

Finite Priority Queueing System with Service Interruption

Richa Sharma^{1*} and Madhu Jain²

¹Assistant Professor, Department of Mathematics, JKLU, Jaipur 302026, Rajasthan, India

²Associate Professor, Department of Mathematics, IIT, Roorkee 247667, Uttarakhand, India

(*Corresponding Author) E-mail-id: *aligarh.richa@gmail.com/richasharma@jkl.edu.in,

²madhufma@iitr.ernet.in/drmadhuja@yahoo.co.in

Abstract: In this paper, we consider a finite buffer queueing model with preemptive resume priority and service interruption. The system may fail anytime while providing service to the priority and non-priority customers. The arrival patterns of two types of the customers are independent Poisson processes wherein the service times are exponentially distributed. The service discipline of the customers is first come first served within their own queue. Further, Runge–Kutta (R–K) method of fourth order is used to obtain the solution of the system of transient equations governing the model. Various performance measures are obtained in terms of transient probabilities such as expected number of the customers in the system, throughput, expected delay, probability of the server being idle, busy or broken down and others. The numerical results are obtained by taking an illustration to examine the effects of various system parameters on the performance measures and total expected cost.

Keywords: Finite queue, Preemptive priority, Service interruption, Queue length, Throughput, Runge–Kutta method

1. INTRODUCTION

Priority queueing systems have numerous applications in many real life situations encountered in computer networks, communication systems, transportation, distribution systems and many others. There are two broad categories of priority situations namely (a) preemptive priority and (b) non-preemptive priority. In preemptive priority model, the service of a lower priority customer is interrupted to provide service to higher priority customer, whereas in non-preemptive priority model, the priority customer is taken for service on the completion of ongoing service of the customer. The present investigation deals with a single-server queueing system with two classes of customers, namely the class one of higher priority customers and class two of lower priority customers. The higher priority customers have preemptive resume priority over the lower priority customers, that is the service of lower priority customer is interrupted on arrival of higher priority customer and resumes from the same point where it is interrupted when the service of all higher priority customers are completed.

Priority queues have received considerable attention in the literature due to their wide applications in the numerous fields. For real-life example, we refer the doctor's clinic wherein the precedence is given to that patient who requires emergency treatment than to that patient who comes for a normal treatment. The notable contributions in the related area by prominent researchers are given as follows. A priority M/D/k queue with finite buffers was analysed by Bose and Pal [1]. A two-class single-server preemptive priority queueing system was examined by Tarabia [2]. Feng and Umemura [3] considered a finite buffer queueing model with two servers and two non-preemptive priority service classes. Subramanian *et al.* [4] examined a single-server M/M/1 queueing system with preemptive priority service. Further Kim [5] proposed a new priority discipline called the (N, n) -preemptive priority discipline according to which, the preemption of the service of a low-class customer is determined by two thresholds N and n of the queue length of high-class customers. A two-stage multi-server priority tandem queue with two types of processed customers has been considered by Kim and Dudin [6]. Moreover, Kim [7] proposed a new priority discipline called the T-preemptive discipline for M/G/1 queue.

Over the past few decades, single-server queueing systems with service interruption have been studied by many researchers in the context of a variety of real life congestion scenarios including telecommunication and computer networks, service and distribution centres, manufacturing and production systems and others. The service interruption due to server breakdown is quite common phenomena which influences the performance measures of any queueing system. The study of queues with service interruptions and repairs is not only important with theoretic point of view but also has enormous engineering applications. Ke () [8] focused attention on bi-level control policy for unreliable server queue. Sherman and Kharoufeh [9] analysed an unreliable M/M/1 retrial queue with infinite capacity orbit and normal queue. The Markovian single-server queue with breakdowns and repairs was studied by Economou and Kanta [10]. Economou and Kapodistria [11] considered a single-server unreliable queue represented by a two-dimensional continuous-time Markov chain. Recently, Jain *et al.* [12] dealt with vacation queue and service interruption wherein the customers arrive at the service station according to Poisson process with rates dependent upon the status of the server. Further, an unreliable M/M/2/K queueing system under N- and F-policy with multi-optional phase repair was examined by Jain *et al.* [13].

Priority queue with combination of unreliable server has also been studied by a few researchers in different frameworks. The first published paper on preemptive priority discipline with unreliable server was due to White and Christie [14]. They studied the preemptive priority system with two priority classes by using the method of generating function to obtain the steady state distribution. A single-server queue with two classes of customers was considered by Towsley and Tripathi [15]. Further, Jain [16] studied a single-server priority queue with service interruption due to failure of the server. The analysis of unreliable server queue with two class non-preemptive priority subscribers was examined by Jain and Bhargava [17]. A preemptive resume is followed by the server who is subjected to random failures. A single-server retrial queue with priority subscribers and server breakdown was studied by Wang [18]. A single-server non-preemptive priority queue subjected to server interruptions was examined by Kamoun [19]. Recently, Dimitriou [20] considered an unreliable single-server queue with preemptive resume priorities and vacations.

In recent past years, a variety of numerical methods was developed to solve many sophisticated mathematical models by numerical approximation. Runge-Kutta (R-K) method is a useful technique for solving the set of differential equations. To increase accuracy and to reduce

calculation time, we can apply a fourth order R–K method, as it is stable at large time steps, accurate and relatively fast. R–K method has been used to provide transient solution of queueing problems in different frameworks by many researchers. Van As [21] dealt with transient system behaviour of the queueing system using the fourth order R–K method. R–K method was used by Jain *et al.* [22, 23] to find the transient probabilities of various states by solving the set of governing differential equations of queueing models in different frameworks. Maheshwari *et al.* [24] developed the queueing model for the unreliable flexible manufacturing cell which operates under stochastic environment. Using R–K method, the probabilities for different system states have been evaluated. The transient behaviour of multi-server queue with server vacations has been examined by Jain and Upadhyaya [25] using numerical approach based on R–K method of the fourth order.

In this investigation, we study two-class preemptive priority queueing systems with unreliable server. The service of the priority and non-priority customers may be interrupted because of the failure of the server. R–K method of the fourth order is used to obtain various performance measures in terms of transient probabilities of the system states. The rest of the paper is organised as follows. Section 2 is devoted to the description of the model. The governing equations of the model for both finite capacity and finite population are given in Section 3. Various performance measures are discussed in Section 4. In Section 5, cost analysis has been discussed. The sensitivity analysis is carried out in Section 6. Finally, the conclusions are drawn in Section 7. The future research directions have also been discussed.

2. MODEL DESCRIPTION

We develop M/M/1 preemptive priority queueing model wherein the service of both priority and non-priority customers may be interrupted because of the server breakdown. To describe the Markov model of two-class preemptive priority queueing system with service interruption, the following assumptions are made:

- We consider the two types of customers, namely type-1 (higher priority) and type-2 (lower priority) customers who arrive according to Poisson fashion in the system.
- A single server serves the customer under preemptive resume priority discipline. The arrival of higher priority customer interrupts the service of a lower priority customer who on its re-entry resumes service from the preempted point. The service times follow exponential distribution with parameter $\mu_1(\mu_2)$ for higher (lower) priority customer.
- Both types of customers arrive with state-dependent arrival rates. Higher priority customers arrive with rate λ_1 and $\lambda_{1,i} = (L - i)\lambda_1$, $0 \leq i \leq L$ for finite capacity and finite population models, respectively. On the other hand, lower priority customers arrive with rate λ_2 and $\lambda_{2,j} = (K - j)\lambda_2$, $0 \leq j \leq K$ for finite capacity and finite population models, respectively. Here, $L(K)$ denotes the capacity/population size of type-1 (type-2) customers for finite capacity/finite population model.
- The queue discipline is first come first served within their own queue. During the service of lower priority customer, if higher priority customer joins the system, then the service of the lower priority customer is interrupted. The service of lower priority customer restarts again when there is no higher priority customer present in the system. The server preempts the lower priority customer whose service is going on and higher priority customer joins the system. This discipline is known as *preemptive priority*.

- The server may experience random breakdowns; it is sent for repairing as soon as failure occurs. The server will be fully recovered after completion of the repair. After repairing, the server restarts service to the customer whose service was interrupted because of the failure of the server.
- The life time and repair time of the server are governed by the exponential distributions with parameter α and β , respectively.

