Analysis of MAP/PH/1 Retrial Queue With Constant Retrial Rate, Bernoulli Schedule Vacation, Bernoulli Feedback, Breakdown and Repair

G. Ayyappan, R. Gowthami
•
Department of Mathematics, Pondicherry Engineering College
ayyappanpec@hotmail.com, gowthami9424@gmail.com

Abstract

A retrial queueing model in which the inter arrival times follow Markovian Arrival Process(MAP) and the service times follow phase type distribution is studied. At the end of receiving the service, the customer has two options namely, either he may go to orbit with probability q_1 to get the service again, if he is not satisfied or with probability p_1, he may depart the system. Similarly, at the end of providing service, the server can either opt to take vacation with probability p_2 or be idle with probability q_2. During the busy period, the server may experience breakdown. Both the breakdown times and repair times of the server follow exponential distribution with parameter σ and δ respectively. The resulting QBD process is analysed in the steady state by employing matrix analytic method. The busy period analysis of our model has also been done. Finally, the numerical and graphical illustration of our model has been given.

Keywords: Markovian arrival process, Phase type distribution, Retrial queues, Bernoulli vacation, Bernoulli feedback, Breakdown and Repair.

I Introduction

On contrary to the normal queueing model, customers arriving to the retrial queueing system join the orbit, if all the servers are busy. They stay in the orbit for some amount of time which is usually exponentially distributed and then they try to know whether there is a possibility to receive service. If the server is free at that moment, they start to receive the service; if not, they come back to the orbit and repeat the process again. In this situation, if the server is available, then there is also a possibility for the new arrival to receive service directly without joining the orbit. Due to this complexity, it is usually tedious to derive analytic results for the retrial models. However, a vast amount of numerical and approximation methods are available to study the retrial queueing systems.

One of the most versatile modelling tools in the theory of point processes is the Markovian Arrival Process(MAP). Neuts(1979) introduced a new concept namely Versatile Markovian PointProcesses(VMPP) to model the arrival processes which are not essentially renewal processes. In order to understand VMPP in a clear and simpler way, Lucantoni et al (1990) coined two new terms, namely MAP and Batch MAP. Chakravarthy (2010) have greatly discussed about MAP in the Encyclopaedia of Operations Research and Management Science. A MAP is usually characterized by the parameter matrices (D_0, D_1), each of which of dimension n in which D_0 governs the change over related to no arrivals whereas D_1 governs the change over related to arrivals. The generator matrix of the resulting continuous time Markov chain is given by $Q = D_0 + D_1$. The point process characterized by the MAP is a peculiar class of semi-Markov process whose transition probability matrix is given by

$$\int_0^t e^{D_0 x} D_1 dx = [I - e^{D_0 t}](-D_0)^{-1} D_1.$$

Let the transition probability vector of the generator matrix $D = D_0 + D_1$ be denoted by π such that $\pi Q = 0$ and $\pi e = 1$. Then, the average arrival count per unit time in the stable version of the MAP, also termed as the fundamental rate is defined as $\lambda = \pi D_1 e_m$. Latouche et al (1999) have deeply discussed about PH-distributions and QBD.

If the duration between consecutive retrials are exponentially distributed with rate nα, where n is the size of the orbit, then such retrial queues are said to follow classical retrial policy and most of the retrial queues follow this policy. But, present situation in communication protocols and local area networks indicates that there are a vast number of queueing situations in which the retrial rate does not depend on the orbit size. This is known as constant retrial policy and it was first studied by Fayolle (1986) while examining the telephone exchange model as a classical single server Markovian retrial queue. He has modelled his system in such a way that the orbital customers make a queue and the request for service can be made by the foremost customer in the waiting line and the retrial times follow exponential distribution with rate α. Artalejo et al (2000) studied the Markovian retrial queueing system with multiserver and constant retrial rate.

If a server gets breakdown, then the customer who is currently receiving service has the choice of being in the system till the repair process gets over or permanently move out of the system or go back to the orbit to resume the service. Such situations are mostly seen in computer and communication networks, at airport with stacked aircraft and in retail shops. Aissani (1998), Kulkarni and Choi (1990) have studied about the effect of unreliable server and the repair process on retrial queues. Reliability of the non-Markovian queueing system when the server is prone to breakdown was analysed by Aissani and Artalejo (1998).

Wang et al (2009) have greatly analyzed the discrete time retrial queue in which server is subject to breakdown and repair and obtained generating function for both system size and orbit size. Retrial queues with batch Markovian Arrival Process, breakdown and repair has been investigated by Li et al (2006) They studied the system by combining matrix-analytic method and the censoring technique with the supplementary variable method. By using Generalized Stochastic Petri Nets(GSPN), Gharbi et al (2011) have proposed an approach for analyzing the finite source retrial systems in which server is prone to breakdown and repair. Efrosin et al (2011) have examined the Markovian retrial queueing model with constant retrial rate and an unreliable server. Dimitriou et al (2010) have investigated the repairable retrial model and derived the stability condition for it.

The concept of Bernoulli schedule vacation is that after providing service to the customer, the server may opt to take vacation with probability p or starts the service to next customer with probability $1 - p$. Keilson et al (1986) have analyzed the non-Markovian vacation queueing model with Bernoulli schedules. The average waiting time for the non-Markovian cyclic service queues with Bernoulli schedules has been derived by Servi (1986).

The concept of Bernoulli feedback is that after receiving the service, the customer can go to orbit with probability p to get the service again or departs the system permanently with probability $q(p + q = 1)$. While investigating the single server non-Markovian model, Tak`αcs (1963) has introduced the concept of feedback mechanism. Krishnakumar et al (2010) have studied the single server Markovian retrial queueing system with feedback and derived the joint distribution of the server state and the orbit size by using generating function technique. Choudhury et al (2005) have derived the distribution for queue size and busy period for the non-Markovian queueing model with two stages of non-homogeneous services and Bernoulli feedback mechanism. Chen et al (2015) have greatly analysed the concept of Bernoulli vacation policy and Bernoulli feedback for the retrial model in discrete time. Badamchi Zadeh et al (2008) have discussed about the non-Markovian queueing system with Bernoulli feedback and Bernoulli vacation. Retrial queues in which the server is prone to experience breakdown and repair have been studied by Kulkarni and Choi (1990). The steady state probability vectors of our model are computed by employing matrix geometric method which was introduced by M. F. Neuts (1981). The rate matrix is obtained by means of the logarithmic reduction algorithm specified by Latouche et al (1999).

Hunter (1983) has shown that the queue length process examined at every distinct embedding is a Markov renewal process. He has given more emphasis on finding discrete time outputs for the queue length processes. Finally, he has examined the Markovian models with finite/infinite waiting space and the instantaneous Bernoulli feedback. A non-Markovian queueing model with Bernoulli feedback mechanism has been studied by Disney et al. (1980). They have shown that the output of their model is also Markov-renewal. Foley et al. (1983) have studied the non-Markovian queueing system with two servers and delayed feedback mechanism. They have given a choice for the customers who have received service from the lower server, either to depart the system or to go to the upper server to get the re-service, They have shown that their system is uniquely weakly lumpable to a Poisson process.

A non-Markovian batch arrival queueing model with Bernoulli vacation under multiple vacation policy has been investigated by Choudhury et al. (2018). The busy period distribution and the waiting time distribution have been derived for their model under steady state. Chakravarthy (2008) has studied the multi server queueing model in which arrival follows MAP and the vacation times follow phase type distribution. He has obtained the waiting time distribution and has presented various numerical examples. Chang et al. (2018) have studied the multi server Markovian retrial queueing model with feedback customers and unreliable servers. They have calculated the stationary distribution by developing a new recursive algorithm and have derived the cost function. Finally, they have made a comparison to validate the exactness of the approximate optimal solution.

