International Journal of System and Software Engineering
Volume 5 Issue 1, June 2017
ISSN.: 2321-6107

Sensitivity Analysis of Markovian Queue with Discouragement, Additional Servers and Threshold Policy

Neetu Singh
World College of Technology and Management, Gurugram, Haryana, India.
E-mail: neetusingh1608@gmail.com

Abstract: In the present paper, an attempt has been made to study the optimal threshold policy for Markovian queueing model having additional servers along with permanent server. The incorporation of customer's balking and reneging behavior has been done. The customers arrive in Poisson fashion and their service times are exponentially distributed. The first server starts service when N (≥1) or more customers are accumulated and turns off when the system is empty. The (j+1)th (j=1, 2,...., r-1) server turns on when there are Nj+1 customers in the system and will be removed as soon as the number of customers drops to threshold level Nj. We use Laplace transform technique to derive transient probabilities and some other system characteristics such as the expected number of jobs in the system, throughput, and probability that jth (j=1,2,3,…,r) server being busy in rendering the service, etc.. The effects of system parameters on the performance characteristics have been examined by taking numerical illustrations.

Keywords: Threshold policy, Markovian queue, Balking, Reneging, Additional servers, Transient analysis, Throughput, Delay time.

I. Introduction

In the economically competitive world, the N-policy for queueing systems has drawn the attention of many researchers since last four decades and so, as if server is idle till next arrival there may be waste of valuable resources. In N-policy, the server turns on at threshold level N of waiting customers and once server starts service, it continues until there is no customer present in the system.

Optimal N-policy for removable service station was first introduced by Yadin and Naor [36]. Wang [35] studied optimal operation of a single removable and non-reliable server in Markovian queue under optimal N-policy. Jain [12] studied optimal N-policy for single removable and unreliable server subject to breakdown and repair and obtained various operational characteristics. An analysis of the M/G/1 queueing system with N and T policy was discussed by Hur et al. [11]. Tian and Zhang [31] studied performance and optimization issues in a vacation system with (D,N) threshold policy. Arumunganathan and Jeyakumar [4] analyzed N-policy for a steady state bulk queue with multiple vacations, set up times and closed down times. Wang and Pearn [34] investigated N policy M/G/1 queueing system with unreliable server to analyze system performance measures and give the expected cost function per unit time to determine the optimal threshold N at a minimum cost. Diaz and Moreno [7] studied a queueing system where the service station operates under an N-policy. N policy queueing systems have been taken into consideration by several researchers in different frameworks [20], [30], [32].

The queueing models have tremendous potentialities in promoting industrial as well as economic growth. The interest in studies of queueing theory is increasing rapidly due to its wide utility in many real life congestion situations. There may be some situations when the arriving customers get discouraged due to long queue, and may not like to join the queue upon arrival if service cannot be provided immediately or if queue is too long. This behavior is known as balking. They may also renege due to impatience after waiting in the queue for some time. Several authors have successfully studied the balking and reneging behavior of the customers in different frameworks. The stationary solution of multi-server model with discouragement was given by Reynolds [23]. Shawky [28] developed M/M/C/K/N model with spares by implementing the customer's balking and reneging behavior. Drekic et al. [8] investigated a preemptive priority queue with balking. Wang et al. [33] analyzed M/M/R machine repair problem with balking and reneging. Queueing systems with balking and reneging behavior of customers have been analyzed by several researchers from time to time [22], [2], [3], [9].

The combination of permanent and removable additional servers is used to achieve the higher grade of service. Also to meet out the situation of backlog and discouragement, the provision of additional servers is recommended. Abou Et Ata and Shawky [1] considered the single server Markovian over flow queue with balking, reneging and additional servers for longer queues. Reliability analysis of a repairable system with spares and a removable repairman has been done by Hsieh and Wang [10]. Shawky [27] presented single server machine interference model with balking, reneging and additional server. M/M/m queueing system with discouragement and additional servers was studied by Jain [13]. The M/M/C/K/N repairable system under the effect of balking and reneging with the provision of additional server and spares was analyzed by Jain et al. [18]. Jain et al. [17] incorporated balking and reneging for M/M/R machine interference model. A batch arrival queue with an additional service channel under N-Policy was analyzed by Chaudhury and Paul [6]. Recently, queueing model with additional server has been investigated by [19].

The transient behavior of queueing systems is not easy to tackle even for markovian models. However a few researchers have made attempt to present transient analysis; a detailed account for the same can be found in a reference book by Sharma [25]. Transient behavior of M/M/R machine repair problem with spares has been taken in to consideration by Sridaran and Jayashree [29]. Sharma and Tarabia [26] gave the transient analysis of M/M/1/N queue. Jain [14] analyzed N- policy for redundant repairable system with additional repairmen, which turn on one by one according to a threshold policy depending upon the queue size. Jain et al. [15] investigated transient analysis of N policy for a machine repair system with spares and reneging. Parthasarathy and Sudhesh [21] analyzed time dependent single server queueing system with state dependent rates. Transient analysis of multi-server queueing systems was done by Chandrika [5], Jain et al. [16] and Al-seedy et al. [2]. A detailed survey and classification of the literature on time dependent queueing systems can be found in Schwarz et al. [24].

The purpose of this paper is to investigate Markovian queue with discouragement under optimal threshold policy. Only one server serves the arriving customers when there are more than N but less than N_1 customers present in the system. When the number of customers exceeds a predefined threshold level at arrival epochs queueing system, an extra server is added and is removed when the number of customers in the system becomes less than certain threshold value. The rest paper is arranged in various sections systematically as follows. The mathematical model, formulation, underlying assumptions and transient equations are given in section 2. The solution of governing transient state equations is facilitated in section 3. Section 4 provides some system characteristics. In section 5, we construct a relationship among the costs to determine the optimal threshold parameters N_j. The sensitivity analysis is carried out in section 6. Finally, conclusion is given in section 7.

II. Model Description

We address finite capacity Markovian queueing model with discouragement and additional heterogeneous servers. The customers arrive according to exponential distribution with rate $\underline{\lambda}$ and are served according to first come first served discipline. The service facility consists of r heterogeneous servers who turn on one by one according to a threshold policy. The service times are assumed to be exponentially distributed. The first server turns on when at least N (≥ 1) customers are present in the system and turns off when all customers get serviced and system becomes empty. The first server serves the customers with rate μ_1 when there are greater than N and less than or equal to N_1 (i.e. $N < n \le N_1$) customers in the system whereas he serves with rate μ_0 when there are greater than or equal to 1 and less than or equal to N (i.e. $1 \le n \le N$) customers in the system. The (j+1) th (j=1, 2, 3, …, r-1) server activates when there is the workload of N_j+1 customers in the system and provides service with rate μ_{j+1} (1,2,…,r-1). When the number of customers in the system reduces to threshold level N_j, the last included (j+1)th server will be removed.

The customers may balk with probability (1-b_1) from the system i.e. b_1 is the joining probability of customers when first server is on and there are greater than or equal to N and less than or equal to N_1 (i.e. $N < n \le N_1$) customers in the system; however when there are greater than or equal to 1 and less than N (i.e. $1 \le n < N$) customers in the system, the joining probability is b_0. The joining probability of customers is assumed to be b_j (j=2,3,…,r) when j servers are rendering service only.

The customers waiting in the queue may get impatient and renege from the system without getting service according to exponential distribution. When first server is on and there are greater than N and less than or equal to N_1 (i.e. $N < n \le N_1$) customers in the system, the reneging parameter is α_1; whereas when there are greater than or equal to 1 and less than or equal to N (i.e. $1 \le n \le N$) customers in the system, the reneging parameter is α_0. When j servers are providing service, the reneging parameter is assumed to be α_j (j=2,3,…,r).

The state space of the system at time t is {C(t), X(t); t ≥ 0} where C(t) denotes the number of busy servers at time t and X(t) be the number of customers in the system at time t.