To formulate the mathematical model in terms of transient state probabilities, the following notations are used:

$P_{i,j,1}(t)$ – Probability that there are $i(0 \leq i \leq L)$ priority customers and $j(0 \leq j \leq K)$ non-priority customers present in the system when the server is in busy state at time t .

$P_{i,j,2}(t)$ – Probability that there are $i(0 \leq i \leq L)$ priority customers and $j(0 \leq j \leq K)$ non-priority customers present in the system when the server is in broken-down state at time t .

3. GOVERNING EQUATIONS

In this section, we provide the differential difference governing equations for both, that is finite capacity and finite population models as given in the following subsections.

3.1. For Finite Capacity Model

The differential difference equations governing the finite capacity model are constructed by taking appropriate transition rates (see Figure 1) as follows:

$$\frac{dP_{0,0,1}(t)}{dt} = -(\lambda_1 + \lambda_2 + \alpha)P_{0,0,1}(t) + \mu_1 P_{1,0,1}(t) + \mu_2 P_{0,1,1}(t) + \beta P_{0,0,2}(t) \quad (1)$$

$$\begin{aligned} \frac{dP_{0,j,1}(t)}{dt} = & -(\lambda_1 + \lambda_2 + \alpha + \mu_2)P_{0,j,1}(t) + \mu_1 P_{1,j,1}(t) + \mu_2 P_{0,j+1,1}(t) + \lambda_2 P_{0,j-1,1}(t) \\ & + \beta P_{0,j,2}(t), \quad 1 \leq j \leq K-1 \end{aligned} \quad (2)$$

$$\frac{dP_{0,K,1}(t)}{dt} = -(\lambda_1 + \alpha + \mu_2)P_{0,K,1}(t) + \mu_1 P_{1,K,1}(t) + \lambda_2 P_{0,K-1,1}(t) + \beta P_{0,K,2}(t) \quad (3)$$

$$\frac{dP_{0,0,2}(t)}{dt} = -(\lambda_1 + \lambda_2 + \beta)P_{0,0,2}(t) + \alpha P_{0,0,1}(t) \quad (4)$$

$$\frac{dP_{0,j,2}(t)}{dt} = -(\lambda_1 + \lambda_2 + \beta)P_{0,j,2}(t) + \alpha P_{0,j,1}(t) + \lambda_2 P_{0,j-1,2}(t), \quad 1 \leq j \leq K-1 \quad (5)$$

$$\frac{dP_{0,K,2}(t)}{dt} = -(\lambda_1 + \beta)P_{0,K,2}(t) + \alpha P_{0,K,1}(t) + \lambda_2 P_{0,K-1,2}(t), \quad 1 \leq j \leq K-1 \quad (6)$$

$$\begin{aligned} \frac{dP_{i,0,1}(t)}{dt} = & -(\lambda_1 + \lambda_2 + \alpha + \mu_1)P_{i,0,1}(t) + \mu_1 P_{i+1,0,1}(t) + \lambda_1 P_{i-1,0,1}(t) \\ & + \beta P_{i,0,2}(t), \quad 1 \leq i \leq L-1 \end{aligned} \quad (7)$$

$$\begin{aligned} \frac{dP_{i,j,1}(t)}{dt} = & -(\lambda_1 + \lambda_2 + \alpha + \mu_1)P_{i,j,1}(t) + \mu_1 P_{i+1,j,1}(t) + \lambda_1 P_{i-1,j,1}(t) + \lambda_2 P_{i,j-1,1}(t) \\ & + \beta P_{i,j,2}(t), \quad 1 \leq j \leq K-1, \quad 1 \leq i \leq L-1 \end{aligned} \quad (8)$$

$$\frac{dP_{i,K,1}(t)}{dt} = -(\lambda_1 + \alpha + \mu_1)P_{i,K,1}(t) + \mu_1 P_{i+1,K,1}(t) + \lambda_1 P_{i-1,K,1}(t) + \lambda_2 P_{i,K-1,1}(t) + \beta P_{i,K,2}(t), \quad 1 \leq i \leq L-1 \quad (9)$$

$$\frac{dP_{L,0,1}(t)}{dt} = -(\lambda_2 + \alpha + \mu_1)P_{L,0,1}(t) + \lambda_1 P_{L-1,0,1}(t) + \beta P_{L,0,2}(t) \quad (10)$$

$$\frac{dP_{L,j,1}(t)}{dt} = -(\lambda_2 + \alpha + \mu_1)P_{L,j,1}(t) + \lambda_1 P_{L-1,j,1}(t) + \lambda_2 P_{L,j-1,1}(t) + \beta P_{L,j,2}(t), \quad 1 \leq j \leq K-1 \quad (11)$$

$$\frac{dP_{L,K,1}(t)}{dt} = -(\alpha + \mu_1)P_{L,K,1}(t) + \lambda_1 P_{L-1,K,1}(t) + \lambda_2 P_{L,K-1,1}(t) + \beta P_{L,K,2}(t) \quad (12)$$

$$\frac{dP_{i,0,2}(t)}{dt} = -(\lambda_1 + \lambda_2 + \beta)P_{i,0,2}(t) + \alpha P_{i,0,1}(t) + \lambda_1 P_{i-1,0,2}(t), \quad 1 \leq i \leq L-1 \quad (13)$$

$$\frac{dP_{i,j,2}(t)}{dt} = -(\lambda_1 + \lambda_2 + \beta)P_{i,j,2}(t) + \alpha P_{i,j,1}(t) + \lambda_1 P_{i-1,j,2}(t) + \lambda_2 P_{i,j-1,2}(t), \quad 1 \leq j \leq K-1, \quad 1 \leq i \leq L-1 \quad (14)$$

$$\frac{dP_{i,K,2}(t)}{dt} = -(\lambda_1 + \beta)P_{i,K,2}(t) + \alpha P_{i,K,1}(t) + \lambda_1 P_{i-1,K,2}(t) + \lambda_2 P_{i,K-1,2}(t), \quad 1 \leq i \leq L-1 \quad (15)$$

$$\frac{dP_{L,0,2}(t)}{dt} = -(\lambda_2 + \beta)P_{L,0,2}(t) + \alpha P_{L,0,1}(t) + \lambda_1 P_{L-1,0,2}(t) \quad (16)$$

$$\frac{dP_{L,j,2}(t)}{dt} = -(\lambda_2 + \beta)P_{L,j,2}(t) + \alpha P_{L,j,1}(t) + \lambda_1 P_{L-1,j,2}(t) + \lambda_2 P_{L,j-1,2}(t), \quad 1 \leq j \leq K-1 \quad (17)$$

$$\frac{dP_{L,K,2}(t)}{dt} = -\beta P_{L,K,2}(t) + \alpha P_{L,K,1}(t) + \lambda_1 P_{L-1,K,2}(t) + \lambda_2 P_{L,K-1,2}(t) \quad (18)$$

3.2. For Finite Population Model

By taking appropriate transition rates shown in Figure 2, we construct the differential difference equations governing the model for finite population as follows:

$$\frac{dP_{0,0,1}(t)}{dt} = -(\lambda_{1,0} + \lambda_{2,0} + \alpha)P_{0,0,1}(t) + \mu_1 P_{1,0,1}(t) + \mu_2 P_{0,1,1}(t) + \beta P_{0,0,2}(t) \quad (19)$$

$$\frac{dP_{0,j,1}(t)}{dt} = -(\lambda_{1,j} + \lambda_{2,j} + \alpha + \mu_2)P_{0,j,1}(t) + \mu_1 P_{1,j,1}(t) + \mu_2 P_{0,j+1,1}(t) + \lambda_{2,j-1} P_{0,j-1,1}(t) + \beta P_{0,j,2}(t), \quad 1 \leq j \leq K-1 \quad (20)$$

$$\frac{dP_{0,K,1}(t)}{dt} = -(\lambda_{1,K} + \alpha + \mu_2)P_{0,K,1}(t) + \mu_1 P_{1,K,1}(t) + \lambda_{2,K-1} P_{0,K-1,1}(t) + \beta P_{0,K,2}(t) \quad (21)$$

$$\frac{dP_{0,0,2}(t)}{dt} = -(\lambda_{1,0} + \lambda_{2,0} + \beta)P_{0,0,2}(t) + \alpha P_{0,0,1}(t) \quad (22)$$