The main motivating factor for our model is from online shopping. Nowadays, most of the people give more preferences to online shopping. Every company is selling different varieties of goods like dresses, electronic gadgets, house hold articles, etc,. These items may be considered as different phases. A customer may get service(booking a product) from any of these phases. If a customer is arriving during the busy period, he/she has to wait for some time in the invisible queue(orbit). After providing service to the customer, the server may either go for other jobs like updating their websites, launching new products, etc., (which may be considered as vacation) or the server may remain idle. During busy period, the server may get breakdown(the server problem). The customer who is receiving service at that moment has to join with the orbital customers and retry to get the service again. Moreover, after booking a particular item, if the customer is not satisfied with it(that is in the sense of price or quality of the item), the customer may again join the orbit and retry to get the service. We have framed our model in such a way that it will match with this situation.

On analysing the literature of queueing theory, we could find many articles that have discussed about the Bernoulli schedule vacation, Bernoulli feedback, breakdown and repair individually and also in different combinations for various arrival and service patterns. In our work, we have modelled our system in such a way that with all these attributes, we have considered MAP for arrival and phase type distribution for service times.

The rest of our work is structured in this manner: Section 2 briefly discusses about the model under study. The infinitesimal generator matrix is obtained in the Section 3. The analysis of this paper has been done in the steady state in the Section 4. The busy period analysis of our model has been done in the Section 5. Some of the performance measures are evaluated in Section 6. Finally, under the Section 7, the behavioural aspects of our queueing model has been analysed with the support of numerical values and graphical representations.

II Model Description

We consider a retrial queueing model with single server in which the customers come to the system as indicated by the Markovian Arrival Process with D_0 and D_1 as its parameter matrices of dimension n. The service times are supposed to follow phase type distribution with representation (α, T) of order m such that $T_0 + Te = 0$. On an arrival of a customer, if the server is available, then he provides service instantly. If not, the customer has to join the orbit of infinite capacity. Irrespective of the orbit size, each unit makes retrial at a constant rate from the orbit. The inter retrial times follow exponential distribution with parameter μ. At the end of providing service to the customer, the server may either go for vacation with probability $\mathfrak{p}_2$ or remains idle with probability $\mathfrak{q}_2$ where $\mathfrak{p}_2 + \mathfrak{q}_2 = 1$. Similarly, after receiving the service, if the customer is satisfied, then he leaves the system with probability $\mathfrak{p}_1$. Otherwise, if the customer is not satisfied, then he joins the orbit with probability $\mathfrak{q}_1$ to get the service, where $\mathfrak{p}_1 + \mathfrak{q}_1 = 1$. But, he has no priority over the orbital customers and he has to compete with them to get the service. During the busy period, the server may get breakdown. As a result, the customer who is receiving service at that time has to join the orbit of infinite capacity. After the completion of repair process, the server becomes idle. The breakdown times, the repair times and the vacation times are all supposed to follow exponential distribution with parameters σ, δ and η respectively.

III The generator matrix

The formulation of the generator matrix for our queueing model has been done in this section. We will begin our study by describing the following notations which are needed.

Notations:

- $N(t)$: Number of customers in the orbit
- I_n: An n-dimensional identity matrix
- **e**: A column vector (of needed dimension) with each of its entries as 1
- $\otimes$: Kronecker multiplication of two matrices
- $\oplus$: Kronecker addition of two matrices
- $Y(t)$ - Nature of the server at time t,

where

$Y(t) = \{0$, the server is in vacation.
1, the server is idle.
2, the server is offering service.
3, the server is in breakdown.

- S(t): Phase of the service process at time t
- M(t): Phase of the Markovian Arrival Process at time t
- λ: Rate of arrival and is defined as $\lambda = \pi D_1 e$ where π is the invariant probability vector

of

the generator matrix $D = D_0 + D_1$

- γ: Rate of service, where $\gamma = [\alpha(-T)^{-1}e]^{-1}$

It is obvious and can proved that$\{(N(t), Y(t), S(t), M(t))$:t 0$\}$ is a continuous time Markov chain (CTMC) whose state space is given below:

$$\mathbf{Y} = l^* \cup l(i)$$

where

$l(i) = \{(i, 0, l): i \geq 1, 1 \leq l \leq n\} \cup \{(i, 1, l): i \geq 1, 1 \leq l \leq n\} \cup \{(i, 2, k, l): i \geq 1, 1 \leq k \leq m, 1 \leq l \leq n\}$

$$\cup \{(i, 3, l): i \geq 1, 1 \leq l \leq n\}$$

The generator matrix of the Markov chain is as follows:

$$Q = \begin{bmatrix} B_{00} & B_{01} & 0 & 0 & 0 & 0 & 0 & B_{10} & A_1 & A_0 \\ 0 & 0 & 0 & 0 & 0 & A_2 & A_1 & A_0 & 0 & 0 \\ 0 & 0 & 0 & A_2 & A_1 & A_0 & 0 & 0 & \cdots & \cdots \\ \cdots & \ddots & \ddots & \ddots & \cdots & & & & & \end{bmatrix}$$

where,

$B_{00} = [D_0 - \eta I_n \quad \eta I_n \quad 00 \quad D_0 \quad \alpha \otimes D_1 p_1 p_2 T^0 \otimes I_n \quad p_1 q_2 T^0 \otimes I_n \quad T \oplus (D_0 - \sigma I_n)] B_{01} =$
$[D_1 \quad 0 \quad 0 \quad 00 \quad 0 \quad 0 \quad 0 q_1 p_2 T^0 \otimes I_n \quad q_1 q_2 T^0 \otimes I_n \quad I_m \otimes D_1 \quad e_m \otimes \sigma I_n]$

$B_{10} = [0 \quad 0 \quad 00 \quad 0 \quad \mu\alpha \otimes I_n 0 \quad 0 \quad 00 \quad 0 \quad 0] A_1 =$
$\begin{bmatrix} D_0 - \eta I_n & \eta I_n & 0 & 00 & D_0 - \mu I_n & \alpha \otimes D_1 & 0 p_1 p_2 T^0 \otimes I_n & p_1 q_2 T^0 \otimes I_n & T \oplus (D_0 - \sigma I_n) & 00 \\ \delta I_n & 0 & D_0 - \delta I_n & & & & & & & \end{bmatrix}$

$A_2 = \begin{bmatrix} 0 & 0 & 0 & 00 & 0 & \mu\alpha \otimes I_n & 00 & 0 & 0 & 00 \\ 0 & 0 & 0 & & & & & & & \end{bmatrix} A_0 =$
$\begin{bmatrix} D_1 & 0 & 0 & 00 & 0 & 0 & 0 q_1 p_2 T^0 \otimes I_n & q_1 q_2 T^0 \otimes I_n & I_m \otimes D_1 & e_m \otimes \sigma I_n 0 \\ 0 & 0 & D_1 & & & & & & & \end{bmatrix}$

IV Analysis of the System

The analysis of the our model has been discussed in this section under steady state.

4.1 Stability Condition

Let us define $A = A_0 + A_1 + A_2$. Then, clearly A is an infinitesimal generator matrix and as a result, we can find an invariant probability vector Ψ of A which obeys

$$\Psi A = 0; \Psi e = 1$$

where the vector Ψ is given by $\Psi = (\psi_0, \psi_1, \psi_2, \psi_3)$.

