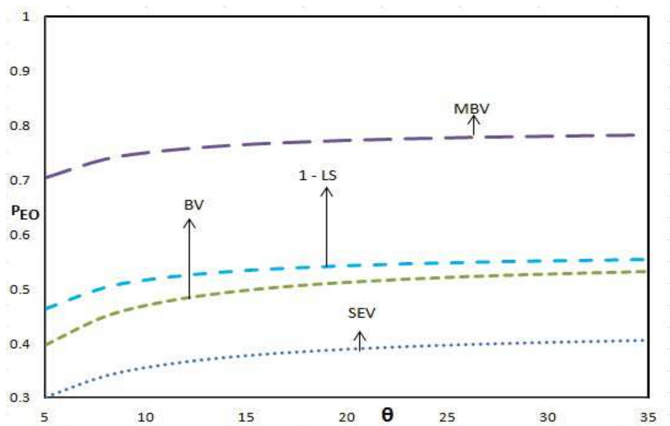
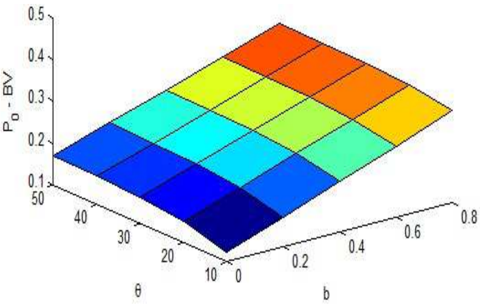
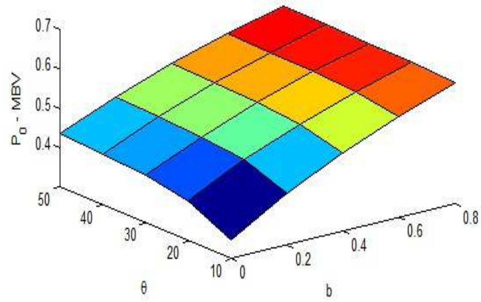


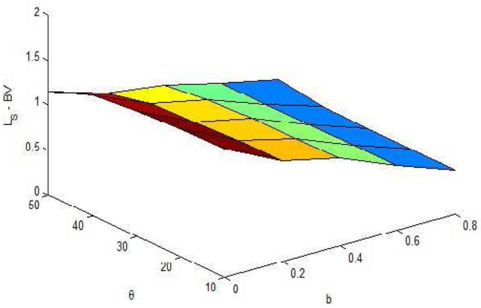
FIGURE 12. P_{EO} *versus* θ for $\lambda = 1, b = 0.8$.