$$\frac{dP_{0,j,2}(t)}{dt} = -(\lambda_{1,j} + \lambda_{2,0} + \beta)P_{0,j,2}(t) + \alpha P_{0,j,1}(t) + \lambda_{2,j-1} P_{0,j-1,2}(t), \quad 1 \leq j \leq K-1 \quad (23)$$

$$\frac{dP_{0,K,2}(t)}{dt} = -(\lambda_{1,K} + \beta)P_{0,K,2}(t) + \alpha P_{0,K,1}(t) + \lambda_{2,K-1}P_{0,K-1,2}(t), \quad 1 \leq j \leq K-1 \quad (24)$$

$$\begin{aligned} \frac{dP_{i,0,1}(t)}{dt} = & -(\lambda_{1,i} + \lambda_{2,i} + \alpha + \mu_1)P_{i,0,1}(t) + \mu_1 P_{i+1,0,1}(t) + \lambda_{1,i-1}P_{i-1,0,1}(t) \\ & + \beta P_{i,0,2}(t), \quad 1 \leq i \leq L-1 \end{aligned} \quad (25)$$

$$\begin{aligned} \frac{dP_{i,j,1}(t)}{dt} = & -(\lambda_{1,i} + \lambda_{2,j} + \alpha + \mu_1)P_{i,j,1}(t) + \mu_1 P_{i+1,j,1}(t) + \lambda_{1,i-1}P_{i-1,j,1}(t) \\ & + \lambda_{2,j-1}P_{i,j-1,1}(t) + \beta P_{i,j,2}(t), \quad 1 \leq j \leq K-1, \quad 1 \leq i \leq L-1 \end{aligned} \quad (26)$$

$$\begin{aligned} \frac{dP_{i,K,1}(t)}{dt} = & -(\lambda_{1,i} + \alpha + \mu_1)P_{i,K,1}(t) + \mu_1 P_{i+1,K,1}(t) + \lambda_{1,i-1}P_{i-1,K,1}(t) \\ & + \lambda_{2,K-1}P_{i,K-1,1}(t) + \beta P_{i,K,2}(t), \quad 1 \leq i \leq L-1 \end{aligned} \quad (27)$$

$$\frac{dP_{L,0,1}(t)}{dt} = -(\lambda_{2,0} + \alpha + \mu_1)P_{L,0,1}(t) + \lambda_{1,L-1}P_{L-1,0,1}(t) + \beta P_{L,0,2}(t) \quad (28)$$

$$\begin{aligned} \frac{dP_{L,j,1}(t)}{dt} = & -(\lambda_{2,j} + \alpha + \mu_1)P_{L,j,1}(t) + \lambda_{1,L-1}P_{L-1,j,1}(t) + \lambda_{2,j-1}P_{L,j-1,1}(t) \\ & + \beta P_{L,j,2}(t), \quad 1 \leq j \leq K-1 \end{aligned} \quad (29)$$

$$\frac{dP_{L,K,1}(t)}{dt} = -(\alpha + \mu_1)P_{L,K,1}(t) + \lambda_{1,L-1}P_{L-1,K,1}(t) + \lambda_{2,K-1}P_{L,K-1,1}(t) + \beta P_{L,K,2}(t) \quad (30)$$

$$\begin{aligned} \frac{dP_{i,0,2}(t)}{dt} = & -(\lambda_{1,i} + \lambda_{2,0} + \beta)P_{i,0,2}(t) + \alpha P_{i,0,1}(t) \\ & + \lambda_{1,i-1}P_{i-1,0,2}(t), \quad 1 \leq i \leq L-1 \end{aligned} \quad (31)$$

$$\begin{aligned} \frac{dP_{i,j,2}(t)}{dt} = & -(\lambda_{1,i} + \lambda_{2,j} + \beta)P_{i,j,2}(t) + \alpha P_{i,j,1}(t) + \lambda_{1,i-1}P_{i-1,j,2}(t) \\ & + \lambda_{2,j-1}P_{i,j-1,2}(t), \quad 1 \leq j \leq K-1, \quad 1 \leq i \leq L-1 \end{aligned} \quad (32)$$

$$\begin{aligned} \frac{dP_{i,K,2}(t)}{dt} = & -(\lambda_{1,i} + \beta)P_{i,K,2}(t) + \alpha P_{i,K,1}(t) + \lambda_{1,i-1}P_{i-1,K,2}(t) \\ & + \lambda_{2,i}P_{i,K-1,2}(t), \quad 1 \leq i \leq L-1 \end{aligned} \quad (33)$$

$$\frac{dP_{L,0,2}(t)}{dt} = -(\lambda_{2,0} + \beta)P_{L,0,2}(t) + \alpha P_{L,0,1}(t) + \lambda_{1,L-1}P_{L-1,0,2}(t) \quad (34)$$

$$\begin{aligned} \frac{dP_{L,j,2}(t)}{dt} = & -(\lambda_{2,j} + \beta)P_{L,j,2}(t) + \alpha P_{L,j,1}(t) + \lambda_{1,L-1}P_{L-1,j,2}(t) \\ & + \lambda_{2,j-1}P_{L,j-1,2}(t), \quad 1 \leq j \leq K-1 \end{aligned} \quad (35)$$

$$\frac{dP_{L,K,2}(t)}{dt} = -\beta P_{L,K,2}(t) + \alpha P_{L,K,1}(t) + \lambda_{1,L-1}P_{L-1,K,2}(t) + \lambda_{2,K-1}P_{L,K-1,2}(t) \quad (36)$$

4. PERFORMANCE CHARACTERISTICS

In this section, we establish various performance indices in terms of transient probabilities. For finding these probabilities, the above-mentioned transient equations are solved using R-K method.

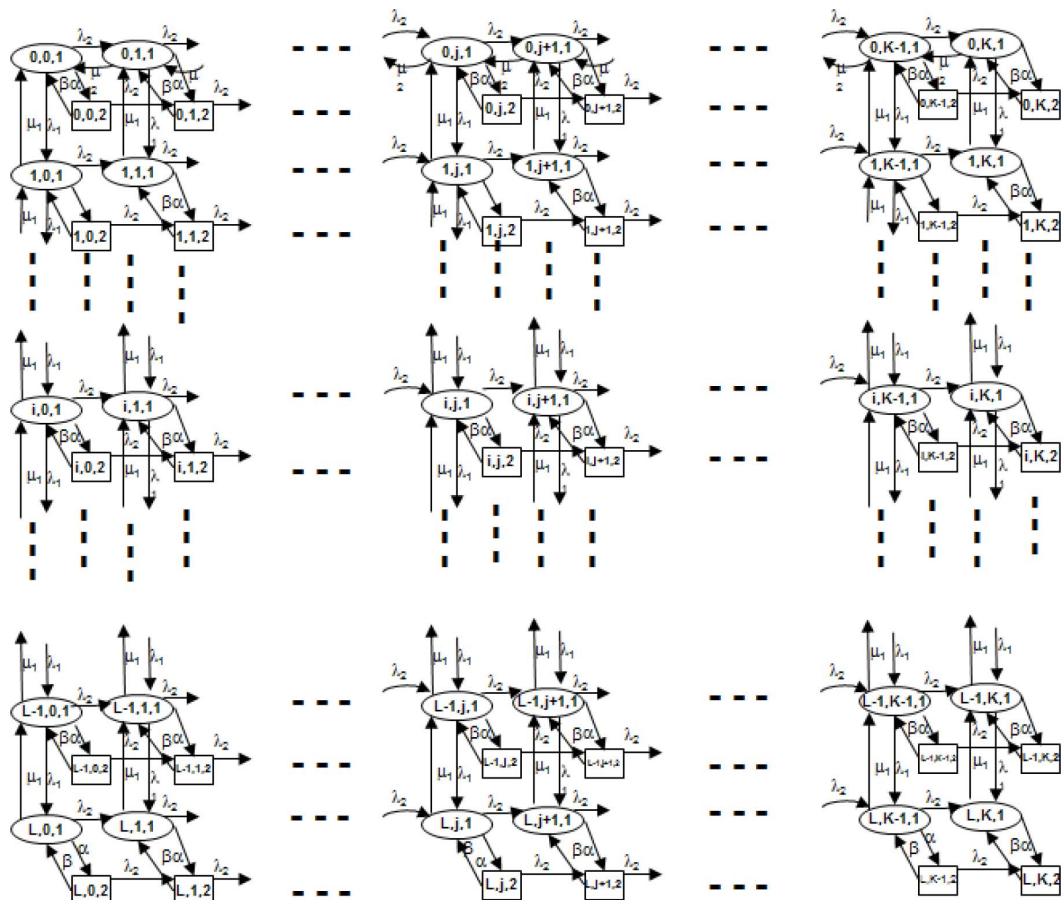


Figure 1. State transition diagram for finite capacity model.