Let

$$C(t) = \begin{cases} 0; & \text{if all the servers are idle}; 0 < n < N \\ 1; & \text{if only } 1^{st} \text{ server is providing service}; 1 < n \le N_1 \\ 2; & \text{if the } j \text{ servers are providing service}; N_{j-1} < n \le N_j, j = 2,3,\ldots,r \end{cases}$$

We define transient probabilities as

$$P_{0,n}(t) = \Pr ob.\{C(t) = 0;\ X(t) = n, 1 < n \le N\}$$

$$P_{j,n}(t) = \Pr ob.\{C(t) = j\ (j = 1,2,\ldots,r); X(t)$$
$$= n, N_{j-1} < n \le N_j,\ N_0 = 1\}$$

By using appropriate transition rates, the state transition flow diagram is constructed in figure 1 where for the brevity we have used the notations:

$$\Lambda^k = \lambda b_j, k = 0,1,\ldots,r; \quad \mu_n^0 = \mu_0 + (n-1)\alpha_0; \quad \mu_n^k$$
$$= \sum_{i=1}^{j} \mu_i + (n-j)\alpha_j, \quad k = 1, 2, 3, \ldots\ldots, r$$

The transient state probabilities satisfy the following Chapman-Kolmogorov equations:

$$\frac{dP_{0,0}(t)}{dt} = -\lambda P_{0,0}(t) + \mu_0 P_{1,1}(t) \qquad \ldots(1)$$

$$\frac{dP_{0,n}(t)}{dt} = -\lambda P_{0,n}(t) + \lambda P_{0,n-1}(t),\ 0 < n \leq N-1 \qquad \ldots(2)$$

$$\frac{dP_{1,1}(t)}{dt} = -(\lambda b_0 + \mu_0)P_{1,1}(t) + (\mu_0 + \alpha_0)P_{1,2}(t) \qquad \ldots(3)$$

$$\frac{dP_{1,n}(t)}{dt} = -\left[\lambda b_0 + \mu_0 + (n-1)\alpha_0\right]P_{1,n}(t) + \lambda b_0 P_{1,n-1}(t) + (\mu_0 + n\alpha_0)P_{1,n+1}(t), 1 < n \leq N-1 \qquad \ldots(4)$$

$$\frac{dP_{1,N}(t)}{dt} = -\left[\lambda b_1 + \mu_0 + (N-1)\alpha_0\right]P_{1,N}(t) + \lambda b_0 P_{1,N-1}(t) + \lambda P_{0,n-1}(t) + (\mu_1 + N\alpha_1)P_{1,N+1}(t) \qquad \ldots(5)$$

$$\frac{dP_{1,n}(t)}{dt} = -\left[\lambda b_1 + \mu_1 + (n-1)\alpha_1\right]P_{1,n}(t) + \lambda b_1 P_{1,n-1}(t) + (\mu_1 + n\alpha_1)P_{1,n+1}(t),\ N \leq n < N_1 \qquad \ldots(6)$$

$$\frac{dP_{1,N_1}(t)}{dt} = -\left[\lambda b_1 + \mu_1 + (N_1-1)\alpha_1\right]P_{1,N_1}(t) + \lambda b_1 P_{1,N_1-1}(t) + \left[\sum_{i=1}^{2}\mu_i + (N_1-1)\alpha_2\right]P_{2,N_1+1}(t) \qquad \ldots(7)$$

$$\frac{dP_{j+1,N_j+1}(t)}{dt} = -\left[\lambda b_{j+1} + \sum_{i=1}^{j+1}\mu_i + (N_j - j)\alpha_{j+1}\right] P_{j+1,N_j+1}(t) + \lambda b_j P_{j,N_j}(t) + \left[\sum_{i=1}^{j+1}\mu_i + (N_j + 1 - j)\alpha_{j+1}\right] P_{j+1,N_j+2}(t),\ \ j = 1, 2, ..., r-1 \qquad \ldots(8)$$

$$\frac{dP_{j+1,n}(t)}{dt} = -\left[\lambda b_{j+1} + \sum_{i=1}^{j+1}\mu_i + (n-j-1)\alpha_{j+1}\right] P_{j+1,n}(t) + \lambda b_{j+1} P_{j+1,n-1}(t) + \left[\sum_{i=1}^{j+1}\mu_i + (n-j)\alpha_{j+1}\right]P_{j+1,n+1}(t)\ , N_j + 1 < n \leq N_{j+1}, j = 1, 2, ..., r-2 \qquad \ldots(9)$$

$$\frac{dP_{r,n}(t)}{dt} = -\left[\lambda b_r + \sum_{i=1}^{r}\mu_i + (n-r)\alpha_r\right] P_{r,n}(t) + \lambda b_r P_{r,n-1}(t) + \left[\sum_{i=1}^{r}\mu_i + (n+1-r)\alpha_r\right]P_{r,n+1}(t),\ N_{r-1} + 1 < n < N_r \qquad \ldots(10)$$

$$\frac{dP_{r,N_r}(t)}{dt} = -\left[\sum_{i=1}^{r}\mu_i + (N_r - r)\alpha_r\right] P_{r,N_r}(t) + \lambda b_r P_{r,N_r-1}(t) \qquad \ldots(11)$$

Let us assume that initially the system is empty so the initial conditions are given by

$$P_{0,0}(0) = 1; P_{0,n}(0) = 0\ (n = 1, 2, ..., N-1); P_{1,n}(0) = 0\ (n = N, N+1, ..., N_1);$$
$$P_{j,n}(0) = 0\ (n = 1, 2, ..., N_r; j = 2, 3, ..., r). \qquad \ldots(12)$$

III. The Analysis

Let $\tilde{P}_{i,n}(s)$ denotes the Laplace transform of $P_{i,n}(t)$ and is defined as:

$$\tilde{P}_{i,n}(s) = \int_0^\infty e^{-st} P_{i,n}(t)dt \text{ , i=0,1,2,…,r.}$$

Taking Laplace transforms of Eqs. (1)- (11), we obtain

$$(s+\lambda)\tilde{P}_{0,0}(s) - \mu_0 \tilde{P}_{1,1}(s) = P_{0,0}(0) \qquad \ldots(13)$$

$$(s+\lambda)\tilde{P}_{0,n}(s) - \lambda \tilde{P}_{0,n-1}(s) = P_{0,n}(0),\ 0 < n \leq N-1 \qquad \ldots(14)$$

$$(s + \lambda b_0 + \mu_0)\tilde{P}_{1,1}(s) - (\mu_0 + \alpha_0)\tilde{P}_{1,2}(s) = P_{1,1}(0) \qquad \ldots(15)$$

$$\left[s + \lambda b_0 + \mu_0 + (n-1)\alpha_0\right]\tilde{P}_{1.n}(s) - \lambda b_0 \tilde{P}_{1,n-1}(s) - (\mu_0 + n\alpha_0)\tilde{P}_{1,n+1}(s) = P_{1,n}(0), 1 < n \leq N-1 \qquad \ldots(16)$$

$$\left[s + \lambda b_1 + \mu_0 + (N-1)\alpha_0\right]\tilde{P}_{1,N}(s) - \lambda b_0 \tilde{P}_{1,N-1}(s) - \lambda \tilde{P}_{0,N-1}(s) - (\mu_1 + N\alpha_1)\tilde{P}_{1,N+1}(s) = P_{1,N}(0) \qquad \ldots(17)$$

$$\left[s + \lambda b_1 + \mu_1 + (n-1)\alpha_1\right]\tilde{P}_{1,n}(s) - \lambda b_1 \tilde{P}_{1,n-1}(s) - (\mu_1 + n\alpha_1)\tilde{P}_{1,n+1}(s) = P_{1,n}(0),\ \ N \leq n < N_1 \qquad \ldots(18)$$

$$\left[s + \lambda b_2 + \mu_1 + (N_1-1)\alpha_1\right]\tilde{P}_{1,N_1}(s) - \lambda b_1 \tilde{P}_{1,N_1-1}(s) - \left(\sum_{i=1}^{2}\mu_i + (N_1-1)\alpha_2\right)\tilde{P}_{2,N_1+1}(s) = P_{1,N_1}(0) \qquad \ldots(19)$$