The vector Ψ, partitioned as $\Psi = (\psi_0, \psi_1, \psi_2, \psi_3)$ is computed by solving the following equations:

$$\psi_0[D - \eta I_n] + \psi_2[p_2 T^0 \otimes I_n] = 0$$

$$\psi_1[\eta I_n] + \psi_1[D_0 - \mu I_n] + \psi_2[q_2 T^0 \otimes I_n] + \psi_3[\delta I_n] = 0$$

$$\psi_1[\alpha \otimes D_1 + \mu\alpha \otimes I_n] + \psi_2[T \oplus (D - \sigma I_n)] = 0$$

$$\psi_2[e_m \otimes \sigma I_n] + \psi_3[D - \delta I_n]$$

subject to

$$\psi_0 + \psi_1 + \psi_2 + \psi_3 = 1.$$

The necessary and sufficient condition required by the system to attain stability is $\Psi A_0 e < \Psi A_2 e$

i.e.,

$$\psi_0[D_1 e_n] + \psi_2[q_1 T^0 \otimes e_n + e_m \otimes D_1 e_n + e_m \otimes \sigma e_n] + \psi_3 D_1 e_n < \psi_1 \mu e_n.$$

4.2 The Transition Probability Vector

Let the transition probability vector of the infinitesimal generator Q be indicated as **x**.

This probability vector can be partitioned as: $x = (x_0, x_1, x_2, \ldots)$, where x_0 is of dimension $2n + mn$ and x_i is of dimension $3n + mn$, for $i \geq 1$.

Since **x** is a transition probability vector of Q, the following two conditions will be satisfied by it:

$$xQ = 0 and xe = 1$$

Once the condition for the system to be stable is achieved, the invariant probability vector **x** can be computed using

$$x_{i+1} = x_1 \mathrm{R}^i, i \geq 0$$

and the remaining vectors namely, $x_0 and x_1$ can be evaluated by solving the equations given below:

$$x_0 \mathrm{B}_{00} + x_1 \mathrm{B}_{10} = 0$$

$$x_0 \mathrm{B}_{01} + x_1 [\mathrm{A}_1 + \mathrm{RA}_2] = 0$$

based on the normalizing condition

$$x_0 e_{2n+mn} + x_1 [\mathrm{I} - \mathrm{R}]^{-1} e_{3n+mn} = 1$$

The rate matrix R can be evaluated by making use of Logarithmic Reduction Algorithm proposed by Latouche et al.(1999).

Logarithmic Reduction Algorithm:

$$Step0: H \leftarrow (-A_1)^{-1} A_0, L \leftarrow (-A_1)^{-1} A_2, G = L and T = H.$$

$$Step1:$$

$$U = HL + LH$$

$$M = H^2$$

$$H \leftarrow (I - U)^{-1} M$$

$$M \leftarrow L^2$$

$$L \leftarrow (I - U)^{-1} M$$

$$G \leftarrow G + TL$$

$$T \leftarrow TH$$

$$continue Step1 until \parallel e - Ge \parallel_\infty < \varepsilon.$$

$$Step2: R = -A_0 (A_1 + A_0 G)^{-1}$$

V Analysis of the busy period

Under this section, we perform the analysis of busy period of our system.

The duration of time between the customer's arrival to the void system and the instant at which the system size reaches zero for the first time is termed as the busy period. Hence, the first passage time from level 1 to level 0 is the analogue of the busy period.

The first return time to level zero with the condition that there should be atleast one visit to a state in any other levels is termed as the busy cycle. Let us first bring in the idea of fundamental period for the purpose of analysing the busy period. As far as the QBD process is concerned, it is nothing but the first passage time from the level i to the level $i-1$, where $i \geq 2$. The discussion has to be done separately for the boundary states (i.e.) for the cases $i = 0,1$. It can be easily seen that there are $3n + mn$ states for each level i where $i \geq 1$. Therefore, while arranging the states in the lexicographic order, j^{th} state of level i may be indicated as (i, j).

NOTATIONS:

- $G_{jj'}(\mathrm{k}, \mathrm{x})$ - The conditional probability that the QBD process enters the level $i-1$ by making precisely k transitions to the left and also by entering the state (i, j') given that it started in the state (i, j) at time $t = 0$.
- $\bar{G}_{jj'}(\mathrm{z}, \mathrm{s}) = \sum_{k=1}^{\infty} \mathrm{z}^{\mathrm{k}} \int_0^{\infty} e^{-\mathrm{sx}} dG_{jj'}(\mathrm{k}, \mathrm{x}) : |\mathrm{z}| \leq 1, \mathrm{Re}(\mathrm{s}) \geq 0$
- $\bar{G}(\mathrm{z}, \mathrm{s}) - The matrix (\bar{G}_{jj'}(\mathrm{z}, \mathrm{s}))$
- $G = (G_{jj'}) = \bar{G}(1,0)$- The matrix which concerns about the first passage times without including the boundary states.
- $G_{jj'}^{(1,0)}(\mathrm{k}, \mathrm{x})$ - The conditional probability that the QBD process enters the level 0 by making precisely k transitions to the left given that it started in the level 1 at time $t = 0$.
- $G_{jj'}^{(0,0)}(\mathrm{k}, \mathrm{x})$ - The first return time to the level 0.
- $\mathbb{E}_{1j}$ - The expected first passage time from the level i to the level $i-1$, given that at time $t = 0$, the process is in the state (i, j).
- $\vec{\mathbb{E}}_1$ - The column vector with $\mathbb{E}_{1j}$ as its entries.
- $\mathbb{E}_{2j}$ - The expected number of customers who received service during the first passage

time from the level i to the level $i-1$, given that the first passage time begins in the state (i, j).

- $\vec{\mathbb{E}}_2$ - The column vector with $\mathbb{E}_{2j}$ as its entries.
- $\vec{\mathbb{E}}_1^{(1,0)}$ - The vector which gives the expected first passage times from the level 1 to the level 0.
- $\vec{\mathbb{E}}_2^{(1,0)}$ - The vector which gives the expected number of service completions in the first passage time from the level 1 to the level 0.
- $\vec{\mathbb{E}}_1^{(0,0)}$ - The expected first return time to the level 0.
- $\vec{\mathbb{E}}_2^{(0,0)}$ - The expected number of service completions in the course of first return time to the level 0.

It can be easily seen that the matrix $\bar{G}(\mathrm{z}, \mathrm{s})$ satisfies the following equation:

$$\bar{G}(\mathrm{z}, \mathrm{s}) = \mathrm{z}[\mathrm{sI} - \mathrm{A}_1]^{-1}\mathrm{A}_2 + [\mathrm{sI} - \mathrm{A}_1]^{-1}\mathrm{A}_0\bar{G}^2(\mathrm{z}, \mathrm{s})$$

Once the rate matrix R is evaluated, we can easily find the matrix G by making use of the result

$$G = -[\mathrm{A}_1 + \mathrm{RA}_2]^{-1}\mathrm{A}_2$$

The matrix G may also be evaluated by employing logarithmic reduction algorithm.