R-K technique is implemented by developing programme in MATLAB software. After obtaining transient probabilities, some performance indices are calculated by using the formulae as given below:

- The probability of the servers being idle at time t is

$$P_I(t) = P_{0,0,1}(t) \quad (37)$$

- The probability of the server being busy at time t is

$$P_B(t) = \sum_{j=1}^K P_{0,j,1}(t) + \sum_{i=1}^L \sum_{j=0}^K P_{i,j,1}(t) \quad (38)$$

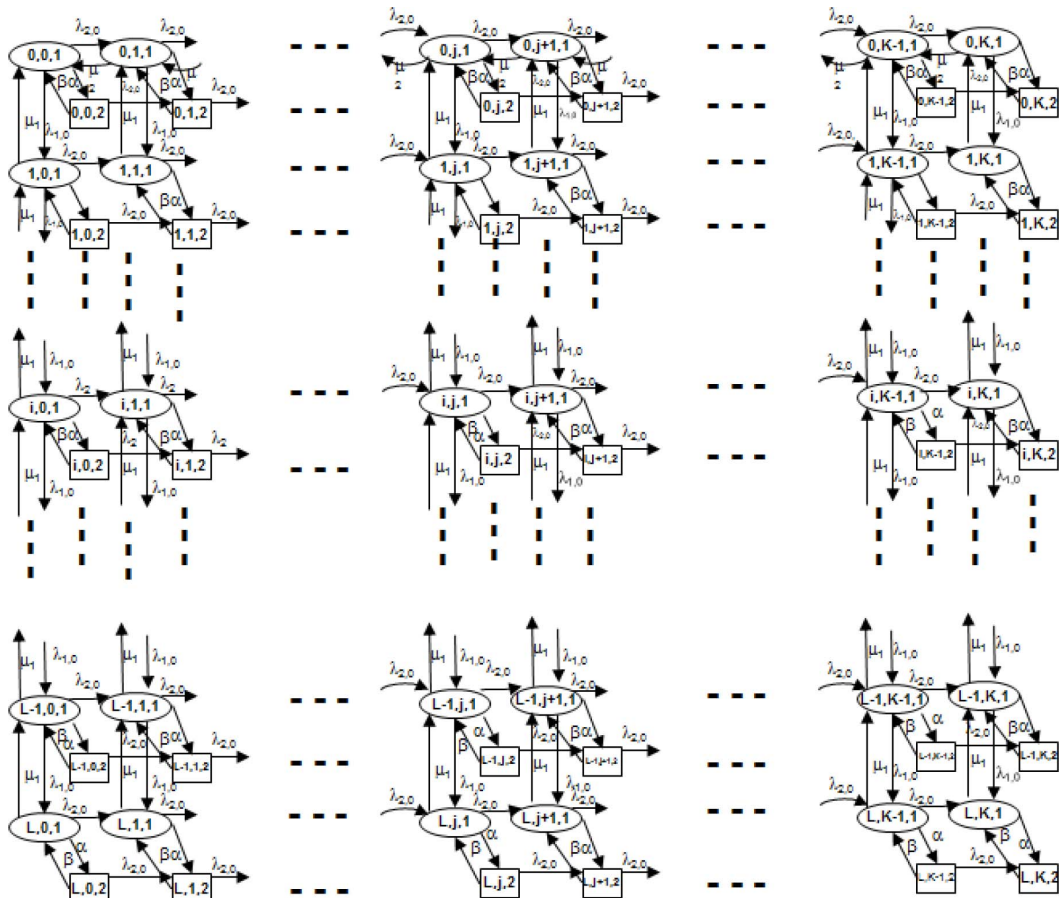


Figure 2. State transition diagram for finite population model.

- The probability of the server being broken down at time t is

$$P_D(t) = \sum_{i=0}^L \sum_{j=0}^K P_{i,j,2}(t) \quad (39)$$

- The expected number of the customers in the system at time t is given by

$$EN_S(t) = \sum_{i=0}^L \sum_{j=0}^K j \{P_{i,j,1}(t) + P_{i,j,2}(t)\} \quad (40)$$

- The throughput at time t is obtained using

$$TP(t) = \mu_1 \sum_{i=1}^L \sum_{j=0}^K P_{i,j,1}(t) + \mu_2 \sum_{j=1}^K P_{0,j,1}(t) \quad (41)$$

- The expected delay at time t is determined as

$$ED(t) = \frac{EN_S(t)}{TP(t)} \quad (42)$$

5. COST ANALYSIS

In this section, we construct the cost function in terms of appropriate performance measures and corresponding cost elements as follows:

The total cost $E\{C(t)\}$ is given by

$$E\{C(t)\} = C_H EN_S(t) + C_I P_I(t) + C_B P_B(t) + C_D P_D(t) \quad (43)$$

where C_H is the holding cost per unit time per customer present in the system, C_I is the cost per unit time for keeping the server idle, C_B is the cost per unit time for keeping the server busy and C_D is the repair cost per unit time of the broken down server.

6. SENSITIVITY ANALYSIS

In this section, we demonstrate the validation of the computational procedure on the basis of R–K method. To solve the set of transient state equations for the above model, the coding of the computer programme has been done with the help of software ‘MATLAB’ using ‘ode 45’ function for computing the transient probabilities. The numerical results are obtained for various performance indices, namely the probability of the server being idle, busy, broken down, expected number of the customers in the system and expected delay at time ‘ t ’ for both finite capacity and finite population models. In Tables 2–5, we summarise the numerical results for the probability of the server being idle, busy, broken down for both finite capacity and finite population models. Moreover, the results for the expected number of customers $EN_S(t)$, expected delay $ED(t)$ and throughput $TP(t)$ with the variation of various system parameters have been displayed in Figures 3(a)–14(b) for both finite capacity and finite population models. For the computation purpose, we fix the default parameters as $L = 5$, $K = 5$, $\lambda_1 = 0.3$, $\lambda_2 = 0.1$, $\mu_1 = 1$, $\mu_2 = 1$, $\alpha_2 = 0.01$, $\beta = 0.1$ and $\lambda_1 = 0.2$, $\lambda_2 = 0.1$, $\mu_1 = 1$, $\mu_2 = 1$, $\alpha = 0.1$, $\beta = 1$. We consider three different sets of cost elements as shown in Table 1 for making the cost analysis.

The effect on total expected cost is shown in Figures 15(a) and 16(d) for finite capacity model by varying arrival rate and failure rate.

Table 1. The different cost sets

Cost Sets	C_H	C_I	C_B	C_D
Set 1	\$20	\$50	\$50	\$25
Set 2	\$40	\$10	\$50	\$50
Set 3	\$800	\$250	\$400	\$400

Table 2. Effect of arrival rate (λ_x) and service rate (μ_y) on various performance measures for finite capacity model

λ_1	t	$P_I(t)$	$P_B(t)$	$P_D(t)$	μ_2	t	$P_I(t)$	$P_B(t)$	$P_D(t)$
0.1	1.0	0.872	0.118	0.009	0.5	1.0	0.723	0.268	0.005
	1.3	0.852	0.135	0.012		1.3	0.665	0.326	0.007
	1.6	0.837	0.148	0.014		1.6	0.607	0.386	0.008
0.3	1.0	0.765	0.226	0.010	0.9	1.0	0.756	0.235	0.001
	1.3	0.731	0.262	0.013		1.3	0.716	0.277	0.006
	1.6	0.702	0.296	0.015		1.6	0.680	0.317	0.007

Table 3. Effect of failure rate (α) and repair rate (β) on various performance measures for finite capacity model