$$\left[s + \lambda b_{j+1} + \sum_{i=1}^{j+1}\mu_i + (N_j - j)\alpha_{j+1}\right]\tilde{P}_{j+1,N_j+1}(s) - \lambda b_j \tilde{P}_{j,N_j}(s) - \left(\sum_{i=1}^{j+1}\mu_i + (N_j + 1 - j)\alpha_{j+1}\right)\tilde{P}_{j+1,N_j+2}(s) = P_{j+1,N_j+1}(0),\ \ j = 2, 3, ..., r-2 \qquad \ldots(20)$$

$$\left[s + \lambda b_{j+1} + \sum_{i=1}^{j+1}\mu_i + (n-j-1)\alpha_{j+1}\right]\tilde{P}_{j+1,n}(s) - \lambda b_{j+1}\tilde{P}_{j+1,n-1}(s) - \left(\sum_{i=1}^{j+1}\mu_i + (n-j)\alpha_{j+1}\right)\tilde{P}_{j+1,n+1}(s) = P_{j+1,n}(0),\ N_j + 1 < n \leq N_{j+1},\ \ j = 1, 2, 3, ..., r-2 \qquad \ldots(21)$$

$$\left[s + \lambda b_r + \sum_{i=1}^{r}\mu_i + (n-r)\alpha_r\right]\tilde{P}_{r,n}(s) - \lambda b_r \tilde{P}_{r,n-1}(s) - \left(\sum_{i=1}^{r}\mu_i + (n-r+1)\alpha_r\right)\tilde{P}_{r,n+1}(s) = P_{r,n}(0), N_{r-1} + 1 < n < N_r \qquad \ldots(22)$$

$$\left[s+\sum_{i=1}^{r}\mu_i+(N_r-r)\alpha_r\right]\tilde{P}_{r,N_r}(s)-\lambda b_r\tilde{P}_{r,N_r-1}(s) \qquad \ldots(23)$$

$$=P_{r,N_r}(0)$$

The matrix form of Eqs. (13)-(23) can be written as

$$B(s)\tilde{P}(s)=P(0) \qquad \ldots(24)$$

where $\tilde{P}(s)$ and $P(0)$ are column vectors of order (N+N$_r$), and are given as follows:

$$\tilde{P}(s)=[\tilde{P}_{0,0}(s),\tilde{P}_{0,1}(s),\ldots,\tilde{P}_{0,N-1}(s),\tilde{P}_{1,1}(s),\ldots,\tilde{P}_{1,N}(s),\ldots,\tilde{P}_{1,N_1-1}(s),\tilde{P}_{1,N_1}(s),\tilde{P}_{2,N_1+1}(s),\ldots,\tilde{P}_{j,n}(s),\ldots,\tilde{P}_{r,N_r-1}(s),\tilde{P}_{r,N_r}(s)] \qquad \ldots(25)$$

and

$$P(0)=[P_{0,0}(0),P_{0,1}(0),\ldots,P_{0,N-1}(0),P_{1,1}(0),\ldots,P_{1,N}(0),\ldots,P_{1,N_1-1}(0),P_{1,N_1}(0),P_{2,N_1+1}(0),\ldots,P_{j,n}(0),\ldots,P_{r,N_r-1}(0),P_{r,N_r}(0)] \qquad \ldots(26)$$

The matrix B(s) is a symmetric matrix of order $(N + N_r) \times (N+N_r)$.

Using the matrix approach (cf. Hsieh and Wang [5]), we obtain an explicit expression for $\tilde{P}_{j,n}(s)$ as

$$\tilde{P}_{j,n}(s)=\frac{|B_n(s)|}{|B(s)|};\ j=0,1,\ldots,r;\quad n=0,1,\ldots,N,N+1,\ldots,N_r \qquad \ldots(27)$$

where $|B(s)|$ is the determinant of matrix B(s) and $|B_n(s)|$ is the determinant obtained by replacing the n$^{\text{th}}$ (n=0, 1, 2, …, N, …, N$_r$) column of matrix B(s) by the initial vector $P(0)=[1,0,0,\ldots,0]^T$.

To obtain the explicit expression for $P_{j,n}(t)$, we proceed as follows:

We observe that s=0 is a root of $|B(s)|=0$. Also by putting $|B(s)-aI|=0$, we get the distinct eigen values (real or complex) of the matrix. We assume that there are q non-zero real distinct eigen values $a_1,a_2,\ldots,a_q$ and p pairs of distinct conjugate complex eigen values denoted as $\left(a_{q+1},\bar{a}_{q+1}\right),\left(a_{q+2},\bar{a}_{q+2}\right),\ldots\ldots,\left(a_{q+r},\bar{a}_{q+p}\right)$.

It is noted that $q+2p=N+N_r-1$

Thus we get

$$|B(s)|=s\left[\prod_{l=1}^{q}(s+a_l)\right]\left[\prod_{l=1}^{p}\left\{s^2+(a_{q+l}+\bar{a}_{q+l})s+a_{q+l}\bar{a}_{q+l}\right\}\right] \qquad \ldots(28)$$

and $\tilde{P}_{j,n}$(s) can be written as

$$\tilde{P}_{j,n}(s)=\frac{|B_n(s)|}{s\left[\prod_{l=1}^{q}(s+a_l)\right]\left[\prod_{l=1}^{p}\left\{s^2+(a_{q+l}+\bar{a}_{q+l})s+a_{q+l}\bar{a}_{q+l}\right\}\right]};\ j=0,1,\ldots,r;\ n=0,1,\ldots,N,N+1,\ldots,N_r \qquad \ldots(29)$$

The partial fractions of $\tilde{P}_{j,n}$(s) are obtained as

$$\tilde{P}_{j,n}(s)=\frac{b_0}{s}+\sum_{i=1}^{q}\frac{b_i}{s+a_i}+\sum_{i=1}^{p}\frac{c_i s+d_i}{s^2+\left(a_{q+i}+\bar{a}_{q+i}\right)s+a_{q+i}\bar{a}_{q+i}} \qquad \ldots(30)$$

where b_0 and b_i, (i= 1, 2, …, q) are real numbers given by

$$b_0=\frac{B_n(0)}{\left[\prod_{l=1}^{q}a_l\right]\left[\prod_{l=1}^{p}a_{q+l}\bar{a}_{q+l}\right]} \qquad \ldots(31)$$

$$b_i=\frac{|B_n(-a_i)|}{(-a_i)\left[\prod_{\substack{l=1\\l\neq u}}^{q}(a_l-a_i)\right]\left[\prod_{l=1}^{p}\left\{a_i^2+(a_{q+l}+\bar{a}_{q+l})(-a_i)+a_{q+l}\bar{a}_{q+l}\right\}\right]},$$

$$i=1,2,\ldots,q \qquad \ldots(32)$$

and $c_i(-a_{q+i})+d_i$ is obtained as

$$c_i(-a_{q+i})+d_i=\frac{|B_n(-a_{q+i})|}{(-a_{q+i})\left[\prod_{\substack{l=1\\l\neq u}}^{q}(a_l-a_{q+i})\right]\left[\prod_{l=1}^{p}\left\{(-a_{q+i})^2+(a_{q+i}+\bar{a}_{q+i})(-a_{q+i})+a_{q+i}\bar{a}_{q+i}\right\}\right]},$$

$$i=1,2,\ldots,p \qquad \ldots(33)$$

We denote g_u and $\underline{h_u}$ as the real and imaginary parts of complex eigen value $(-a_{q+i})$. Then the inverse Laplace transforms of Eq. (29) is given by

$$P_{j,n}(t)=b_0+\sum_{i=1}^{q}b_i e^{-a_i l}+\sum_{i=1}^{p}\left[c_i e^{-g_i l}\cos(h_i l)+\frac{d_i-c_i g_i}{h_i}e^{-g_i l}\sin(h_i l)\right] \qquad \ldots(34)$$

IV. SOME PERFORMANCE INDICES

In this section, we establish the expressions for some measures of performance characterizing the system in terms of transient state probabilities obtained in previous section.

