# Finite Queueing Model with Multitask Servers and Blocking

**Chandra Shekhar[1,\*], Madhu Jain[2]**

[1]Department of Mathematics, Birla Institute of Technology and Science, Pilani Campus, Pilani, Rajasthan, 333 031, India
[2]Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee, Uttarakhand, 247 667, India

**Abstract**  This paper deals with a Markovian queueing system having a multi-task service counters and finite queue in front of each counter. The total service of a customer is completed in three stages provided by two servers at three counters. The first server (S1) can serve the counter I and III alternatively, whereas second server (S2) provides the service at counter II. The server (S1) gives the priority to the customers waiting for stage third for the service at counter III since they are in last phase of service completion. The steady state queue size distribution has been obtained. The expressions for mean number of customers in the system, average queue length and blocking probability have been obtained. Sensitivity analysis has been carried out to study the effect of variation of different parameters.

**Keywords**  Queue, Tandem, Multi-task, Markov Process, Queue Size Distribution, Blocking Probability

## 1. Introduction

Queueing theory has been widely used to model various service systems to provide Quality of Service (QoS) dealing with many congestion situations of day – to - day life. Queueing situations at any service station provide a variety of fascinated challenges to the mathematicians who are interested in applications of probability theory and optimization. Question of interest in the design of optimal service system is to predict the average delay at various interacting and sequenced counters managed by multitask servers and by incorporating customer's behaviour. In such a system formation of tandem queues, blocking is very common. Tandem queues with blocking have been addressed by many theorists and researchers. Optimal ordering of stations in tandem queue with blocking was considered by Ding and Greenberg[1]. Altiok[2], Altiok and Perros[3], Brandwajn and Jow[4] developed algorithms which can be used for the analysis of tandem networks with exponential service times and Poisson external arrivals. Morris and Perros[5] presented a discrete-time tandem network of cut-through queues and introduced a new bursty arrival process. They analyzed the tandem network using single-node decomposition. Charles and LaPadula[6] analyzed the performance of multi-queue single-server systems consisting of many queues. They approximated polling system consisting of several queues. The explicit

time – dependent distribution of the number of tasks for a parallel multi-processor system with task-splitting and feedback was obtained by Haghighi[7]. Zhang and Tian[8] studied a Markovian queueing system with multi-task servers where each server can perform two types of job. Using the matrix analytic method, they provided a new computational algorithm for the stationary distributions of the queue length and waiting time. The analysis of networks of queues under repetitive service blocking mechanism had been presented by Awan et al.[9]. They assumed that the nodes are connected according to an arbitrary configuration and each node in the networks employs an active queue management (AQM) based queueing policy to guarantee certain quality of service for multiple class external traffic. Bhaskar and Lallement[10] investigated supply chain which is represented as a two-input, three-stage queueing network. They computed the minimum response time for the delivery of items to the final destination along the three stages of the network. Yang et al.[11] considered a finite capacity M/M/R queue with second optional channel. Using the matrix - geometric method, they obtained the steady-state probability distributions and various system performance measures.

Intermediate buffer space plays a very important role in managing tandem queues which arise due to variation in service facility at different counters. Kavusturucu and Gupta[12] modeled finite buffer tandem manufacturing system under N policy using open queueing networks. They computed the throughput of the system using decomposition, isolation and expansion methodologies. Vidalis and Papadopoulos[13] obtained the exact solution of the large sparse linear system by the use of the Gauss–Seidel method for reliable multi-station series queueing networks where

buffers of non-identical finite capacities are allowed between successive stations. Aweya et al.[14] proposed a new active queue management scheme for a network device (e.g. router, switch, etc.) with a shared buffer where the buffer is logically organized into multiple queues. Sharma and Virtamo[15] considered a queue with finite buffer where the buffer size limits the amount of work that can be stored in the queue. They obtained the stability, the rates of convergence to the stationary distribution and functional limit theorems for this system. In addition, they also obtained algorithms to compute the stationary density of the workload process, the waiting times and the probability of packet loss. Chydzinski and Winiarczyk[16] investigated the blocking probability in a finite–buffer queue whose arrival process is given by the batch Markovian arrival process. Using the supplementary variable and imbedded Markov chain techniques, Goswami et al.[17] obtained the queue-length distributions at pre-arrival and arbitrary epochs for finite - buffer multi - server bulk - service queueing system. Diamantidis and Papadopoulos[18] used exact Markovian analysis to examine the model of a serial flow line with two workstations and an intermediate buffer where each workstation has multiple unreliable and non-identical parallel machines.