As far as the boundary states are concerned, namely 0 and 1, we have the following equations which are satisfied by $\bar{G}^{(1,0)}(\mathrm{z}, \mathrm{s})$ and $\bar{G}^{(0,0)}(\mathrm{z}, \mathrm{s})$ respectively.

$$\bar{G}^{(1,0)}(\mathrm{z}, \mathrm{s}) = \mathrm{z}[\mathrm{sI} - \mathrm{A}_1]^{-1}\mathrm{B}_{10} + [\mathrm{sI} - \mathrm{A}_1]^{-1}\mathrm{A}_0\bar{G}(\mathrm{z}, \mathrm{s})\bar{G}^{(1,0)}(\mathrm{z}, \mathrm{s})$$

$$\bar{G}^{(0,0)}(\mathrm{z}, \mathrm{s}) = [\mathrm{sI} - \mathrm{B}_{00}]^{-1}\mathrm{B}_{01}\bar{G}^{(1,0)}(\mathrm{z}, \mathrm{s})$$

. Since, the three matrices namely, G, $\bar{G}^{(1,0)}(1,0)$ and $\bar{G}^{(0,0)}(1,0)$ are stochastic, we may easily evaluate the following moments:

$$\vec{\mathbb{E}}_1 = -\frac{\partial}{\partial \mathrm{s}}\bar{G}(\mathrm{z}, \mathrm{s})|_{\mathrm{s}=0;\mathrm{z}=1} = -[\mathrm{A}_1 + \mathrm{A}_0(G + \mathrm{I})]^{-1}e$$

$$\vec{\mathbb{E}}_2 = \frac{\partial}{\partial \mathrm{z}}\bar{G}(\mathrm{z}, \mathrm{s})|_{\mathrm{s}=0;\mathrm{z}=1,} = -[\mathrm{A}_1 + \mathrm{A}_0(G + \mathrm{I})]^{-1}\mathrm{A}_2 e$$

$$\vec{\mathbb{E}}_1^{(1,0)} = -\frac{\partial}{\partial \mathrm{s}}\bar{G}^{(1,0)}(\mathrm{z}, \mathrm{s})|_{\mathrm{s}=0;\mathrm{z}=1} = -[\mathrm{A}_1 + \mathrm{A}_0 G]^{-1}[\mathrm{A}_0\vec{\mathbb{E}}_1 + e]$$

$$\vec{\mathbb{E}}_2^{(1,0)} = \frac{\partial}{\partial \mathrm{z}}\bar{G}^{(1,0)}(\mathrm{z}, \mathrm{s})|_{\mathrm{s}=0;\mathrm{z}=1} = -[\mathrm{A}_1 + \mathrm{A}_0 G]^{-1}[\mathrm{B}_{10}e + \mathrm{A}_0\vec{\mathbb{E}}_2]$$

$$\vec{\mathbb{E}}_1^{(0,0)} = -\frac{\partial}{\partial \mathrm{s}}\bar{G}^{(0,0)}(\mathrm{z}, \mathrm{s})|_{\mathrm{s}=0;\mathrm{z}=1} = -\mathrm{B}_{00}^{-1}[e + \mathrm{B}_{01}\vec{\mathbb{E}}_1^{(1,0)}]$$

$$\vec{\mathbb{E}}_2^{(0,0)} = \frac{\partial}{\partial \mathrm{z}}\bar{G}^{(0,0)}(\mathrm{z}, \mathrm{s})|_{\mathrm{s}=0;\mathrm{z}=1} = -\mathrm{B}_{00}{}^{-1}\mathrm{B}_{01}\vec{\mathbb{E}}_2^{(1,0)}$$

VI Performance Measures

In order to examine the behaviour of our model in the steady state, a few performance measures for our model are enumerated in this section.

- Probability of orbit being empty:
  $$P_{empty} = \sum_{j=0}^{1}\sum_{l=1}^{n} x_{0jl} + \sum_{k=1}^{m}\sum_{l=1}^{n} x_{02kl}$$
- Probability of server to be idle:
  $$P_{idle} = \sum_{i=0}^{\infty}\sum_{l=1}^{n} x_{i1l}$$
- Probability of server to be in vacation:
  $$P_{vacation} = \sum_{i=0}^{\infty}\sum_{l=1}^{n} x_{i0l}$$
- Probability of server to be busy:
  $$P_{busy} = \sum_{i=0}^{\infty}\sum_{k=1}^{m}\sum_{l=1}^{n} x_{i2kl}$$

- Probability of server to be in breakdown:

$$P_{breakdown} = \sum_{i=1}^{\infty} \sum_{l=1}^{n} x_{i3l}$$

- Probability of a new arrival getting into service directly:

$$P_s = \frac{1}{\lambda}\{\sum_{i=0}^{\infty} x_{i1} D_1 e\}$$

- Probability of a new arrival getting to receive service directly with a minimum of one customer waiting in the orbit:

$$P_{sw} = \frac{1}{\lambda}\{\sum_{i=1}^{\infty} x_{i1} D_1 e\}$$

- The total retrial rate at which the orbital customers appeal for service:

$$\mu^* = \mu\{\sum_{i=1}^{\infty} \sum_{l=1}^{n} x_{i0l} + \sum_{i=1}^{\infty} \sum_{l=1}^{n} x_{i1l} + \sum_{i=1}^{\infty} \sum_{k=1}^{m} \sum_{l=1}^{n} x_{i2kl} + \sum_{i=1}^{\infty} \sum_{l=1}^{n} x_{i3l}\}$$

- The effective retrial rate :

$$\mu_s = \mu\{\sum_{i=1}^{\infty} \sum_{l=1}^{n} x_{i1l}\}$$

- Expected orbit size

$$E_{orbit} = \sum_{i=1}^{\infty} i x_i e_{3n+mn}$$

- Average system size:

$$E_{system} = E_{orbit} + P_{busy}$$

- Probability of a successful retrial:

$$P_{sr} = \frac{\mu}{\mu+\lambda}\{\sum_{i=1}^{\infty} \sum_{l=1}^{n} x_{i1l}\}$$

- Mean number of successful retrial:

$$E_{srt} = \frac{\mu}{\mu+\lambda}\{\sum_{i=1}^{\infty} \sum_{l=1}^{n} i x_{i1l}\}$$

VII Numerical Results

In this section, we will analyse the behaviour of our model numerically as well as graphically. The following five different MAP representations, all of which have the same mean, say 1, are taken into consideration for the arrival process.

Erlang of order 2:

$$D_0 = [-2 \quad 20 \quad -2]; D_1 = [0 \quad 02 \quad 0]$$

Exponential:

$$D_0 = [-1]; D_1 = [1]$$

Hyperexponential:

$$D_0 = [-1.90 \quad 00 \quad -0.19]; D_1 = [1.710 \quad 0.1900.171 \quad 0.019]$$

Since, all these three arrival process are renewal, their correlation is zero.

Consider the following three phase type distribution for the service times.

Erlang of order 2:

$$\alpha = (1,0); T = [-12 \quad 120 \quad -12]$$

Exponential:

$$\alpha = (1); T = [-6]$$

Hyperexponential:

$$\alpha = (0.8,0.2); T = [-16.8 \quad 00 \quad -1.68]$$

Illustrative Example 1:

In the following tables, we examine the impact of the retrial rate μ against the expected orbit size. Fix $\lambda = 1; \gamma = 6; \mathfrak{p}_1 = 0.6; \eta = 2; \mathfrak{p}_2 = 0.4; \sigma = 1; \mathfrak{q}_1 = 0.4; \delta = 2; \mathfrak{q}_2 = 0.6$.

Table 1: *Expected orbit size - Exponential Service*

μ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
11	10.0011675	7.48381067	28.7824173
13	7.70793519	5.79663583	21.8380220
15	6.53032933	4.92315841	18.2956968
17	5.81352572	4.38926486	16.1464306
19	5.33131245	4.02923081	14.7031629
21	4.98472155	3.77006250	13.6669608

Table 2: *Expected orbit size - Erlang Service*

μ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
11	12.7430548	9.45396003	37.5253785
13	9.29498782	6.94117250	26.9239288
15	7.66359572	5.73985615	21.9503036
17	6.71251431	5.03600320	19.0617824
19	6.08961722	4.57372333	17.1738503
21	5.65003242	4.24691754	15.8431278

Table 3: *Expected orbit size - Hyperexponential Service*

μ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
11	4.86515668	3.72736456	12.8474069
13	4.18208704	3.21276505	10.8735917
15	3.76653401	2.89837776	9.67760920
17	3.48707744	2.68644585	8.87528324
19	3.18123908	2.53393179	8.29971350
21	3.13499689	2.41893523	7.86667776

From the Table 1-3, we have the following observations.