The expected number of customers $L\{\mathbf{N_r},t\}$ at any instant t in the system, having r heterogeneous servers is

$$L\{\mathbf{N_r},t\}=\sum_{n=1}^{N-1}nP_{0,n}(t)+\sum_{n=1}^{N_r}\sum_{j=1}^{r}nP_{j,n}(t) \qquad \ldots(35)$$

The throughput of the system $E\{T(t)\}$ at time t is obtained by using

$$E\{T(t)\}=\sum_{n=1}^{N_r}\sum_{j=1}^{r}\mu_n P_{j,n}(t) \qquad \ldots(36)$$

Average delay time τ(t) is given by

$$\tau(t) = \frac{L\{\mathbf{N_r}, t\}}{E\{T(t)\}} \qquad \ldots(37)$$

The probabilities that at any instant the 1st and jth (j=2,3,….,r) servers are in busy state respectively, are given by

$$P^{(1)}(t) = \sum_{n=1}^{N_r} P_{1,n}(t);$$
$$P^{(j)}(t) = \sum_{n=N_{j-1}+1}^{N_r} P_{j,n}(t) \quad (j = 2,3,...,r) \qquad \ldots(38)$$

V. OPTIMAL RESULTS

In the present section we obtain the cost relationship for M/M/r queueing system with queue dependent heterogeneous servers.

Denote

C_h= holding cost per unit time for each customer present in the system

C_j= cost incurred per unit time for providing jth (1,2,3,…,r) server

In order to evaluate the threshold values N_j (j=1,2,3,…,r-1) we construct the cost function using different cost elements to minimize the total average cost $Z = C_h L\{\mathbf{N_r}, t\} + \sum_{j=1}^{r} C_j P^{(j)}(t)$ …(39)

so that the net profit is maximum.

It is realized that the expenses for waiting customers are more than the expenditure of providing the extra servers for service. For making maximum net profit in providing (j+1)th (j=1,2,3,…,r-1) server one by one in the system, the following cost inequalities must be satisfied:

$$C_h[L\{\mathbf{N_j}, t\} - L\{\mathbf{N_r}, t\}] > \sum_{k=j}^{r-1}\left(C_{k+1}\sum_{n=N_k+1}^{N_{k+1}} P_{k+1,n}(t)\right), \qquad \ldots(40)$$
$$j = 1,2,...,r-1$$

where $\mathbf{N_j} = \{N_1, N_2, ..., N_j\}$ let **ej** denotes the vector of size j, having zero elements except 1 at jth position. From eq. (40), the maximum profit can be obtained by applying the (j+1)th server at N_j only if the following inequalities are satisfied:

$$\begin{aligned} C_h[L\{\mathbf{N_j}, t\} - L\{\mathbf{N_r}, t\}] - \sum_{k=j}^{r-1}\left(C_{k+1}\sum_{n=N_k+1}^{N_{k+1}} P_{k+1,n}(t)\right) \\ \geq C_h[L\{\mathbf{N_j}, t\} - L\{\mathbf{N_r} - \mathbf{e_j}, t\}] \\ -\left(C_{j+1}\sum_{n=N_j}^{N_{j+1}} P_{j+1,n}(t)\right) - \\ \sum_{k=j+1}^{r-1}\left(C_{k+1}\sum_{n=N_k+1}^{N_{k+1}} P_{k+1,n}(t)\right); \\ j = 1,2,3,....,r-1 \end{aligned} \qquad \ldots(41)$$

$$\begin{aligned} C_h[L\{\mathbf{N_j}, t\} - L\{\mathbf{N_r}, t\}] - \sum_{k=j}^{r-1}\left(C_{k+1}\sum_{n=N_k+1}^{N_{k+1}} P_{k+1,n}(t)\right) \\ \geq C_h[L\{\mathbf{N_j}, t\} - L\{\mathbf{N_r} + \mathbf{e_j}, t\} \\ -\left(C_{j+1}\sum_{n=N_j+2}^{N_{j+1}} P_{j+1,n}(t)\right) \\ -\sum_{k=j+1}^{r-1}\left(C_{k+1}\sum_{n=N_k+1}^{N_{k+1}} P_{k+1,n}(t)\right); \\ j = 1,2,3,....,r-1 \end{aligned} \qquad \ldots(42)$$

The inequalities (41)-(42) are simplified as

$$\frac{[L\{\mathbf{N_r}, t\} - L\{\mathbf{N_r} - \mathbf{e_j}, t\}]}{P_{j+1,N_j}(t)} \leq \beta_j$$
$$\leq \frac{[L\{\mathbf{N_r} + \mathbf{e_j}, t\} - L\{\mathbf{N_r}, t\}]}{P_{j+1,N_j+2}(t)};\ j = 1,2,3,...,r-1 \qquad \ldots(43)$$

where $\beta_j = \dfrac{C_{j+1}}{C_h}$.

TABLE 1: DIFFERENT SETS OF PARAMETERS USED IN TABLES 2-5 AND FIGURES 2-13

set	μ_1	μ_2	μ_3	μ_4	set	α_1	α_2	α_3	α_4	set	β_1	β_2	β_3	β_4
A.1	1.2	1.2	.5	.5	B.1	.2	1.2	2.2	3.2	C.1	.9	.8	.7	.6
A.2	2.2	2.2	1	1	B.2	2	3	4	5	C.2	.5	.7	.6	.5
A.3	3.2	3.2	1.5	1.5	B.3	4	3.5	4.5	5.5	C.3	.3	.6	.5	.4
set	N	N_1	N_2	N_3	set	α_1	α_2	α_3	α_4	set	μ_1	μ_2	μ_3	μ_4
D.1	2	5	10	20	E.1	.7	.6	.55	.5	F.1	5.5	2	1	1
D.2	3	10	15	25	E.2	.65	.55	.5	.45	F.2	6	2.5	1.5	1.5
D.3	5	15	25	30	E.3	.6	.5	.45	.4	F.3	6.5	3	2	2

VI. SENSITIVITY ANALYSIS

In this section, the sensitivity analysis is carried out in order to demonstrate the effect of various parameters on the system performance. The algorithm described above has been employed to develop computer program in software MATLAB and the computational results are displayed in tables 2-5 and figures 2-13.

Table 1 displays the sets of different values of parameters that have been considered in figs. 2-13 and tables 2-5. In tables 2-5 we display average response time τ(t) by varying λ, μ_j, α_j and βj for different values of N; it is noted that $\tau(t)$ increases with λ, βj and decreases with α_j, μ_j.

The graphical presentation of expected number of jobs $L\{\mathbf{N_r}, t\}$ and throughput at any instant t, $E\{T(t)\}$ for different values of threshold values, arrival rates, service rates, reneging parameter, joining rates has been done in figs. 2-13.

In figs. 2-4 by fixing the parameters r=3, $\lambda = 7, \beta_0 = .9, \beta_1 = .85, \beta_2 = .75, \beta_3 = .6, \mu_0 = 6, \mu_1 = 2.5, \mu_2 = 1.5, \mu_3 = 1.5, \alpha_0 = .7$, we illustrate the effect of t and λ, μ_j, α_j respectively on $L\{\mathbf{N_r}, t\}$. We note that for smaller values of t, $L\{\mathrm{N_r}, t\}$ increases sharply whereas for larger values of t it is almost constant. From fig. 2 we notice that $L\{\mathbf{N_r}, t\}$ increases with λ. Figs. 3 and 4 show the variation in $L\{\mathbf{N_r}, t\}$ for different sets (A.1, A.2, A.3) of μ_j and sets (B.1, B.2, B.3) of α_j, respectively. By fixing other parameters, we observe that $L\{\mathbf{N_r}, t\}$ decreases for increasing values of μ_j and α_j.