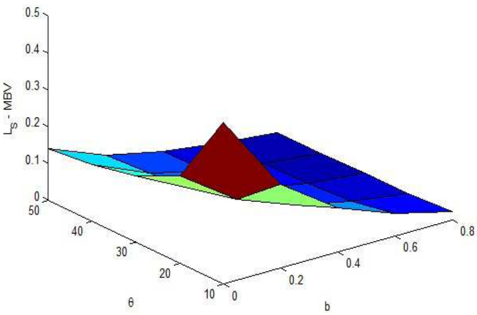
(a) P_0 (BV) versus b and θ



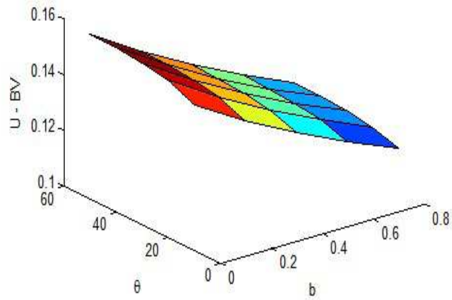
(b) P_0 (MBV) versus b and θ



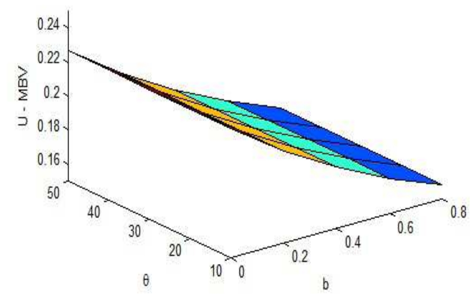
(c) $L_s(BV)$ versus b and θ



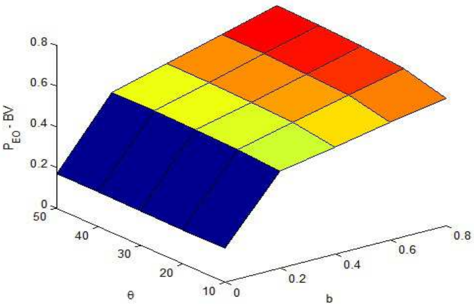
(d) L_s (MBV) versus b and θ



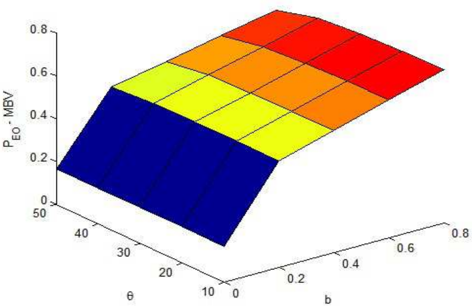
(e) $U(BV)$ versus b and θ



(f) U (MBV) versus b and θ



(g) $P_{EO}(BV)$ versus b and θ



(h) P_{EO} (MBV) versus b and θ

FIGURE 13. Combined effect of b and θ on performance measures.

In Figures 2−4, the trend of the probability P_0 is plotted against λ, θ and b. Figure 2 reveals that P_0 decreases for increasing values of λ. Figure 3 illustrates that P_0 increases for increasing values of the retrial rate θ and Figure 4 shows that P_0 increases as the balking probability b increases. It is evident from Figures 2−4 that the probability that there are no customers in the system, P_0, increases for the increasing values of θ and b and it dereases for the increasing values of λ as expected. Interestingly, it is noticed that in all the four cases, the probability P_0 is higher in the case of MBV.

Figures 5−7 depict the effect of the parameters on the mean system size L_s. In Figure 5, it is seen that L_s increases steadily for SEV, BV, 1-LS and increases slowly for MBV as λ increases. Figure 6 shows that L_s decreases for increasing values of θ. From Figure 7, it is observed that L_s decreases for increasing values of b. Figures 5− 7 reveal the fact that L_s is always lower in MBV than in other cases.

The effect of the parameters on the server utilization U is exhibited in Figures 8−10. Figure 8 shows that the utilization increases for increasing values of λ. In Figure 9, we can see a very slow increase in server utilization in the case of SEV and BV as θ increases but it is steady for 1- LS and MBV. Further the server utilization decreases for increasing values of the balking probability b as expected. It can be further observed that the utilization is higher in modified Bernoulli vacation than in other cases.

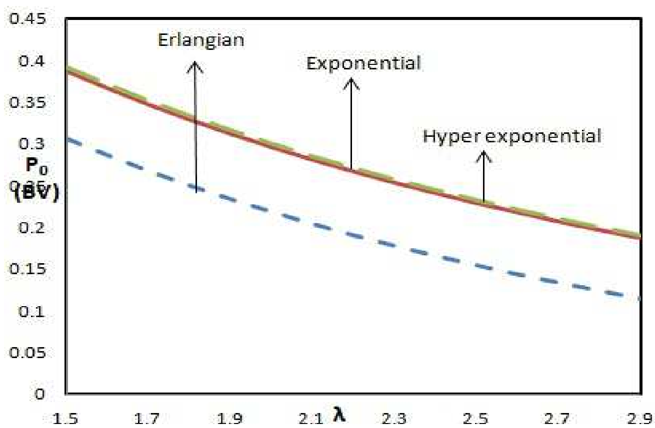


FIGURE 14. P_0 (BV) *versus* λ.