The behavior of the customer in the queue is very important factor for analyzing the quality of service in any service system. Wang et al.[19] developed profit model to determine the optimal number of servers in the system under the constraint of impatience behaviors of the customers. Yang and Choo[20] considered an M/M/s queue with balking, reneging and retrials. They developed an algorithm for the stationary distribution of the Markov chain and presented some numerical results. Al-Seedy et al.[21] used generating function technique to obtain the transient solution for the M/M/c queue with balking and reneging. Perel and Yechiali[22] studied M/M/c queues (c=1, 1<c<∞ and c=∞) in a two-phase (fast and slow) Markovian random environment, with impatient customers. Recently there has been considerable interest in state dependent queueing models. In real life situations, the customer's behaviour may play a deciding role on the service rate of a server. The input traffic may also be influenced by the status of the servers. In the last few years, researches on queueing network theory have been redirected by a series of brilliant and sobering examples. Queues with state dependent arrival rate have wide applications in computers and communication system, production processes, etc. Shagon[23] studied a single server queueing model wherein arrival rate depends on the server status. Using approximation for the general service time distribution by phase type distribution, M/G/1 queue with queue length dependent arrival rate was investigated by Gong et al.[24]. A state dependent queueing model wherein the service rate is adjusted at the beginning of service was studied by Wang[25]. Jain[26] studied optimal N-policy for state dependent Markovian queue with single removable and unreliable server. Hwang et al.[27] proposed a Markov decision theoretic frame-work for a multi-rate network by

approximating a single state-dependent Poisson arrival stream without incurring any significant increase in computational complexity. Jain et al.[28] analyzed the finite queue-dependent heterogeneous multiprocessor service system in which processors are shared by more than one job. They obtained steady-state queue size distribution using recursive method considering Markovian arrival and service times. Lee and Kim[29] considered an M/G/1 queueing system where the speed of the server depends on the amount of work present in the system. By using the level crossing theory and solving the corresponding integral equation, they obtained the stationary distribution of the workload in the system in explicit form. For a single-server retrial queue with state dependent exponential inter-arrival, service and inter-retrial times, Parthasarathy and Sudesh[30] studied the time dependent system size probabilities by employing the continued fractions. Wall and Worthington[31] considered the time dependent behavior of virtual waiting time for modeling of approximate time dependent behavior of queue length of the form M(t)/G/c queueing model. Soares and Latouche[32] developed various models of fluid queues with a level dependency component where the behavior of the phase process changes when the level crosses certain thresholds, as well as the rate at which fluid increases or decreases. Banik[33] considered an infinite-buffer single - server queue with renewal input. Lee[34] considered a class of stochastic networks with state-dependent arrival and service rates. Under the uniform (in state) stability condition, he established several moment stability properties of the system. Banerjee and Gupta[35] analyzed single-server finite-buffer queue where customers arrive according to Poisson process and served in batches of minimum size and maximum threshold limit. Louvel et al.[36] recently presented a non - intrusive and adaptable resource management framework which was developed upon a customized architecture. Zhou et al.[37] gave exact analysis of two-stage tandem queueing network with MAP inputs and buffer sharing using matrix filtration technique. MAP/M/2 queueing system in steady - state was analyzed by Chakravarthy and Karatz[38] with two identical servers in parallel system and pure space sharing among rigid jobs using matrix analytic method.

In this paper we deal with Markovian queueing system wherein jobs are served by two multi-task servers at three counters. The input traffic is assumed to be effected by the status of the server. The state-dependent server may provide service with faster rate to reduce the backlog in case of long queue. The organization of the paper is as follows. The description of the model and underlying assumptions and notations are given in section 2. In section 3 governing equations and queue size distribution are established. The expressions for mean number of jobs in the system and blocking probability are obtained in section 4. The numerical algorithm to develop a computer program is discussed in section 5. Numerical illustrations and sensitivity analysis are also provided. Concluding remarks and scopes of future work are given in the last section 6.
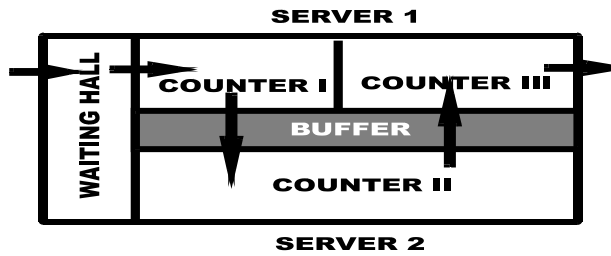
## 2. System Description