To demonstrate the effect of β, N, N_j on $L\{\mathbf{N_r}, t\}$ in figs. 5-7, respectively, we fix parameters as r=3, $\lambda = .2, \beta_0 = .9, \beta_1 = .8, \beta_2 = .7, \beta_3 = .6, \mu_0 = 1.2, \mu_1 = 1.2, \mu_2 = .5, \alpha_0 = .7, \alpha_1 = 1.2, \alpha_2 = 2.2, \alpha_3 = 3.2$. In fig. 5 we illustrate the effect of sets (C.1, C.2, C.3) of β_j on $L\{\mathbf{N_r}, t\}$ and notice that $L\{\mathbf{N_r}, t\}$ increases very slightly and for lower values of β_j the effect is almost negligible. From figs. 6 and 7 we note that initially as time increases, $L\{\mathbf{N_r}, t\}$ increases sharply but for increasing values of t there is a slight change; $L\{\mathbf{N_r}, t\}$ also increases with increasing values of N and different sets (D.1, D.2, D.3) of threshold levels N_j as expected in physical situations.

In figs. 8-10, we plot the variation of throughput vs. t for different values of λ, μ_j, α_j, respectively. It is observed that $E\{T(t)\}$ increases with the increase in t, λ, μ_j, α_j. It is found that $E\{T(t)\}$ increases sharply for λ but moderately for different sets (F.1, F.2, F.3) of μ_j and different sets (E.1, E.2, E.3) of α_j; the effect of t is more prevalent for smaller values of t. In fig. 11 we observe the pattern of $E\{T(t)\}$ by taking sets (C.1, C.2, C.3) of β_j; it is noted that $E\{T(t)\}$ increases very slightly by increasing the values of β_j and for smaller values of t the effect is almost negligible. Figs. 12 and 13 demonstrate the effect of N and N_j (sets D.1, D.2, D.3) respectively on $E\{T(t)\}$; it is noticed that $E\{T(t)\}$ decreases with the increase in threshold values.

- Overall we conclude that at an instant t, the expected number of jobs $L\{\mathbf{N_r}, t\}$ and throughput of the system $E\{T(t)\}$ both increase sharply as time t increases and become almost constant after some time, which is in agreement with real time situations.
- Both the expected number of jobs $L\{\mathrm{N_r}, t\}$ and average delay time τ(t) at any instant t increase with the increasing values of λ, β_j, Nj. However, the increase in μ_j,α_j results in a decrease in $L\{\mathrm{N_r}, t\}$ as well as in τ(t).
- The throughput of the system $E\{T(t)\}$ at time t increases as the parameters λ, μ_j, β_j, α_j increase but the opposite trend with the increasing values of threshold levels N and N_j are quite visible.

VII. DISCUSSION

In this paper we have obtained the transient probabilities and other performance indices for Markovian queue with optimal threshold policy. The incorporation of threshold policy is profitable with the economic point of view. The backlog of customers can be reduced to the desired level with the provision of additional servers. With the provision of additional servers at predefined threshold levels, a system designer may enable to ensure the high grade of service (GOS) at optimum cost. The combination of additional servers and discouragement may also be of great importance in many real life queueing situations such as computer, telecommunication, manufacturing, etc..

REFERENCES

[1] M. O. Abou El Ata, and A. I. Shawky, "The single server Markovian overflow queue with balking, reneging and additional server for longer queues," *Microelectronics and Reliability*, vol. 32, no. 12, pp. 1389-1394, 1992.

[2] R. O. Al-Seedy, A. A. El-Sherbiny, S. A. El-Shehawy, and S. I. Ammar, "Transient solution of the M/M/c queue with balking and reneging," *Computers & Mathematics with Applications*, vol. 57, no. 8, pp. 1280-1285, 2009.

[3] S. I. Ammar, M. M. Helan, and F. T. Al Amri, "The busy period of an M/M/1 queue with balking and reneging," *Applied Mathematical Modelling*, vol. 37, no. 22, pp. 9223-9229, 2013.

[4] R. Arumunganathan, and S. Jeyakumar, "Steady state analysis of a bulk queue with multiple vacations, set up times with N-policy and closedown times," *Applied Mathematical Modelling*, vol. 29, pp. 972-986, 2005.

[5] K. U. Chandrika, Transient analysis of a system with queue dependent servers, *OPSEARCH*, vol. 43, no. 2, pp. 178-189, 2006.

[6] G. Chaudhury, and M. Paul, "A batch arrival queue with an additional service channel under N-policy," *Applied Mathematical Modelling*, vol. 156, no. 1, pp. 115-130, 2004.

[7] A. G. H. Díaz, and P. Moreno, "A discrete-time single-server queueing system with an N-policy, an early setup and a generalization of the Bernoulli feedback", *Mathematical and Computer Modelling: An International Journal*, vol. 49, no. 5-6, pp. 977-990, 2009.

[8] S. D. Woolford, and G. Douglas, "A preemptive priority queue with balking," *European Journal of Operational Research*, vol. 164, no. 2, pp. 387-401, 2005.

[9] D. Guha, V. Goswami, and A. D. Banik, "Algorithmic computation of steady-state probabilities in an almost observable GI/M/c queue with or without vacations under state dependent balking and reneging," *Applied Mathematical Modelling*, vol. 40, no. 5-6, pp. 4199-4219, 2016.

[10] Y. C. Hsieh, and K. H. Wang, "Reliability of a repairable system with spares and a removable repairman," *Microelectronics Reliability*, vol. 35, no. 2, pp. 197-207, 1995.

[11] S. Hur, J. Kim, and C. Kang, "An analysis of the M/G/1 system with N and T policy", *Applied Mathematical Modelling*, vol. 27, no. 8, pp. 665-675, 2003.

[12] M. Jain, "Optimal N-policy for single server Markovian queue with breakdown, repair and state dependent arrival rate," *International Journal of Management Systems*, vol. 13, no. 3, pp. 245-260, 1997.

[13] M. Jain, "M/M/m queue with discouragement and additional servers," *Journal of German Studies Review*, vol. 36, no. 1-2, pp. 31-42, 1998.

[14] M. Jain, "N- Policy for redundant repairable system with additional repairman," *OPSEARCH*, vol. 40, no. 2, pp. 97-114, 2003.

[15] M. Jain, Rakhee, and S. Maheshwari, "N policy for a machine repair system with spares and reneging," *Applied Mathematical Modelling*, vol. 28, no. 6, pp. 513-531, 2004.

[16] M. Jain, G. C. Sharma, and N. Singh, "Transient analysis of M/M/R machining system with mixed standbys, switching failures, balking, reneging and additional removable repairmen," *International Journal of Engineering*, vol. 20, no. 2, pp. 169-182, 2007.

[17] M. Jain, G. C. Sharma, and M. Singh, "M/M/R machine interference model with balking, reneging, spares and two modes of failure," *OPSEARCH*, vol. 40, no. 1, pp. 24-41, 2003.

[18] M. Jain, M. Singh, and K. P. S. Baghel, "M/M/C/K/N machine repair problem with balking, reneging, spares and additional repairman," *Journal of German Studies Review*, vol. 26-27, pp. 49-60, 2000.

[19] C. Kim, and A. Dudin, "Analysis of a queueing model with contingent additional server," In Gaj P., Kwiecień A., Stera P. (eds.) *Computer Networks*, Communications in Computer and Information Science, vol. 608, 2016. Springer, Cham.

[20] D. H. Lee, and W. S. Yang, "The N-policy of a discrete time Geo/G/1 queue with disasters and its application to wireless sensor networks," *Applied Mathematical Modelling*, vol. 37, no. 23, pp. 9722-9731, 2013.

[21] P. R. Parthasarathy, and R. Sudhesh, "Time-dependent analysis of a single server retrial queue with state dependent rates," *Operations Research Letters*, vol. 35, no. 5, pp. 601-611, 2007.

[22] A. I. Pazgal, and S. Radas, "Comparison of customer balking and reneging behavior to queueing theory predictions: An experimental study**,"** *Computers and Operations Research*, vol. 35, no. 8, pp. 2537-2548, 2008.

[23] J. F. Reynolds, "The stationary solution of a multi-server model with discouragement," *Operations Research*, vol. 16, pp. 64-71, 1968.

[24] J. A. Schwarz, G. Selinka, and R. Stolletz, "Performance analysis of time-dependent queueing systems: Survey and classification," *Omega*, vol. 63, pp. 170-189, 2016.

[25] O. P. Sharma, *Markovian Queues*, Ellis Horwood, London, 1990.