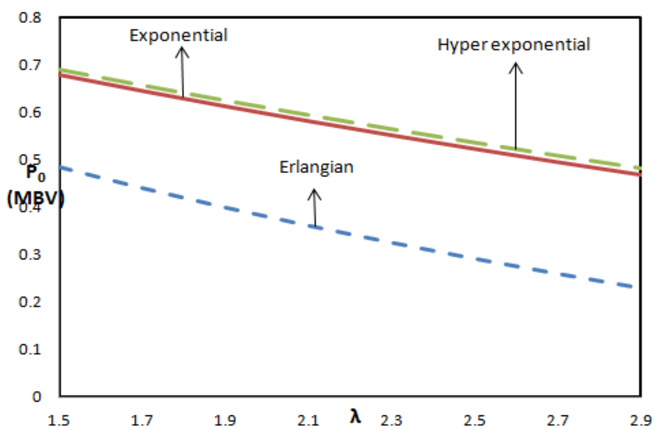


FIGURE 15. P_0 (MBV) *versus* λ.

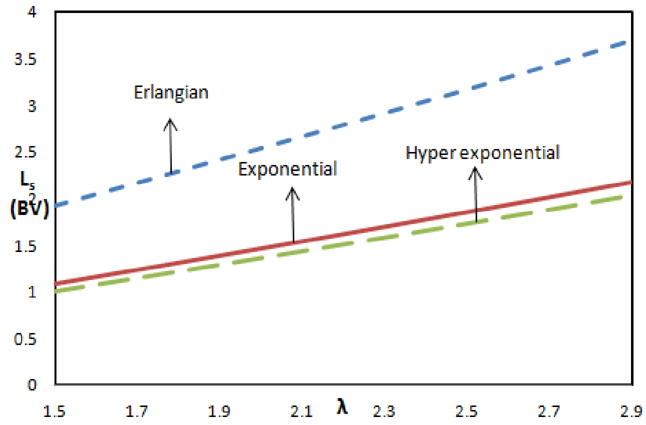


FIGURE 16. $L_s(BV)$ *versus* λ.

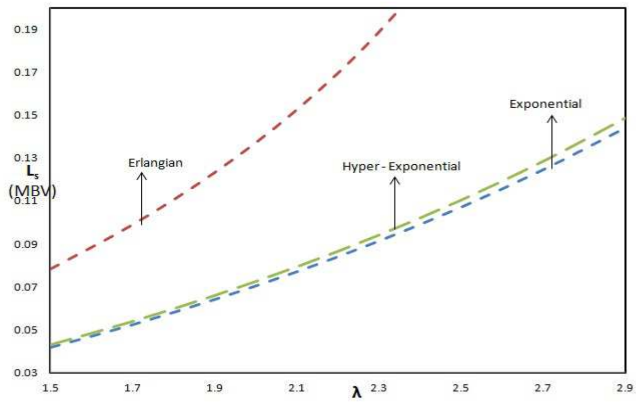


FIGURE 17. L_s (MBV) *versus* λ.

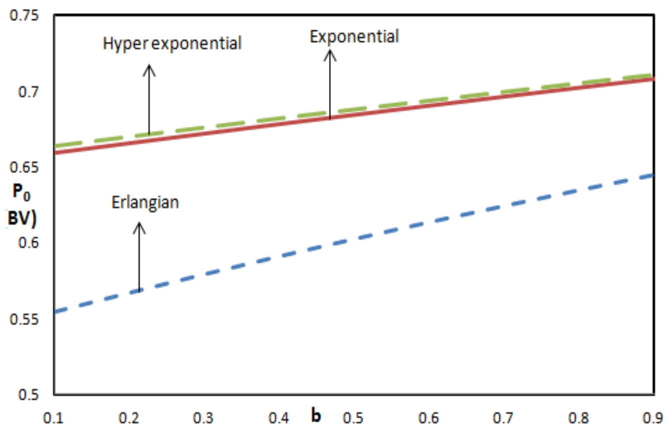


FIGURE 18. P_0 (BV) *versus* b.