**Figure 1.**    Three counters service system with two servers

In the present investigation, we consider a three counters service system served by two multi-purpose servers as shown in Figure 1. For modeling purpose the following assumptions and notations are used:

• The arrival of jobs for service at counter I follows Poisson distribution with rate λ. The jobs are independent to each other.

• The arriving jobs may balk only at counter 1 with probability β (0<β<1) when queue size reaches to a threshold value $N_1$ ($N_1$>1).

• The job entering in the system has to be served by all (three) counters i.e. the jobs form tandem queue. It is not allowed to leave in between without getting complete service. The flow of jobs in the service facility is shown in the figure 1.

• Counters I and III are operated by the first server (S1) whereas second server (S2) operates single counter II. Both servers are independent to each others.

• The service times at counters I, II, and III are exponential distributed with rates $\mu_1$, $\mu_2$, and $\mu_3$, respectively.

The state-dependent servers at counter II and III switch to faster rate $\mu_2'$ and $\mu_3'$ respectively, if there are more than $N_2$ ($N_2$>1) jobs at counter II and at least one job at counter I respectively.

• The discipline for service at all counters is First Come First Served (FCFS).

• The queue is allowed to be formed in front of each counter within the buffer and waiting space capacity of system i.e. at most M jobs are allowed in the system.

• The times taken for transaction of jobs between counters in sequence are assumed to be negligible.

• The switching time of first server (S1) between counter I and counter III is also negligible.

The following notations are used for mathematical formulation of birth-death process:

$P_{0,0,0}$ Steady-state probability that the system is idle i.e. there is no job at any counter.

$P_{i,j,k}$ Steady-state probability that there are i, j and k jobs at counters I, II, and III respectively, for service.

$P_I$ Fraction of time during which the server (S1) is busy at counter I.

$P_{III}$ Fraction of time during which the server (S1) is busy at counter III.

## 3. Mathematical Analysis

The steady-state Chapman-Kolmogrov difference equations governing the present model are as follows:

$$-\lambda P_{0,0,0} + \mu_3 P_{0,0,1} = 0 \tag{1}$$

$$-(\lambda + \mu_3) P_{0,0,k} + \mu_2 P_{0,1,k-1} - \mu_3 P_{0,0,k+1} = 0, 1 \le k \le M - 1 \tag{2}$$

$$\mu_2 P_{0,1,M-1} - \mu_3 P_{0,0,M} = 0 \tag{3}$$

$$-(\lambda + \mu_2) P_{0,j,0} + \mu_1 P_{1,j-1,0} + \mu_3 P_{0,j,1} = 0, 1 \le j \le N_2 \tag{4}$$

$$-(\lambda + \mu_2') P_{0,j,0} + \mu_1 P_{1,j-1,0} + \mu_3 P_{0,j,1} = 0, N_2 < j \le M - 1 \tag{5}$$

$$-\mu_2' P_{0,M,0} + \mu_1 P_{1,M-1,0} = 0, N_2 < M \tag{6}$$

$$-(\lambda + \mu_1) P_{i,0,0} + \lambda P_{i-1,0,0} + \mu_3' P_{i,0,1} = 0, 1 \le i \le N_1 \tag{7}$$

$$-(\lambda\beta + \mu_1) P_{i,0,0} + \lambda P_{i-1,0,0} + \mu_3' P_{i,0,1} = 0, i = N_1 + 1 \tag{8}$$

$$-(\lambda\beta + \mu_1) P_{i,0,0} + \lambda\beta P_{i-1,0,0} + \mu_3' P_{i,0,1} = 0, N_1 + 1 < i \le M - 1 \tag{9}$$