[26] O. P. Sharma, and A. M. K. Tarabia, "A simple transient analysis of an M/M/1/N queue," *Sankhya: The Indian Journal of Statistics*, vol. 62, no 2, pp. 273-281, 2000.

[27] A. I. Shawky, "The single server machine interference model with balking, reneging and an additional server for longer queues," *Microelectronics Reliability*, vol. 37, no. 1, pp. 355-357, 1997.

[28] Shawky, A.I. (2000), "The machine interference model: M/M/C/K/N with balking, reneging and spares", *OPSEARCH*, vol. 37, no. 1, pp. 25-35, 2000.

[29] V. Sridaran, and P. R. Jayashree, "A note on the transient behavior and expected profit of a M/M/R machine repair problem with spares," *Stochastic Processes and their Applications*, (eds. A. Vijayakumar and M. Sreenivasan), Narosa Pub. House, New Delhi, India, pp. 215-231, 1999.

[30] W. Sun, S. Li, and E. Cheng-Guo, "Equilibrium and optimal balking strategies of customers in Markovian queues with multiple vacations and N-policy," *Applied Mathematical Modelling*, vol. 40, no. 1, pp. 284-301, 2016.

[31] N. Tian, and Z. G. Zhang, "A two-threshold vacation policy in queueing systems," *European Journal of Operational Research*, vol. 168, no. 1, pp. 153-163, 2004.

[32] J. Wang, X. Zhang, and P. Huang, "Strategic behavior and social optimization in a constant retrial queue

with the N-policy," *European Journal of Operational Research*, vol. 256, no. 3, pp. 841-849, 2017.

[33] K. H. Wang, J. B. Ke, and J. C. Ke, "Profit analysis of the M/M/R machine repair problem with balking, reneging and standby switching failures," *Computers and Operations Research*, vol. 34, no. 3, pp. 835-847, 2007a.

[34] K. H. Wang, T. Y. Wang, and W. L. Pearn, "Optimal control of the N policy M/G/1 queueing system with server breakdowns and general startup times", *Applied Mathematical Modelling*, vol. 31, no. 10, pp. 2199-2212, 2007b.

[35] K. H. Wang, "Optimal operation of a Markovian queueing system with a removable and non -reliable server", *Microelectronics Reliability*, vol. 35, no. 8, pp. 1131-1136, 1995.

[36] M. Yadin, and P. Naor, "Queueing system with removable service station," *Opl. Research Quarterly*, vol. 14, pp. 393-405, 1963.