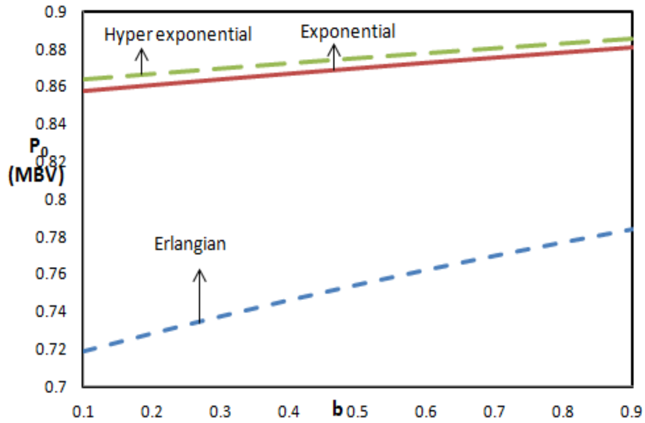


FIGURE 19. P_0 (MBV) *versus* b.

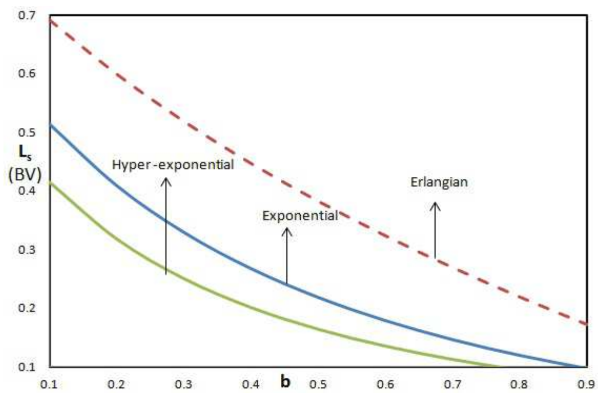


FIGURE 20. L_s (BV) *versus* b.

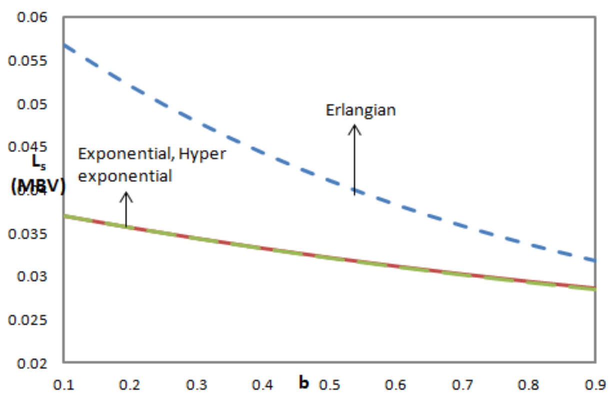


FIGURE 21. L_s (MBV) *versus* b.

Finally, the effect of the parameters on the probability P_{EO} that the orbit is empty is depicted in Figures 11 and 12. Figure 11 shows the effect of arrival rate λ on P_{EO}. The values of P_{EO} decreases for increasing values of λ as expected. Figure 12 shows there is a slow increase in P_{EO} for increasing values of θ.

Figure 13 depicts the combined effect of the retrial rate and balking probability on the performance measures P_0, L_s, U and P_{EO}, for the both BV and MBV for $\mu_1 = 10, \mu_2 = 15, \nu = 4, a = 0.5$ and $p = 0.6$.

In Figures 14−21, the effect of λ and b on L_s and P_0 are showcased for both Bernoulli and modified Bernoulli vacation schedules when the distributions (service times, retrial time, vacation time) are taken as Exponential $\left(a(x) = \mu_1 \mathrm{e}^{-\mu_1 x}\right)$, Erlangian of order two $\left(a(x) = \mu_1^2 x \mathrm{e}^{-\mu_1 x}\right)$ and Hyper Exponential $\left(a(x) = p_1 \mu_1 \mathrm{e}^{-\mu_1 x} + (1 - p_1)\mu_2 \mathrm{e}^{-\mu_2 x},\ 0 < p_1 < 1\right)$ by chosen parametric values as $\lambda = 0.5,\ \mu_1 = 10, \mu_2 = 15,\ p = 0.6, p_1 = 0.4, \theta = 5, \nu = 4, a = 0.5$ and $b = 0.8$.