• While maximizing the retrial rate, the expected orbit size minimizes for different arrangements of service and arrival times.

• While comparing to Erlang and Exponential arrival times, the expected orbit size decreases more rapidly in the case of hyperexponential arrival time. Similarly, the expected orbit size decreases slowly in the case of Erlang arrival time.

Illustrative Example 2:

In the following tables, we examine the impact of the service rate γ on the expected orbit size. Fix $\lambda = 1; \mu = 8; \mathfrak{p}_1 = 0.60; \eta = 2; \mathfrak{p}_2 = 0.40; \sigma = 1; \mathfrak{q}_1 = 0.40; \delta = 2; \mathfrak{q}_2 = 0.60$.

Table 4: *Expected orbit size - Exponential Service*

γ	ARRIVAL		
	EXPONENTIAL	**ERLANG**	**HYPEREXPONENTIAL**
7	8.62406523	6.38645821	24.9106128
8	5.34880881	3.98252157	14.9745343
9	3.99796043	2.98468674	10.9073389
10	3.26376783	2.44125073	8.70686572
11	2.80380376	2.10058231	7.33376198
12	2.48925463	1.86760598	6.39840730
13	2.26091389	1.69851986	5.72205463

Table 5: *Expected orbit size - Erlang Service*

γ	ARRIVAL		
	EXPONENTIAL	**ERLANG**	**HYPEREXPONENTIAL**
7	10.0886405	7.41091127	29.7073972
8	5.78766750	4.28690000	16.4457714
9	4.20009203	3.12293216	11.6004235
10	3.37742675	2.51789282	9.10389213
11	2.87562366	2.14837061	7.58839020
12	2.53827076	1.89982294	6.57424954
13	2.29625267	1.72148762	5.85002618

Table 6: *Expected orbit size - Hyperexponential Service*

γ	ARRIVAL		
	EXPONENTIAL	**ERLANG**	**HYPEREXPONENTIAL**
7	4.97313477	3.77351943	13.3210642
8	3.82708589	2.90365725	10.0520703
9	3.17297465	2.40685867	8.19104494
10	2.75103836	2.08653058	6.99299693
11	2.45682845	1.86336640	6.15918869
12	2.24027528	1.69928391	5.54661758
13	2.07439882	1.57374500	5.07829007

From the Table 4-6, we have the following observations.

• While maximizing the service rate, the expected orbit size minimizes for various possible arrangements of arrival and service times.

• While comparing to Erlang and Exponential arrival times, expected orbit size decreases more rapidly forhyperexponential arrival time. Similarly, the expected orbit size reduces gradually for Erlang arrival time.

Illustrative Example 3:

In the following tables, we examine the impact of the vacation rate η on expected orbit size. Fix $\lambda = 1; \gamma = 6; \mathfrak{p}_1 = 0.60; \mu = 8; \mathfrak{p}_2 = 0.40; \sigma = 1; \mathfrak{q}_1 = 0.40; \delta = 2; \mathfrak{q}_2 = 0.60$.

Table 7: *Expected orbit size - Exponential Service*

η	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
4	3.37760655	2.48794146	9.31560352
6	2.38366783	1.76179679	6.27202950
8	2.04477644	1.51618610	5.22962692
10	1.87519953	1.39392122	4.70752570
12	1.77370727	1.32100067	4.39510928
14	1.70624181	1.27264837	4.18755323
16	1.65818186	1.23826766	4.03979561

Table 8: *Expected orbit size - Erlang Service*

η	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
4	3.76004704	2.74496640	10.6614094
6	2.59592138	1.90098311	7.03743508
8	2.20977550	1.62305018	5.82862515
10	2.01854731	1.48608798	5.22879429
12	1.90472438	1.40484025	4.87154080
14	1.82932015	1.35114669	4.63485557
16	1.77572971	1.31305560	4.46666909

Table 9: *Expected orbit size - Hyperexponential Service*

η	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
4	2.20929153	1.67412072	5.46231400
6	1.65814092	1.26336768	3.87581037
8	1.45463836	1.11330793	3.29231358
10	1.34958578	1.03632469	2.99270300
12	1.28565431	0.989662976	2.81122828
14	1.24271322	0.958408469	2.68981153
16	1.21190604	0.936030889	2.60298246

From the Table 7-9, we have the following observations.

• While maximizing the vacation rate, the expected orbit size minimizes for various possible arrangements of arrival and service times.

• While comparing to Erlang and Exponential arrival times, the expected orbit size decreases more rapidly for hyperexponential arrival time. Similarly, the expected orbit size reduces gradually for Erlang arrival time.

Illustrative Example 4:

In the following tables, we examine the impact of the repair rate δ against the expected orbit size. Fix $\lambda = 1; \gamma = 6; \mathfrak{p}_1 = 0.60; \eta = 2; \mathfrak{p}_2 = 0.40; \sigma = 1; \mathfrak{q}_1 = 0.40; \mu = 8; \mathfrak{q}_2 = 0.60$.

Table 10: *Expected orbit size - Exponential Service*

δ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
4	7.59759409	5.51219114	22.7455157
6	5.87750594	4.25241592	17.5011470
8	5.24414894	3.78930406	15.5590737
10	4.91583865	3.54965983	14.5484770
12	4.71517696	3.40339073	13.9291648
14	4.57989976	3.30488730	13.5108519
16	4.48255049	3.23406112	13.2093878

Table 11: *Expected orbit size - Erlang Service*

δ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
4	9.01354432	6.47340005	27.6296447
6	6.68529225	4.79331402	20.3620356
8	5.87341675	4.20725586	17.8186074
10	5.46167788	3.91034115	16.5247530
12	5.21302166	3.73120654	15.7416246
14	5.04665662	3.61145470	15.2167986
16	4.92756297	3.52578803	14.8406217

Table 12: *Expected orbit size - Hyperexponential Service*

δ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
4	4.31052599	3.21966516	11.8178685
6	3.70054138	2.75276167	10.1062609
8	3.44417977	2.55760891	9.37940438
10	3.30337635	2.45078331	8.97785988
12	3.21444535	2.38346547	8.72330076
14	3.15320451	2.33718412	8.54755003
16	3.10847255	2.30342084	8.41893285

From the Table 10-12, we have the following observations.

- While maximizing the repair rate, the expected orbit size minimizes for various possible arrangements of arrival and service times.
- While comparing to Erlang and Exponential arrival times, the expected orbit size decreases more rapidly in the case hyperexponential arrival time. Similarly, the expected orbit size decreases gradually for Erlang arrival time.

Illustrative Example 5:

In the following tables, we examine the impact of the breakdown rate σ against expected orbit size. Fix $\lambda = 1; \gamma = 6; \mathfrak{p}_1 = 0.60; \eta = 2; \mathfrak{p}_2 = 0.40; \delta = 2; \mathfrak{q}_1 = 0.40; \mu = 8; \mathfrak{q}_2 = 0.60$.