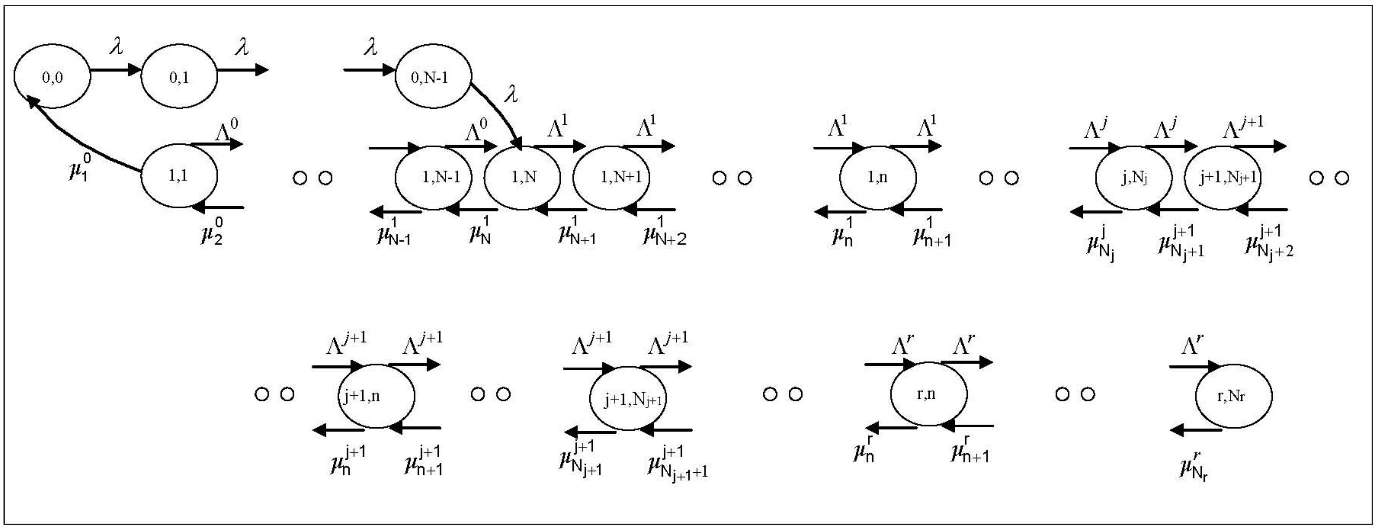

Fig. 1: State Transition Diagram

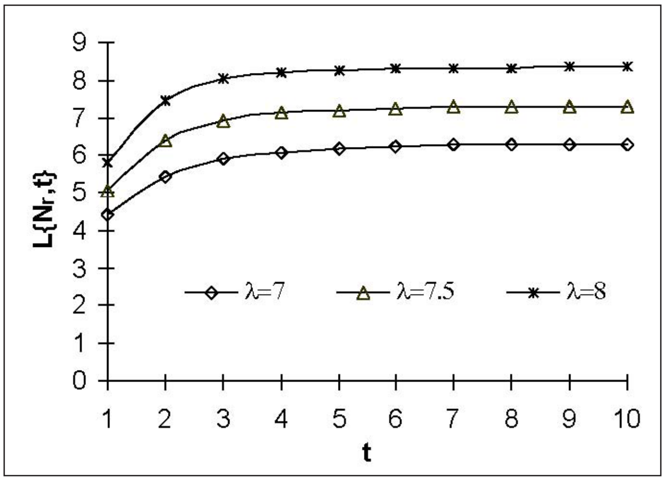

Fig. 2: L{Nr,t} vs. t for Different Values of λ

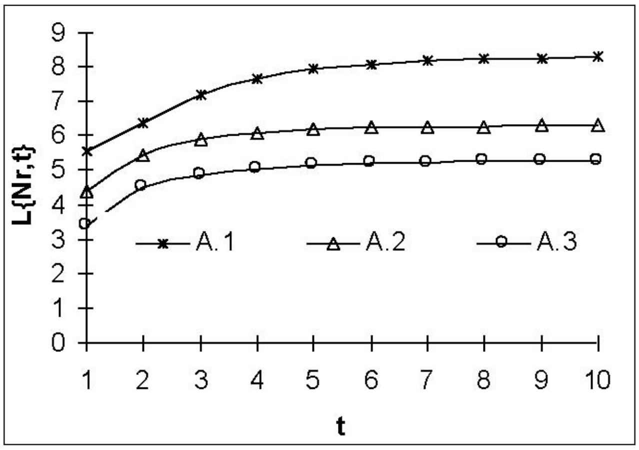

Fig. 3: L{Nr,t} vs. t for Different Values of μ_j

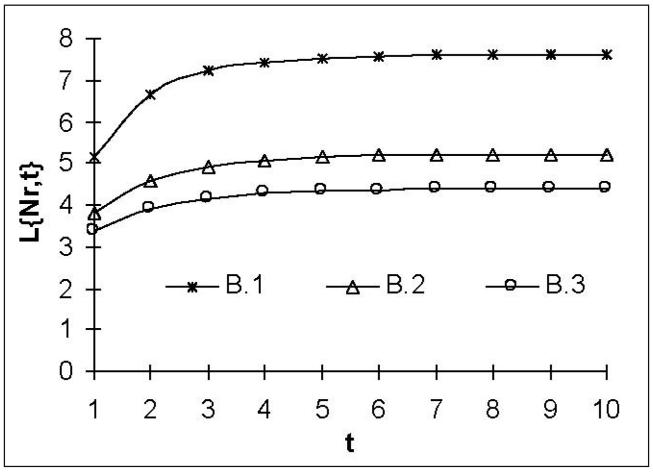

Fig. 4: L{Nr,t} vs. t for Different Values of α_j

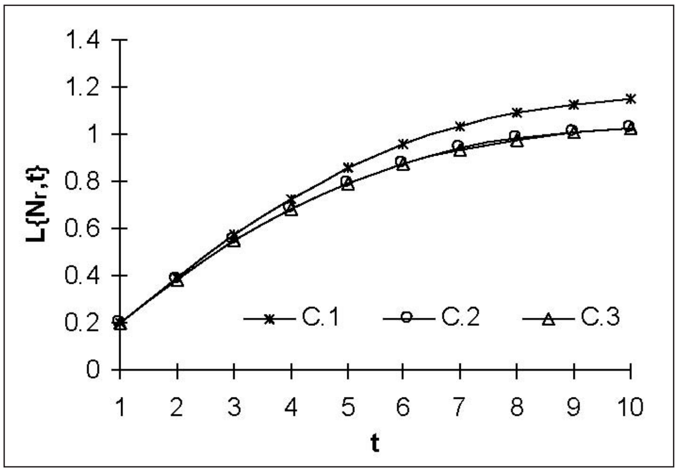

Fig. 5: L{Nr,t} vs. t for Different Values of β_j

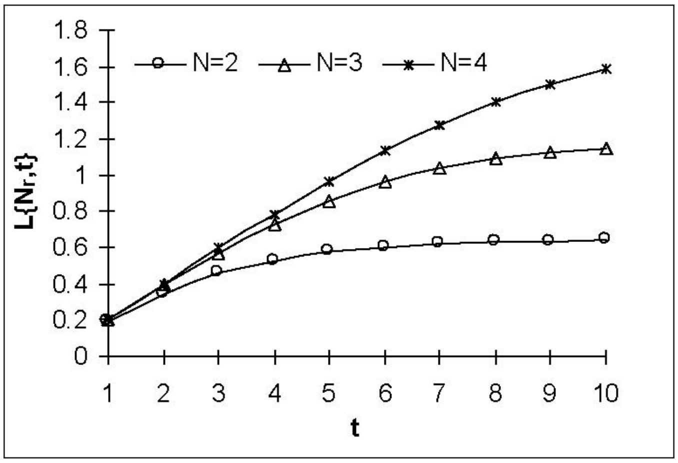

Fig. 6: L{Nr,t} vs. t for Different Values of N

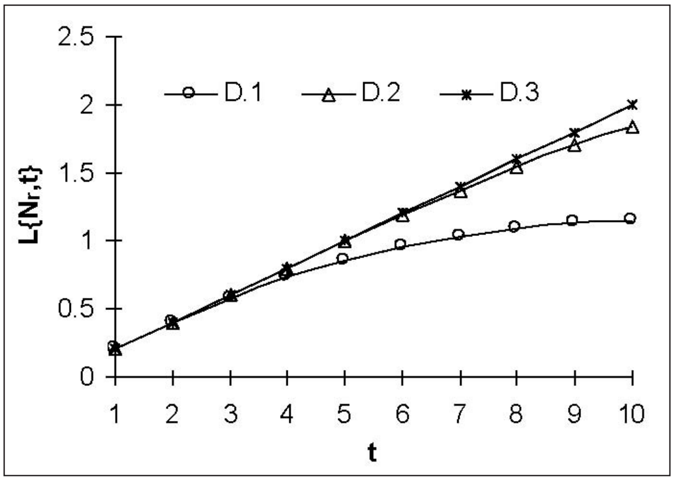

Fig. 7: L{Nr,t} vs. t for Different Values of N_j

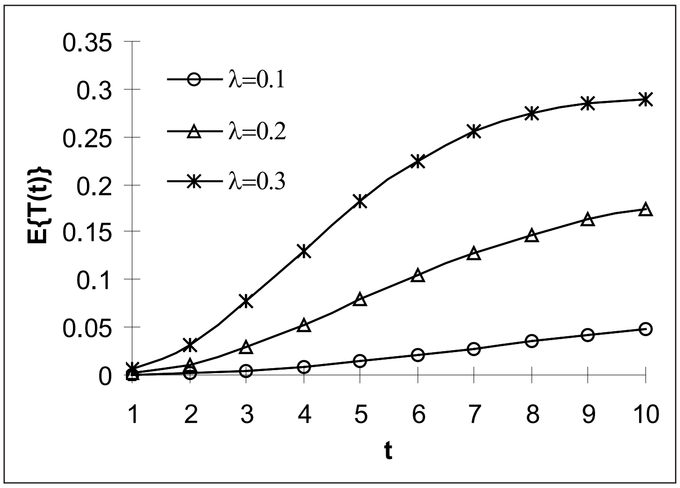

Fig. 8: E{T(t)} vs. t for Different Values of λ

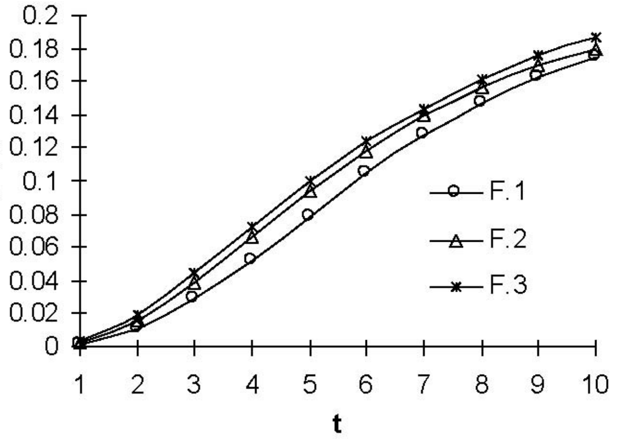

Fig. 9: E{T(t)} vs. t for Different Values of μ_j

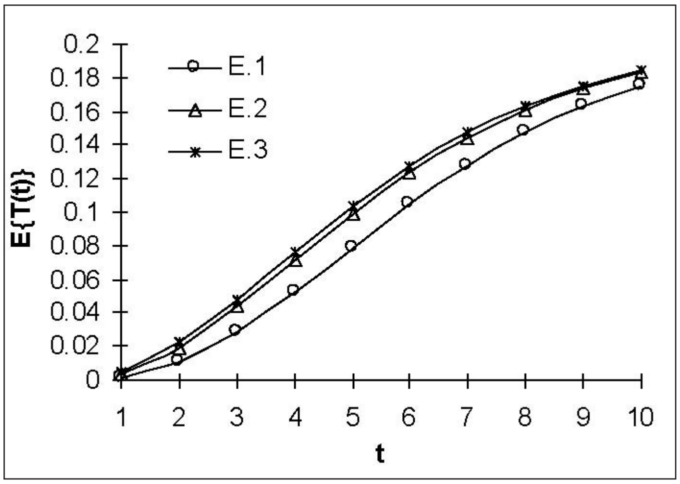

Fig. 10: E{T(t)} vs. t for Different Values of α_j

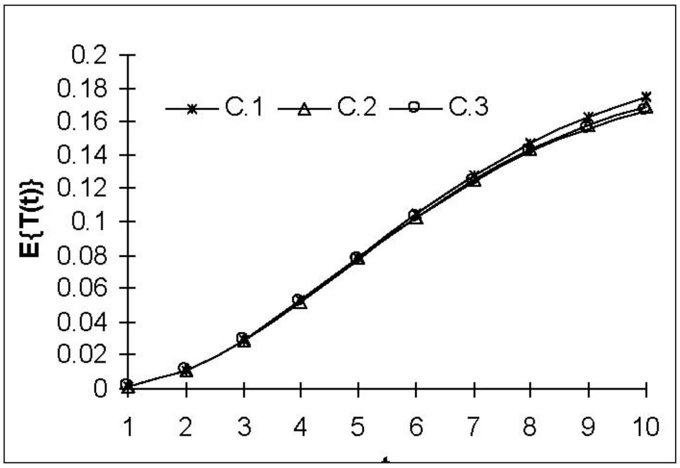

Fig. 11: E{T(t)} vs. t for Different Values of β_j

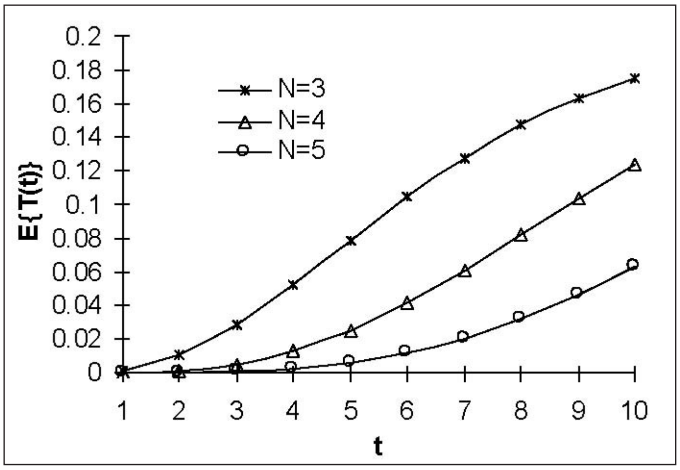

Fig. 12: E{T(t)} vs. t for Different Values of N

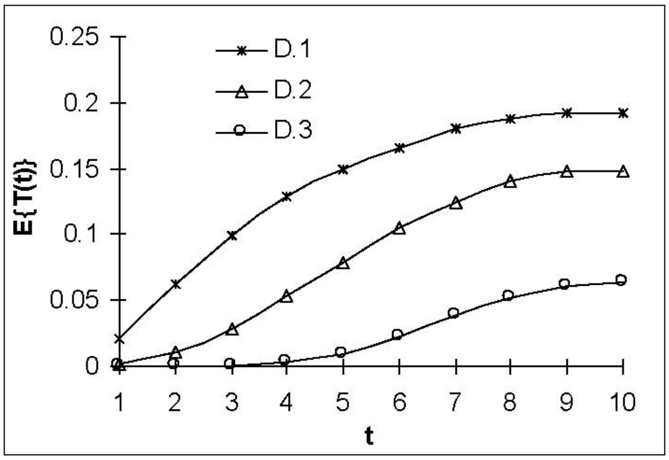

Fig. 13: E{T(t)} vs. t for Different Values of N_j

TABLE 2: AVERAGE DELAY TIME τ(T) BY VARYING ARRIVAL RATE (λ) AND N

		N=3			N=4			N=5	
t	$\lambda = 7$	$\lambda = 7.5$	$\lambda = 8$	$\lambda = 7$	$\lambda = 7.5$	$\lambda = 8$	$\lambda = 7$	$\lambda = 7.5$	$\lambda = 8$
1	0.7611	0.8065	0.8507	0.7112	0.7504	0.7964	0.7526	0.7632	0.7867
2	1.0239	1.0792	1.1318	1.0046	1.0885	1.1774	0.9184	1.0180	1.1378
3	1.1411	1.2113	1.2844	1.1661	1.2907	1.4248	1.0929	1.2585	1.4762
4	1.2054	1.2889	1.3812	1.2669	1.4210	1.5902	1.1983	1.4312	1.7509
5	1.2416	1.3340	1.4396	1.3265	1.5002	1.6931	1.2650	1.5471	1.9489
6	1.2619	1.3597	1.4731	1.3612	1.5469	1.7543	1.3039	1.6192	2.0789
7	1.2732	1.3739	1.4916	1.3809	1.5738	1.7894	1.3260	1.6619	2.1589
8	1.2794	1.3817	1.5016	1.3921	1.5890	1.8093	1.3382	1.6864	2.2060
9	1.2829	1.3860	1.5070	1.3983	1.5975	1.8203	1.3449	1.7002	2.2331
10	1.2848	1.3883	1.5099	1.4018	1.6023	1.8265	1.3485	1.7078	2.2484

Table 3: Average Delay time $\tau(t)$ by Varying Service Rate (μ_j) and N

		N=3			N=4			N=5	
t	μ_F.1	μ_F.2	μ_F.3	μ_F.1	μ_F.2	μ_F.3	μ_F.1	μ_F.2	μ_F.3