From the graphs 2−12, we can find an interesting fact that the performance of the system is better in modified Bernoulli vacation compared with the other vacation schedules in terms of system stability and server utilization. The higher values of P_0 and lower values of L_s for modified Bernoulli vacation ensures that the stability of the system is higher. The server utilization and probability of an empty orbit are also higher in modified Bernoulli schedule than the other vacation polices. In view of the above, we can conclude that among all the vacation polices the modified Bernoulli vcation policy gives the better system performance. The trends observed in Figure 13 strongly support the conclusion drawn from Figures 2−12. Further, Figures 14−21 confirms that the modified Bernoulli vacation policy gives the better system performance when we take different distributions for service times, retrial time and vacation time.

5. Conclusion

For a single server retrial queue with second optional service where the customer is permitted to balk if his service is not immediate, the supplementary variable technique has been used to study the steady state system size and orbit size distribution along with their means. Other system performance measure and orbit characteristics are also computed. Further the general stochastic decomposition law for M/G/1 vacation models is shown to hold good for our system when balking is not permitted. Moreover, extensive numerical illustrations are provided to showcase the effect of the parameters on the performance measures. Interestingly, it is observed from the graphs that the modified Bernoulli vacation schedule gives better performance of the system compared with the other vacation schedules in terms of stability and also utilization.

References

[1] J.R. Artalejo, Accessible bibliography on retrial queue. *Math. Comput. Model.* **30** (1999) 1–6.

[2] J.R. Artaejo, Accessible bibliography on Retrial Queues, progress in 2000–2009. *Math. Comput. Model.* **51** (2010) 1071–1081.

[3] J.R. Artalejo and G. Falin, Standard and retrial queueing systems: a comparative analysis. *Rev. Math. Comput.* **15** (2002) 101–129.

[4] J.R. Artalejo, Analysis of an M/G/1 queue with constant repeated attempts and server vacation. *Comput. Opens. Res.* **24** (1997) 493–504.

[5] D. Arivudainambi and P. Godhandaraman, Retrial queueing system with balking, optional service and vacation. *Ann. Oper. Res.* **229** (2015) 67–84.

[6] J.R. Artalejo and A. Gomez−Corral, Retrial queueing systems: A Computational Approach. Springer, Berlin (2008).

[7] Ch. Banerjee, A. Kundu, A. Agarwal, P. Singh, S. Bhattacharya and R. Dattagupta, Priority based K -Erlang Distribution Method in Cloud Computing. *Int. J. Recent Trends in Engineering and Technology* **10** (2014) 135–144.

[8] B.D. Choi, K.K. Park and C.E.M. Pearce, An M/M/1 retrial queue with control policy and general retrial times. *Queueing Syst.* **14** (1993) 275–292.

[9] R.B. Cooper, Introduction to Queueing Theory. North Holland, New York (1981).

[10] B.T. Doshi, Queueing system with vacations-A survey, *Queueing Syst.* **1** (1986) 29–66.

[11] B.T. Doshi, A note on Stochastic decomposition in a GI/G/1 queue with vacation or set up times. *J. Appl. Prob.* **22** (1985) 419–428.

[12] G.I. Falin and T.G.C. Templeton, Retrial Queues. Chapman and Hall, London (1997).

[13] G.I. Falin, A survey of retrial queues. *Queueing Syst.* **7** (1990) 127–167.

[14] G. Fayolle, A simple telephone exchange with delayed feedbacks. In Teletraffic Analysis and Computer Performance Evaluation, edited by O.J. Boxma, J.W. Cohen, H.C. Tijms. Elsevier, Amsterdam (1986) 245–253.

[15] S.W. Fuhramann and R.B. Cooper, Stochastic Decomposition in M/G/1 queue with generalized vacations. *Oper. Res.* **33** (1985) 1117–1129.