Table 13: *Expected orbit size - Exponential Service*

σ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
0.2	4.08040357	2.96155493	11.8191902
0.3	4.73131254	3.44689911	13.7479083
0.4	5.54494873	4.05191775	16.1683587
0.5	6.59105242	4.82715640	19.2937651
0.6	7.98585710	5.85628764	23.4817132
0.7	9.93858363	7.28856770	29.3811856
0.8	12.8676733	9.41875439	38.3052798

Table 14: *Expected orbit size - Erlang Service*

σ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
0.2	4.08450150	2.94457362	11.9867048
0.3	4.80393168	3.47566946	14.1600116
0.4	5.73618355	4.16175105	16.9879714
0.5	6.99170138	5.08201017	20.8145996
0.6	8.77333791	6.38066758	26.2765620
0.7	11.4985662	8.35096411	34.6983724
0.8	16.1853911	11.6946701	49.3649830

Table 15: *Expected orbit size - Hyperexponential Service*

σ	ARRIVAL		
	EXPONENTIAL	ERLANG	HYPEREXPONENTIAL
0.2	4.02198460	3.02084111	10.8591758
0.3	4.32444739	3.25474173	11.6877493
0.4	4.65258139	3.50785258	12.5914617
0.5	5.01076886	3.78348382	13.5827069
0.6	5.40425830	4.08558108	14.6764439
0.7	5.83942470	4.41891647	15.8909556
0.8	6.32412046	4.78934454	17.2488808

From the Table 13-15, we have the following observations.

- While maximizing the breakdown rate, the mean orbit size also maximizes for all possible arrangements of arrival and service times.
- While comparing to exponential and hyperexponential service times, the mean orbit size increases more rapidly for Erlang service time. Similarly, the mean orbit size increases slowly for hyperexponential service time.

Illustrative Example 6:

In the following figures, we observe the impact of both the retrial rate μ and vacation rate η against expected system size. Fix $\lambda = 1; \mathfrak{p}_1 = 0.60; \gamma = 6; \mathfrak{p}_2 = 0.40; \sigma = 1; \mathfrak{q}_1 = 0.40; \delta = 2; \mathfrak{q}_2 = 0.60$.