α	t	$P_I(t)$	$P_B(t)$	$P_D(t)$	β	t	$P_I(t)$	$P_B(t)$	$P_D(t)$
0.09	1.0	0.700	0.219	0.079	0.5	1.0	0.757	0.235	0.007
	1.3	0.648	0.252	0.098		1.3	0.716	0.277	0.005
	1.6	0.601	0.283	0.114		1.6	0.680	0.386	0.004
0.1	1.0	0.694	0.222	0.088	0.9	1.0	0.757	0.231	0.006
	1.3	0.640	0.255	0.109		1.3	0.717	0.272	0.005
	1.6	0.592	0.279	0.128		1.6	0.681	0.318	0.001

Table 4. Effect of arrival rate (λ_1) and service rate (μ_2) on various performance measures for finite population model

λ_1	t	$P_I(t)$	$P_B(t)$	$P_D(t)$	μ_2	t	$P_I(t)$	$P_B(t)$	$P_D(t)$
0.1	0.6	0.608	0.354	0.048	1.1	0.6	0.487	0.495	0.049
	0.8	0.544	0.410	0.064		0.8	0.398	0.598	0.058
	1.0	0.494	0.449	0.083		1.0	0.310	0.714	0.053
0.2	0.6	0.485	0.495	0.050	0.9	0.6	0.723	0.491	0.048
	0.8	0.411	0.580	0.059		0.8	0.665	0.584	0.055
	1.0	0.356	0.665	0.054		1.0	0.607	0.675	0.051

Table 5. Effect of failure rate (α) and repair rate (β) on various performance measures for finite population model

α	t	$P_I(t)$	$P_B(t)$	$P_D(t)$	β	t	$P_I(t)$	$P_B(t)$	$P_D(t)$
0.1	0.6	0.485	0.495	0.050	1.5	0.6	0.488	0.495	0.047
	0.8	0.411	0.580	0.059		0.8	0.4148	0.578	0.057
	1.0	0.356	0.665	0.054		1.0	0.360	0.661	0.052
0.5	0.6	0.405	0.498	0.216	2.0	0.6	0.490	0.494	0.045
	0.8	0.330	0.597	0.278		0.8	0.417	0.575	0.056
	1.0	0.276	0.687	0.337		1.0	0.363	0.660	0.048

For finite capacity and finite population models, Tables 2–5 show the effect of arrival rate (λ_1), service rate (μ_2), breakdown rate (α) and repair rate (β) on various performance indices for varying values of time t . We notice that $P_B(t)$ and $P_D(t)$ increase, whereas $P_I(t)$ decreases as time t grows. It is clear from Tables 2 and 4 that $P_B(t)$ and $P_D(t)$ increase (decrease) with the increasing values of ($\lambda_1(\mu_1)$) but the reverse trend is observed for $P_I(t)$. On the other hand, in Tables 3 and 5, the

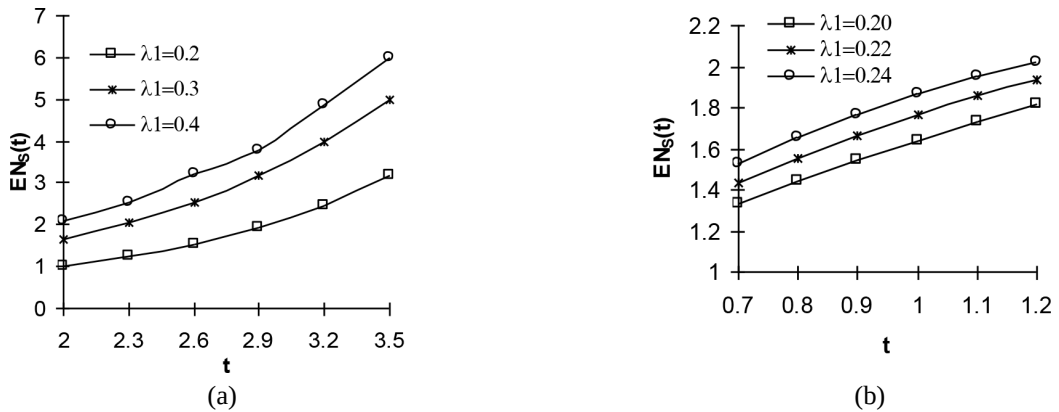


Figure 3. $EN_S(t)$ vs t by varying λ_1 (a) finite capacity model and (b) finite population model

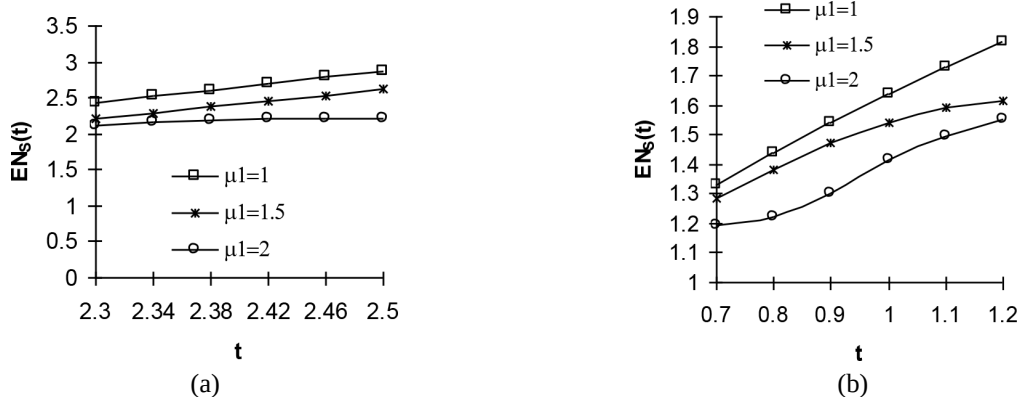


Figure 4. $EN_S(t)$ vs t by varying μ_1 (a) finite capacity model and (b) finite population model

decreasing (increasing) trend is seen for $P_I(t)$ with the increasing values of $(\alpha(\beta))$. $P_B(t)$ and $P_D(t)$ show the increasing (decreasing) trends by increasing the values of failure (repair) rate.

We display the numerical results for the expected number of the customers in the system ($EN_S(t)$), expected delay ($ED(t)$) and throughput ($TP(t)$) against the parameters $\lambda, \mu, \alpha, \beta$ and n in Figures 3(a)–14(b). Figures 3(a)–14(b) show that $EN_S(t)$, $ED(t)$ and $TP(t)$ increase sharply as time t grows. From Figures 3(a)–5(b), an increasing (decreasing) trend is seen for $EN_S(t)$ on increasing the values of $(\lambda_1(\mu_1))$ for both finite capacity and finite population models. It is also observed that initially there is a gradual increment (decrement) but later on, for higher values of $\lambda_1(\mu_1)$, there is a sharp increment (decrement). Figures 5(a) and 6(b) depict how the break-down rate (α) and repair rate (β) of the server affect the system size. It is observed that $EN_S(t)$ increases (decreases) sharply with the increasing values of $\alpha(\beta)$ which agrees with real-life congestion situations for both the models. From Figures 7(a) and 8(b) for both the models, we can easily observe that initially $ED(t)$ decreases (increases) rapidly for increasing values of $\mu_1(\lambda_1)$ then becomes almost constant. As the

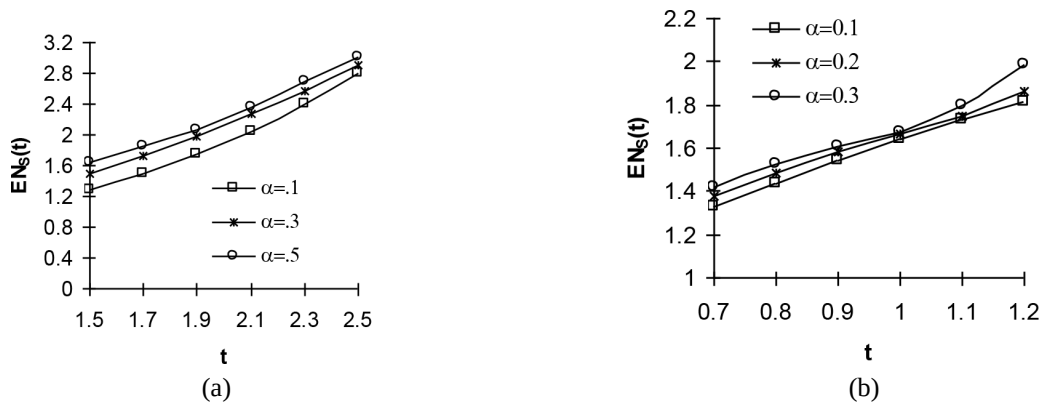


Figure 5. $EN_s(t)$ vs t by varying α for (a) finite capacity model and (b) finite population model