[16] E.N. Gilbert, Retrials and Balks. *IEEE Transactions on Information Theory* **34** (1988) 1502–1508.

[17] A. Gomez−Corral, Stochastic analysis of a single server retrial queue with general retrial times. *Naval Res. Logist.* **46** (1999) 561–581.

[18] S. Hur and S.J. Paik, The effect of different arrival rates on the N-policy of M/G/1 with server set up. *Appl. Math. Model.* **23** (1999) 289–299.

[19] J.C. Ke, C.H. Wu and Z.G. Zhang, Recent Developments in Vacation Queueing models: A short survey. *Inter. J. Oper. Res.* **7** (2010) 3–8.

[20] J.C. Ke and F.M. Chang, Modified vacation policy for M/G/1 retrial queue with balking and feedback. *Comput. Ind. Eng.* **57** (2009) 433–443.

[21] J. Keilson and L.D. Servi, Oscillating random walk models for GI/G/I vacation systems with Bernoulli schedule. *J. Appl. Probab.* **23** (1986) 790–802.

[22] B. Krishna Kumar and D. Arivudainambi, The $M/G/1$ retrial queue with Bernoulli schedule and general retrial time. *Comput. Math. Appl.* **43** (2002) 15–30.

[23] B. Krishna Kumar and S. Pavai Madheswari, $M^X/G/1$ Retrial Queue with Multiple Vacations and Starting Failures. *Opsearch* **40** (2003) 115–137.

[24] B. Krishna Kumar, S. Pavai Madheswari and D. Ariudainambi, On the Busy Period of an $M/G/1$ Retrial Queueing System with Two-Phase Service and Preemptive Resume. *Stochastic Modell. Appl.* **8** (2005) 18–34.

[25] B. Krishna Kumar, S. Pavai Madheswari and A. Vijayakumar, The M/G/1 retrial queue with feedback and starting failure. *Appl. Math. Model.* **26** (2002) 1057–1075.

[26] L.I. Sennott and P.A. Humblet and R.L. Tweedi, Mean drifts and non ergodicity of Markov chains. *Oper. Res.* **31** (1983) 783–789.

[27] J. Shanthikumar, On Stochastic decomposition in M/G/1 type queue with generalized server vacations. *Oper. Res.* **36** (1988) 566–569.

[28] K.C. Madan, An M/G/1 queue with second optional service. *Queueing Syst.* **34** (2000) 37–46.

[29] K.C. Madan and G. Choudhury, A single server queue with two phases of heterogeneous service under Bernoulli schedule and a generalized vacation time. *Int. J. Inform. Manage. Sci.* **16** (2005) 1–16.

[30] M. Martin and J.R. Artalejo, Analysis of an M/G/1 queue with two types of impatient units. *Adv. Appl. Probab.* **27** (1995) 840–861.

[31] J. Medhi, A single server Poisson input queue with a second optional channel. *Queueing Syst.* **42** (2002) 239–242.

[32] H. Takagi, Queueing Analysis, Volume 1:Vacation and priority systems. North-Holland, Amsterdan (1991).

[33] J. Wang, An M/G/1 queue with second optional service and server breakdowns. *Comput. Math. Appl.* **47** (2004) 1713–1723.

[34] J. Wang and J. Li, A repairable $M/G/1$ Retrial Queue with Bernoulli vacation and two phase service. *Quality Technology and Quantitative Manag.* **5** (2008) 179–192.

[35] X. Wu, P. Brill, M. Hlynka and J. Wang, An $M/G/1$ retrial queue with balking and retrial during service. *Int. Oper. J. Res.* (2005) 30–51.

[36] Z. Wenhui, Analysis of a single server retrial queue with FCFS orbit and Bernoulli vacation. *Appl. Math. Comput.* **161** (2005) 353–364.

[37] T. Yang and J.G.C. Templeton, A survey on retrial queue. *Queueing Syst.* **2** (1987) 201–233.