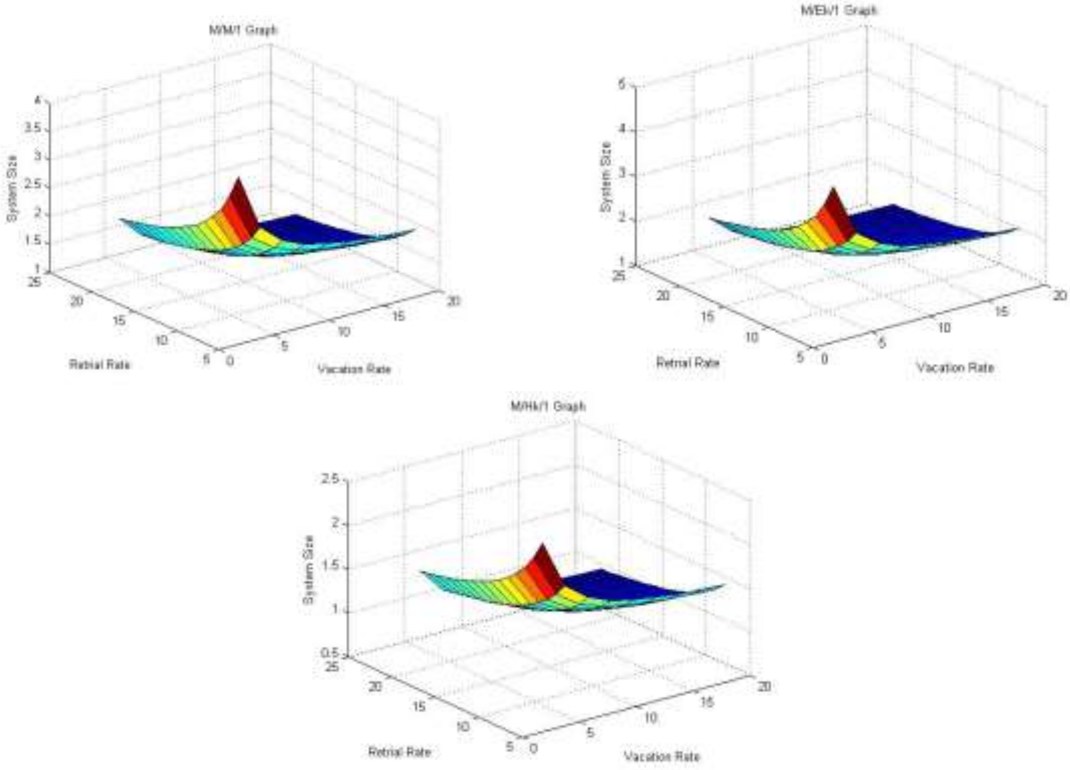

Figure 1: *Exponential arrival*

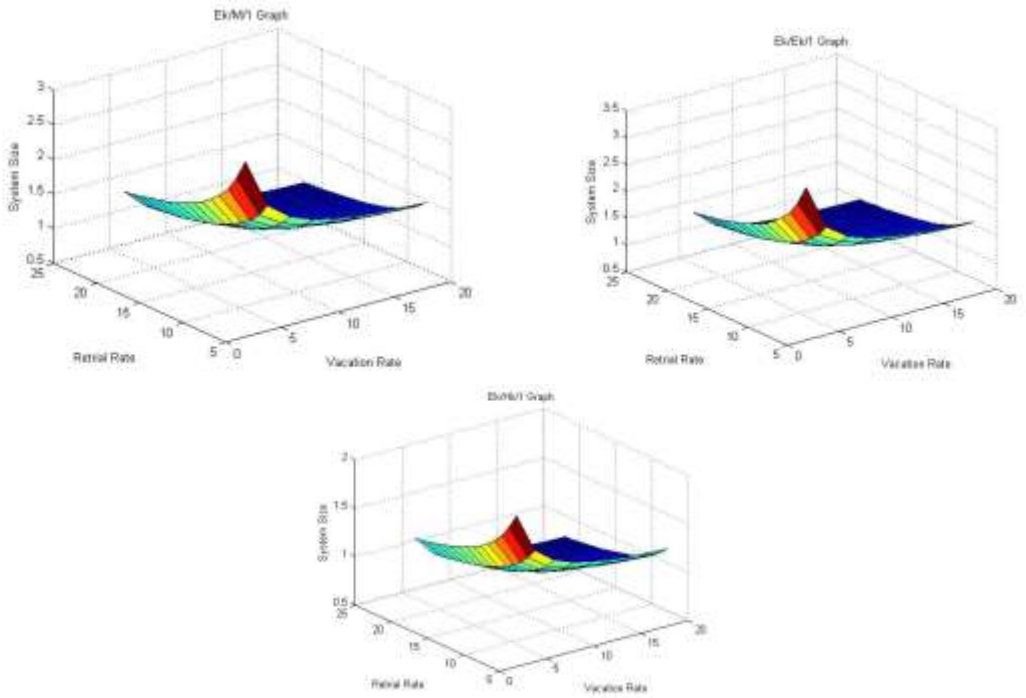

Figure 2: *Erlang arrival*

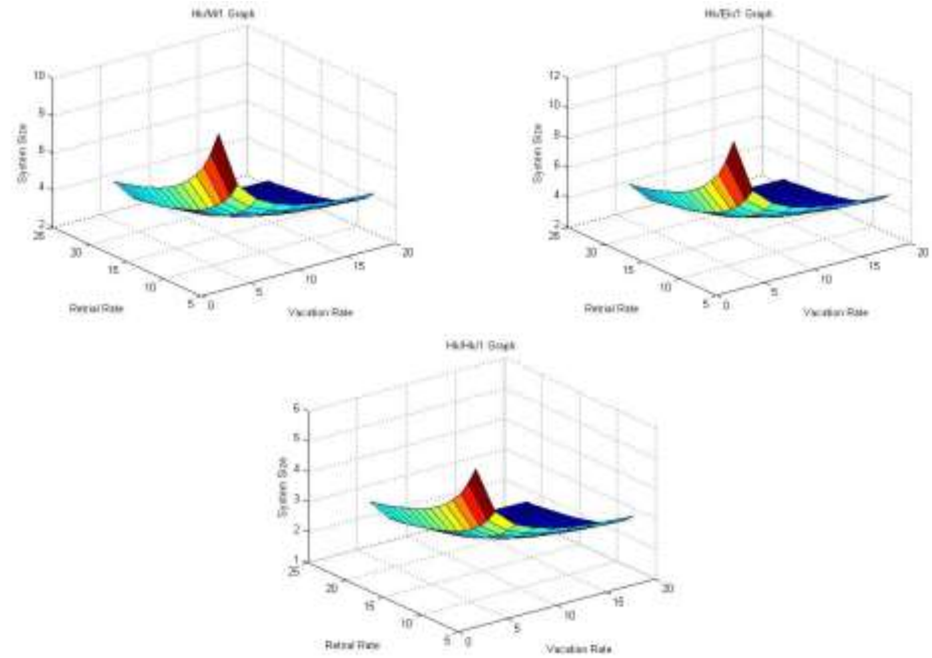

Figure 3: *Hyperexponential arrival*

A quick observation of Figure 1-3 discloses the fact that expected system size decreases while maximizing both the retrial rate and vacation rate for all possible arrangements of arrival and service times.

Illustrative Example 7:

In the following figures, we analyse the influence of both the repair rate δ and the service rate γ against the probability of server being idle.

We fix $\lambda = 1; \mathfrak{p}_1 = 0.60; \eta = 2; \mathfrak{p}_2 = 0.40; \mu = 8; \mathfrak{q}_1 = 0.40; \sigma = 1\mathfrak{q}_2 = 0.60$.