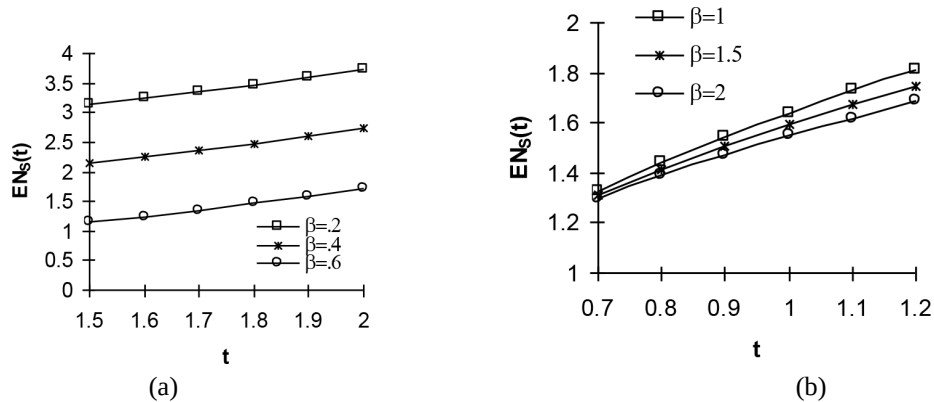


Figure 6. $EN_s(t)$ vs t by varying β for (a) finite capacity model and (b) finite population model

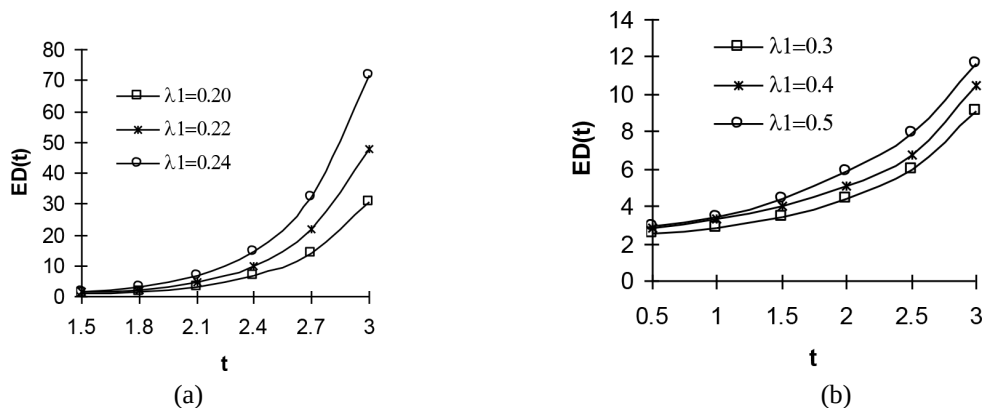


Figure 7. $ED(t)$ vs t by varying λ_1 for (a) finite capacity model and (b) finite population model

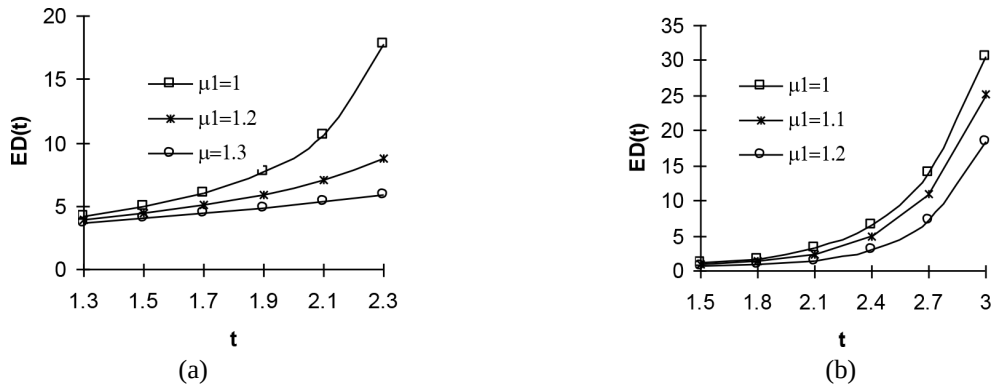


Figure 8. ED (t) vs t by varying μ_1 for (a) finite capacity model and (b) finite population model

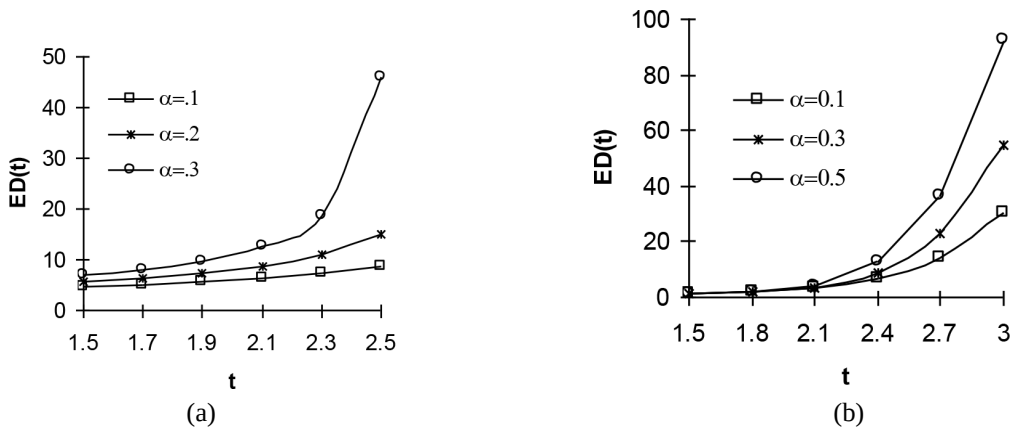


Figure 9. ED (t) vs t by varying α for (a) finite capacity model (b) finite population model

values of $\alpha(\beta)$ increase, an increasing (decreasing) trend is noted for ED(t) for both the models from Figures 9(a) and 10(b), respectively. Figures 11(a) and 12(b) exhibit the increasing (decreasing) trends of throughput for the increasing values of service (arrival) rate for both the models. On the other hand, TP(t) decreases (increases) as the values of $(\alpha(\beta))$ increase as depicted in Figures 13(a) and 14(b) for both finite capacity and finite population models.

In Figures 15(a) and 16(c), we exhibit the effect of time 't' along with arrival rate (λ_2) and failure rate (α) on the expected total cost ($E\{C(t)\}$) for different sets of cost elements as provided in Table 1. From these figures, it can easily be seen that $E\{C(t)\}$ first decreases gradually then increases sharply as time 't' increases for all three sets of cost elements. Total expected cost increases as the values of arrival rate and failure rate increase.

Overall based on numerical experiment performed by taking illustration, we conclude that better service can be expected by the server up to a certain level if we increase the service rate. As expected, the average system size can also be controlled by improving the service and repair rate

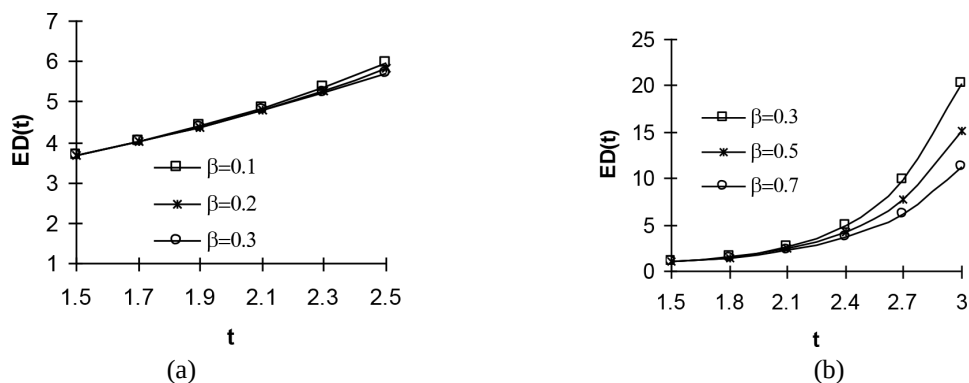


Figure 10. $ED(t)$ vs t by varying β for (a) finite capacity model and (b) finite population model