1	0.8351	0.7441	0.6695	0.8222	0.7166	0.6334	0.9065	0.7810	0.6836
2	1.2073	1.0629	0.9361	1.1777	1.0399	0.9199	1.0717	0.9594	0.8709
3	1.4563	1.2430	1.0655	1.5051	1.2693	1.0728	1.4639	1.2289	1.0342
4	1.6485	1.3600	1.1394	1.7835	1.4360	1.1713	1.8119	1.4225	1.1369
5	1.7940	1.4361	1.1824	2.0020	1.5485	1.2301	2.1314	1.5662	1.1986
6	1.8982	1.4842	1.2074	2.1634	1.6215	1.2644	2.3933	1.6621	1.2332
7	1.9690	1.5139	1.2217	2.2769	1.6675	1.2839	2.5930	1.7228	1.2520
8	2.0153	1.5319	1.2298	2.3538	1.6957	1.2948	2.7360	1.7597	1.2619
9	2.0450	1.5427	1.2344	2.4045	1.7128	1.3009	2.8340	1.7815	1.2670
10	2.0636	1.5491	1.2370	2.4373	1.7230	1.3042	2.8989	1.7942	1.2697

Table 4: Average Delay Time $\tau(t)$ by Varying Reneging Rate (α_j) and N

		N=3			N=4			N=5	
t	α_E.1	α_E.2	α_E.3	α_E.1	α_E.2	α_E.3	α_E.1	α_E.2	α_E.3
1	0.7611	0.7828	0.8055	0.7112	0.7400	0.7714	0.7526	0.7899	0.8317
2	1.0239	1.0689	1.1161	1.0046	1.0562	1.1101	0.9184	0.9742	1.0356
3	1.1411	1.2098	1.2861	1.1661	1.2519	1.3449	1.0929	1.1998	1.3224
4	1.2054	1.2939	1.3976	1.2669	1.3834	1.5160	1.1983	1.3539	1.5445
5	1.2416	1.3448	1.4706	1.3265	1.4667	1.6328	1.2650	1.4617	1.7173
6	1.2619	1.3749	1.5167	1.3612	1.5181	1.7097	1.3039	1.5309	1.8400
7	1.2732	1.3925	1.5451	1.3809	1.5492	1.7588	1.3260	1.5736	1.9226
8	1.2794	1.4027	1.5623	1.3921	1.5676	1.7894	1.3382	1.5990	1.9759
9	1.2829	1.4086	1.5725	1.3983	1.5785	1.8084	1.3449	1.6139	2.0094
10	1.2848	1.4119	1.5786	1.4018	1.5848	1.8199	1.3485	1.6225	2.0301

Table 5: Average Delay Time $\tau(t)$ by Varying (β_j) and N

		N=3			N=4			N=5	
t	β_C.1	β_C.2	β_C.3	β_C.1	β_C.2	β_C.3	β_C.1	β_C.2	β_C.3
1	0.7745	0.7611	0.7441	0.7714	0.7444	0.7166	0.8317	0.8075	0.7810
2	1.0886	1.0239	1.0629	1.1101	1.0732	1.0399	1.0356	0.9955	0.9594
3	1.2636	1.1411	1.2430	1.3449	1.3048	1.2693	1.3224	1.2718	1.2289
4	1.3780	1.2054	1.3600	1.5160	1.4737	1.4360	1.5445	1.4788	1.4225
5	1.4527	1.2416	1.4361	1.6328	1.5888	1.5485	1.7173	1.6369	1.5662
6	1.5001	1.2619	1.4842	1.7097	1.6644	1.6215	1.8400	1.7464	1.6621
7	1.5293	1.2732	1.5139	1.7588	1.7125	1.6675	1.9226	1.8182	1.7228
8	1.5471	1.2794	1.5319	1.7894	1.7424	1.6957	1.9759	1.8635	1.7597
9	1.5578	1.2829	1.5427	1.8084	1.7608	1.7128	2.0094	1.8912	1.7815
10	1.5642	1.2848	1.5491	1.8199	1.7719	1.7230	2.0301	1.9079	1.7942