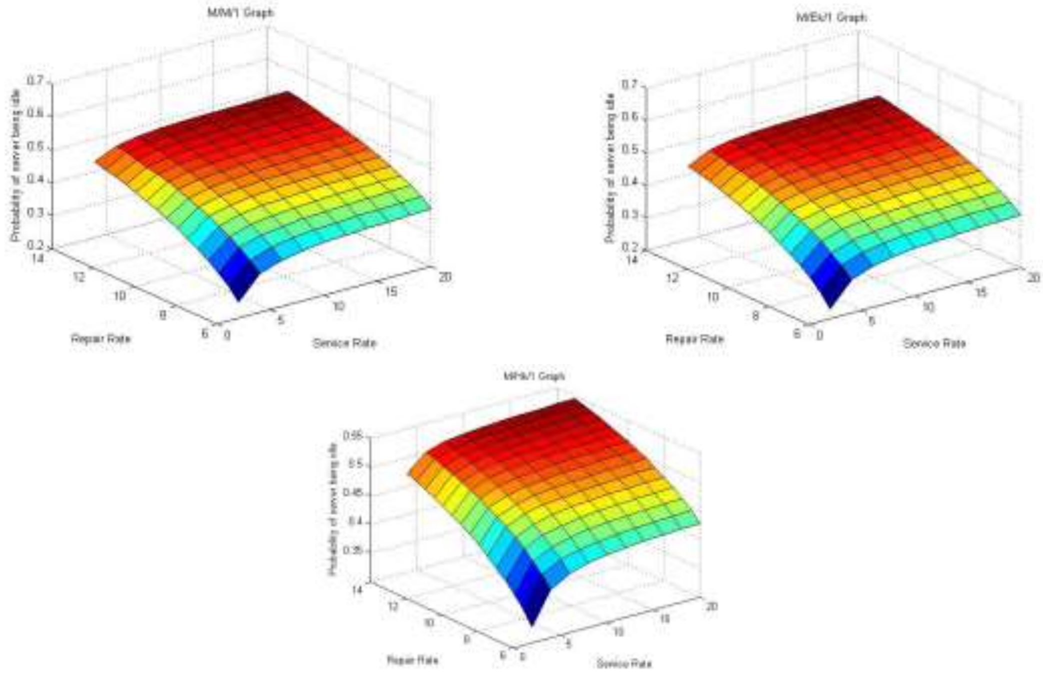

Figure 4: *Exponential arrival*

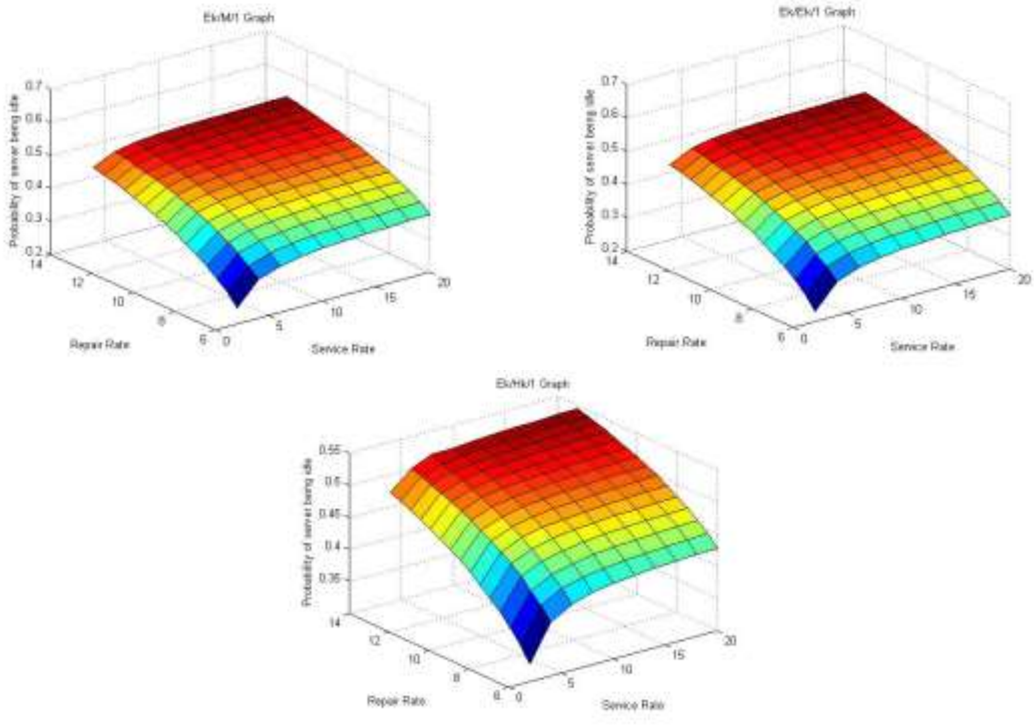

Figure 5: *Erlang arrival*

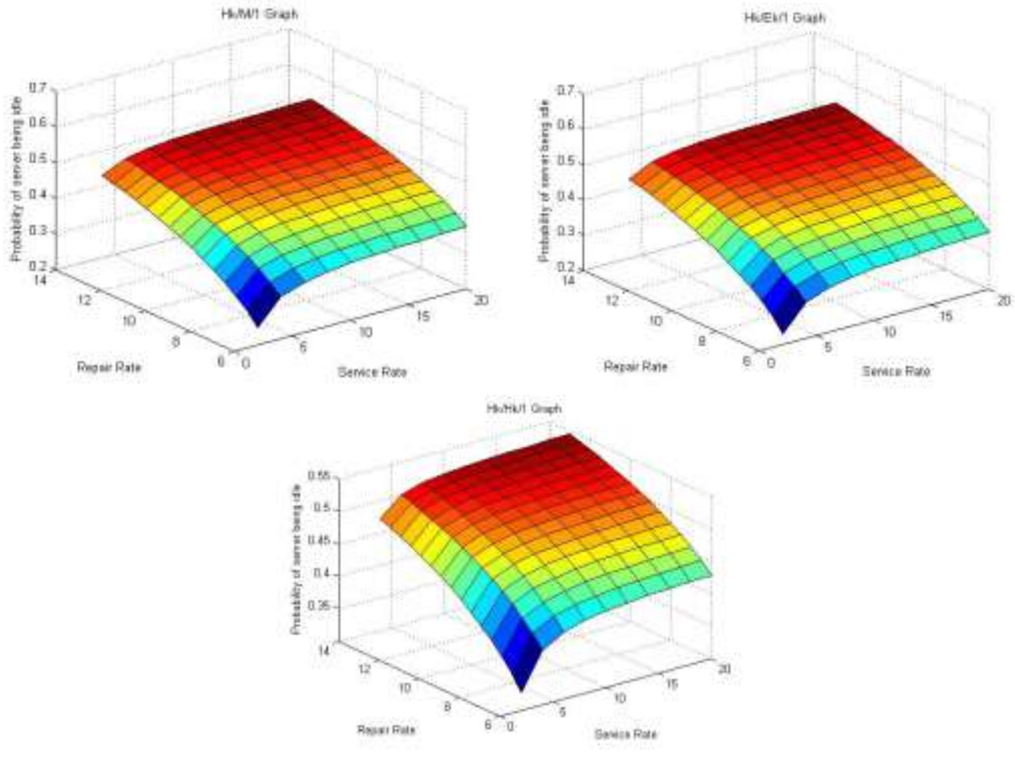

Figure 6: *Hyperexponential arrival*

A quick view of Figure 4-6 reveals the fact that the probability of server being idle increases while increasing both the repair rate and the service rate for all possible arrangements of arrival and service times.

VIII Conclusion

In our work, we have discussed about the retrial queuing system in which arrival follows MAP and service time follows PH-distribution together with Bernoulli schedule vacation, Bernoulli feedback, breakdown and repair. The effect of retrial rate, repair rate, vacation rate, breakdown rate and service rate on the average orbit size has been analysed through numerical values. Also, the influence of both the repair rate and the service rate against the probability of server being idle and the effect of retrial rate and the vacation rate on the average system size have been clearly visualized with the support of graphical representations. We have also analysed the busy period for our model. Our work can be extended to queueing models in which arrival follows BMAP.

References

[1] Aissani, A. and Artalejo, J.R. (1998). On the single server retrial queue subject to breakdowns. *Queueing Systems*, 30:309–321.

[2] Artalejo, J.R., Comez-Corral, A. and Neuts, M.F. (2000). Analysis of multiserver queues with constant retrial rate. *European Journal of Operational Research*, 135:569–581.

[3] Badamchi Zadeh, A. and Shankar, G.H. (2008). A two phase queue system with Bernoulli feedback and Bernoulli schedule server vacation. *Information and Management Science*, 19-2:329–338.

[4] Chakravarthy, S.R. (2009). Analysis of multi-server queue with Markovian arrivals and synchronous phase type vacations. *Asia-Pacific Journal of Operational Research*, 26-1:85–113.

[5] Chakravarthy, S.R. (2010). Markovian arrival processes. *Wiley Encyclopaedia of Operations Research and Management Science*.

[6] Cohen, J.W. (1957). Basic problems of telephone traffic theory and the influence of repeated calls. *Philips Telecommunication Review*, 18-2:49-100.

[7] Dimitry Efrosin and Anastasia Winkler (2011). Queueing system with a constant retrial rate, non-reliable server and threshold-based recovery. *European Journal of Operational Research*, 210:594-605.

[8] Fayolle, G. (1986). A simple telephone exchange with delayed feedbacks. *International Seminar on Teletraffic Analysis and computer Performance Evaluation*, 245–253.

[9] Fu-Min Chang, Tzu-Hsin Liu, Jau-Chuan Ke (2018). On an unreliable server retrial queue with customer feedback and impatience. *Applied Mathematical Modelling*, 55:171–182.

[10] Gautam Choudhury and Madhuchanda Paul (2005). A two phase queueing system with Bernoulli feedback. *Information and Management Science*, 16-1:35-52.

[11] Gautam Choudhury and Mitali Deka (2018). A batch arrival unreliable server delaying repair queue with two phases of service and Bernoulli vacation under multiple vacation policy.. *Quality Technology and Quantitative Management*,15-2:157–186.

[12] Ioannis Dimitriou and Christos Langaris (2010). A repairable queueing model with two-phase service, start-up times and retrial customers. *Computers and Operations Research*, 37:1181–1190.

[13] Jeffrey J. Hunter (1983). Filtering of Markov Renewal Queues, I: Feedback Queues. *Advances in Applied Probability*, 15-2:349-375.

[14] Jin-ting Wang and Peng Zhang (2009). A single-server discrete-time retrial G-queue with server breakdowns and repairs. *Acta Mathematicae Applicatae Sinica, English Series*, 25-4:675–684.

[15] Keilson, J. and Servi, L.D. (1986). Oscillating random walks doles for GI/G/1 vacation system with Bernoulli schedules. *Journal of Applied Probability*, 23-3:790–802.

[16] Krishna Kumar, B., Vijayalakshmi, G., Krishnamoorthy, A. and Sadiq Basha, S. (2010). A single server feedback retrial queue with collisions. *Computers and Operations Research*, 37:1247–1255.

[17] Kulkarni, V.G. and Choi, B.D. (1990). Retrial queues with server subject to breakdowns and repairs. *Queueing Systems*, 7:191–208.

[18] Latouche, G. and Ramaswami, V. (1999). Introduction to Matrix Analytic Methods in Stochastic Modelling. *American Statistical Association, SIAM*, Alexandria, Philadelphia.

[19] Lucantoni, D, Meier-Hellstern, K.S. and Neuts, M.F. (1990). A single- server queue with server vacations and a class of nonrenewal arrival processes. *Advanced Applied Probability*, 22:676–705.

[20] Nawel Gharbi and Claude Dutheillet (2011). An algorithmic approach for analysis of finite-source retrial systems with unreliable server. *Computers and Mathematics with Applications*, 62:2535–2546.

[21] Neuts, M.F. (1981). Matrix-Geometric Solutions in Stochastic Models - An Algorithmic Approach. *The Johns Hopkins University Press*, Baltimore and London.

[22] Neuts, M.F. (1979). A versatile markovian point process. *Journal of Applied Probability*, 16:764–779.

[23] Peishu Chen, Yongwu Zhou, Changawen Li (2015). Discrete-time retrial queue with Bernoulli vacation, preemptive resume and feedback customers. *Journal of Industrial Engineering and Management*, 8-4:1236–1250.

[24] Quan-Lin Li, Yu Ying, Yiqiang Q.Zhao (2006). A BMAP/G/1 retrial queue with a server subject to breakdowns and repairs. *Annals of Operations Research*, 141-1:233–270.

[25] Ralph L. Disney Donald C. McNickle Burton Simon (1980). The M/G/1 queue with instantaneous bernoulli feedback. *Naval Research Logistics: A Journal dedicated to Advances in Operations and Logistics Research*, 27-4:635–644.

[26] Robert D. Foley and Ralph L. Disney (1983). Queues with Delayed Feedback. *Advances in Applied Probability*, 15-1:162–182.

[27] Servi, L.D. (1986) Average delay approximation of M/G/1 cyclic service queues with Bernoulli schedules. *IEEE Journal on Selected Ares in Communications*, 4-6:813–822.

[28] Takács, L. (1963). A single-server queue with feedback. *The Bell system Technical Journal*, 42-2:505–519.