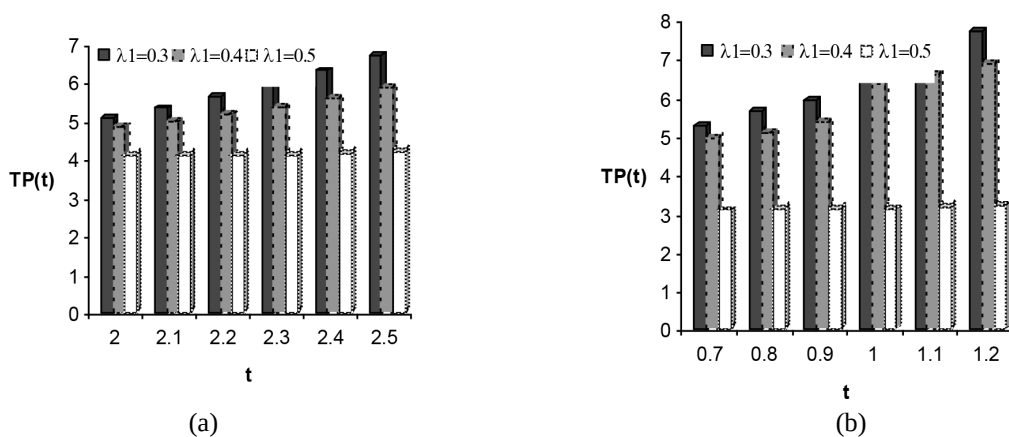


Figure 11. $TP(t)$ vs t by varying λ_1 for (a) finite capacity model and (b) finite population model

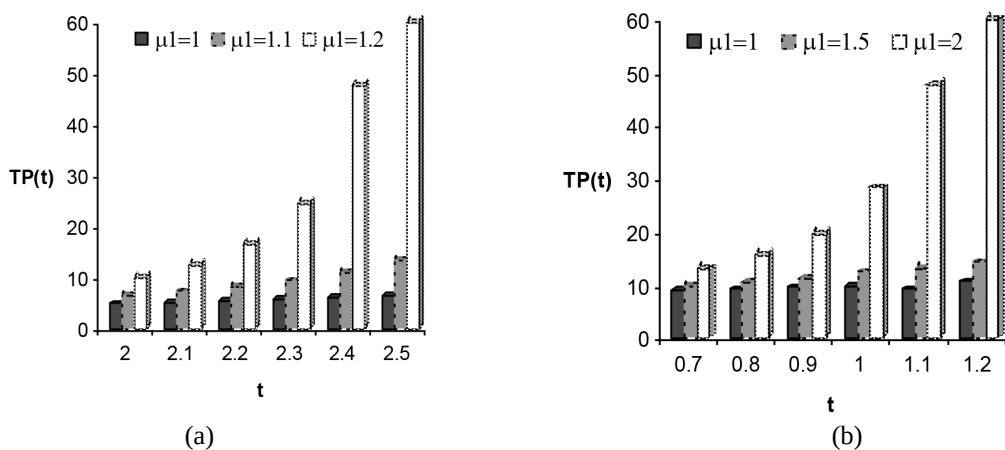


Figure 12. $ED(t)$ vs t by varying μ_2 for (a) finite capacity model and (b) finite population model

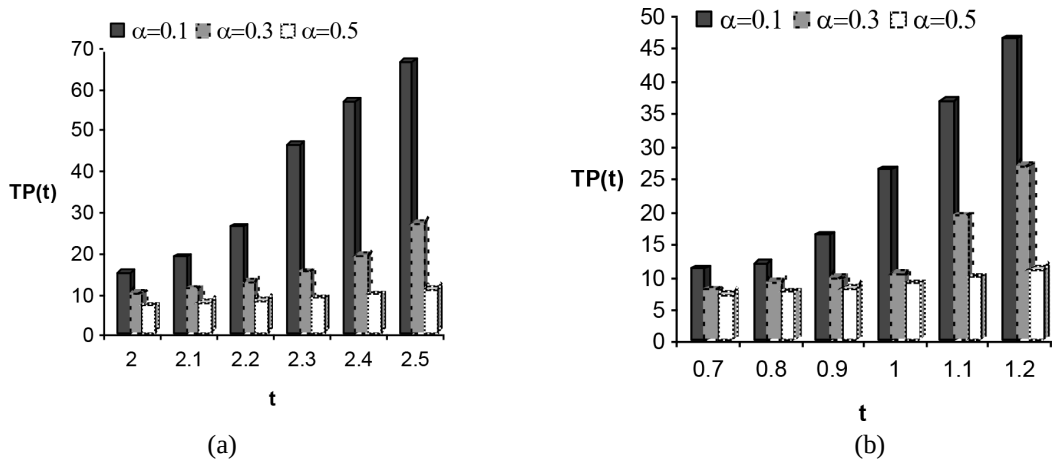


Figure 13. ED (t) vs t by varying α for (a) finite capacity model and (b) finite population model

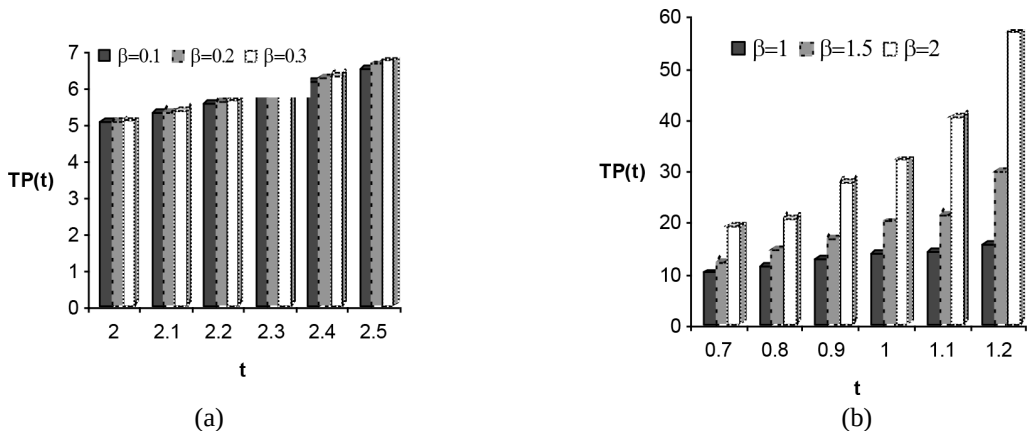


Figure 14. ED (t) vs t by varying β for (a) finite capacity model and (b) finite population model

of the server. Numerical results reveal that expected delay of the system can be decreased to some extent when the server provides service to the customers with high rate and is less prone to breakdowns. In addition, as service rate of the server increases, throughput shows the increasing trend which is quite obvious.

7. CONCLUSIONS

In this investigation, two-class preemptive priority queueing system with service interruption has been studied. The provision of service interruption makes the model more versatile in depicting the many real-time congestion situations encountered in industrial as well as day-to-day problems. The transient model discussed in the present investigation may provide a fruitful insight to the

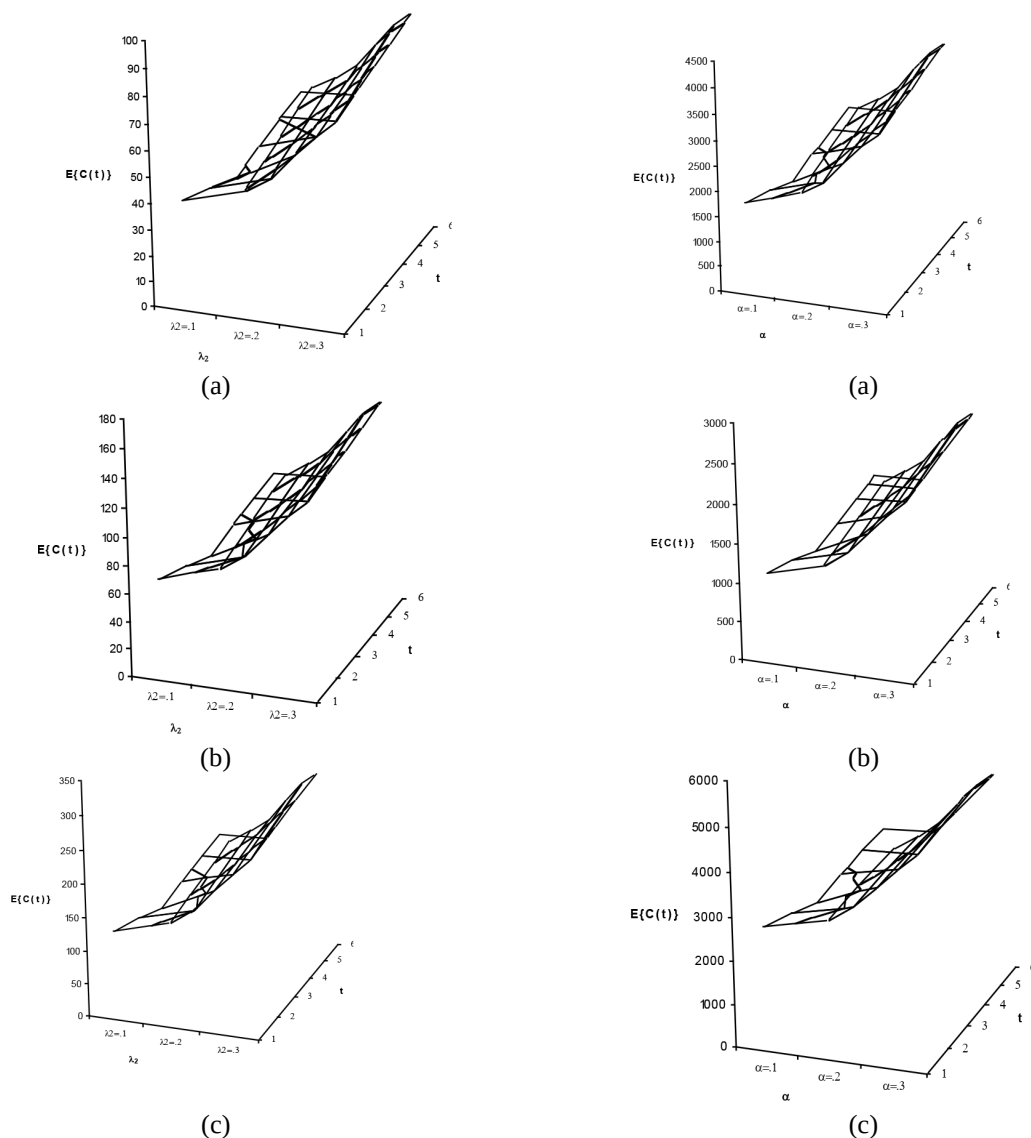


Figure 15. Effect of λ_2 and t on the expected total cost for (a) Set 1, (b) Set 2 and (c) Set 3.
Figure 16. Effect cost for of α and t on the expected total cost for (a) Set 1, (b) Set 2 and (c) Set 3

system developers and queue analysts for assessing the actual congestion problems based upon the performance measures, so that the blocking and delay can be controlled. The future research may include the extension of present model by adding the additional features like bulk arrival and multi-phase optional repair facilities.

REFERENCES

- [1] Bose, I., Pal, R., 2002. Average Waiting Time of Customers in a Priority M/D/k Queue with Finite Buffers. *Computers & Operations Research*, 29(4), pp. 327–339.
- [2] Tarabia, A.M.K., 2007. Two-class Priority Queueing System with Restricted Number of Priority Customers. *International Journal of Electronics and Communications*, 61(8), pp. 534–539.
- [3] Feng, W., Umemura, M., 2009. Analysis of a Finite Buffer Model with Two Servers and Two Nonpreemptive Priority Classes. *European Journal of Operational Research*, 192, pp. 151–172.
- [4] Subramanian, A.M.G., Ayyappan, G., Sekar, G., 2010. M/M/1 Retrial Queueing System with Pre-emptive Priority Service and Standby Server. *International Journal of Computational and Applied Mathematics*, 5(4), pp. 487–499.
- [5] Kim, C., Dudin, S., 2011. Priority Tandem Queueing Model with Admission Control. *Computers & Industrial Engineering*, 61, pp. 131–140.
- [6] Kim, C., Dudin, S., 2011. Priority Tandem Queueing Model with Admission Control. *Computers & Industrial Engineering*, 61, pp. 131–140.
- [7] Kim, K., 2012. T-preemptive Priority Queue and Its Application to the Analysis of an Opportunistic Spectrum Access in Cognitive Radio Networks. *Computers & Operations Research*, 39(7), pp. 1394–1401.
- [8] Ke, J.C., 2004. Bi-level Control for Batch Arrival Queues with an Early Startup and Un-reliable Server. *Applied Mathematical Modelling*, 28(5), pp. 469–485.
- [9] Sherman, N.P., Kharoufeh, J.P., 2006. An M/M/1 Retrial Queue with Unreliable Server. *Operations Research Letters*, 34(6), pp. 697–705.
- [10] Economou, A., Kanta, S., 2008. Equilibrium Balking Strategies in the Observable Single-server Queue with Breakdowns and Repairs. *Operations Research Letters*, 36(6), pp. 696–699.
- [11] Economou, A., Kapodistria, S., 2010. Synchronized Abandonments in a Single Server Unreliable Queue. *European Journal of Operational Research*, 203(1), pp. 143–155.
- [12] Jain, M., Sharma, G.C., Sharma, R., 2011. Working Vacation Queue with Service Interruption and Multi-optional Repair. *International Journal of Information and Management – Science*, 22, pp. 157–175.
- [13] Jain, M., Sharma, G.C., Sharma, R., 2012. Optimal Control of (N,F) Policy for Unreliable Server Queue with Multi Optional Phase Repair and Start Up. *International Journal of Mathematics & Operational Research*, 4(2), pp. 152–174.
- [14] White, H., Christie, L., 1958. Queueing with Preemptive Priorities or with Breakdown. *Operational Research*, 7, pp. 86–103.
- [15] Towsley, D., Tripathi, S.K., 1991. A Single Server Priority Queue with Server Failures and Queue Flushing. *Operations Research Letters*, 10(6), pp. 253–362.
- [16] Jain, M., 1999. Single Server Queue with Server Breakdown Including Priority and Varying Rates. *JKAU: Engineering Science*, 11(2), pp. 51–60.
- [17] Jain, M., Bhargava, C., 2008. Bulk Arrival Retrial Queue with Unreliable Server and Priority Subscribers. *International Journal of Operational Research*, 5(4), pp. 242–259.
- [18] Wang, J., 2008. On the Single Server Retrial Queue with Priority Subscribers and Server Breakdowns. *Journal of Systems Science and Complexity*, 21, pp. 304–315.
- [19] Kamoun, F., 2008. Performance Analysis of a Non-preemptive Priority Queueing System Subject to a Correlated Markovian Interruption Process. *Computers & Operations Research*, 35(12), pp. 3969–3988.
- [20] Dimitriou, L., 2013. A Preemptive Resume Priority Retrial Queue with State Dependent Arrivals, Unreliable Server and Negative Customers. *TOP*, 21(3), pp. 542–571 (In Press).
- [21] Van As, H., 2003. Transient Analysis of Markov Queueing Systems and Its Application to Congestion-control Modeling. *IEEE Journal on Selected Areas in Communications*, 4(6), pp. 891–904.
- [22] Jain, M., Sharma, G.C., Singh, N., 2007. Transient Analysis of M/M/R Machining System with Mixed Standbys, Switching Failures, Balking, Reneging and Additional Removable Repairmen. *International Journal of Engineering (IJE), Transactions A: Basics*, 20(2), pp. 168–182.
- [23] Jain, M., Sharma, G.C., Mittal, R., 2008. Transient Analysis of Channel Allocation in Cellular Radio Network with Balking and Reneging. *Computer Society of India*, 38(4), pp. 11–19.
- [24] Maheshwari, S., Sharma, P., Jain, M., 2010. Unreliable Flexible Manufacturing Cell with Common Cause Failure. *International Journal of Engineering Science and Technology*, 2(8), pp. 4071–4716.
- [25] Jain, M., Upadhyaya, S., 2011. Synchronous Working Vacation Policy Finite-buffer Multi Server Queueing System. *Applied Mathematics and Computation*, 217(24), pp. 9916–9932.