$$-\mu_1 P_{M,0,0} + \lambda\beta P_{M-1,0,0} = 0, N_1 + 1 < M \tag{10}$$

$$-(\lambda + \mu_1 + \mu_2) P_{i,j,0} + \lambda P_{i-1,j,0} + \mu_1 P_{i-1,j-1,0} + \mu_3' P_{i,j,1} = 0,$$
$$1 \le i \le N_1, 1 \le j \le N_2, \ i + j \le M - 1 \tag{11}$$

$$-(\lambda + \mu_1 + \mu_2') P_{i,j,0} + \lambda P_{i-1,j,0} + \mu_1 P_{i-1,j-1,0} + \mu_3' P_{i,j,1} = 0,$$
$$1 \le i \le N_1, \ N_2 < j \le M - 2, \ i + j \le M - 1 \tag{12}$$

$$-(\lambda\beta + \mu_1 + \mu_2) P_{i,j,0} + \lambda P_{i-1,j,0} + \mu_1 P_{i-1,j-1,0} + \mu_3' P_{i,j,1} = 0,$$
$$i = N_1 + 1, 1 \le j \le N_2, \ i + j \le M - 1 \tag{13}$$

$$-\left(\lambda\beta + \mu_1 + \mu_2'\right)P_{i,j,0} + \lambda P_{i-1,j,0} + \mu_1 P_{i-1,j-1,0} + \mu_3' P_{i,j,1} = 0,$$
$$i = N_1 + 1,\, N_2 \le j \le M - 2,\, i + j \le M - 1 \tag{14}$$

$$-\left(\lambda\beta + \mu_1 + \mu_2\right)P_{i,j,0} + \lambda\beta P_{i-1,j,0} + \mu_1 P_{i-1,j-1,0} + \mu_3' P_{i,j,1} = 0,$$
$$N_1 + 1 < i \le M - 2,\, 1 \le j \le N_2,\, i + j \le M - 1 \tag{15}$$

$$-\left(\lambda\beta + \mu_1 + \mu_2'\right)P_{i,j,0} + \lambda\beta P_{i-1,j,0} + \mu_1 P_{i-1,j-1,0} + \mu_3' P_{i,j,1} = 0,$$
$$N_1 + 1 < i \le M - 2,\, N_2 < j \le M - 2,\, i + j \le M - 1 \tag{16}$$

$$-\left(\mu_1 + \mu_2\right)P_{i,M-i,0} + \mu_1 P_{i+1,M-i-1,0} + \lambda P_{i-1,M-i,0} = 0,$$
$$M - N_2 \le i \le N_1 + 1 \tag{17}$$

$$-\left(\mu_1 + \mu_2'\right)P_{i,M-i,0} + \mu_1 P_{i+1,M-i-1,0} + \lambda P_{i-1,M-i,0} = 0$$
$$1 \le i \le N_1 + 1,\, M - i \ge N_2 \tag{18}$$

$$-\left(\mu_1 + \mu_2\right)P_{i,M-i,0} + \mu_1 P_{i+1,M-i-1,0} + \lambda\beta P_{i-1,M-i,0} = 0,$$
$$\max(N_1 + 1,\, M - N_2) \le i \le M - 1 \tag{19}$$

$$-\left(\mu_1 + \mu_2'\right)P_{i,M-i,0} + \mu_1 P_{i+1,M-i-1,0} + \lambda\beta P_{i-1,M-i,0} = 0,$$
$$N_1 + 1 < i \le M - 1,\, M - i \ge N_2 \tag{20}$$

$$-\left(\lambda + \mu_2 + \mu_3\right)P_{0,j,k} + \mu_2 P_{0,j+1,k-1} + \mu_3 P_{o,j,k+1} = 0,$$
$$1 \le j \le N_2 - 2,\, 1 \le k \le M - 2,\, j + k \le M - 1 \tag{21}$$

$$-\left(\lambda + \mu_2 + \mu_3\right)P_{0,j,k} + \mu_2' P_{0,j+1,k-1} + \mu_3 P_{o,j,k+1} = 0,$$
$$j = N_2 - 1,\, 1 \le k \le M - 2,\, j + k \le M - 1 \tag{22}$$

$$-\left(\lambda + \mu'_2 + \mu_3\right)P_{0,j,k} + \mu_2' P_{0,j+1,k-1} + \mu_3 P_{o,j,k+1} = 0,$$
$$N_2 - 1 \le j \le M - 2,\, 1 \le k \le M - 2,\, j + k \le M - 1 \tag{23}$$

$$-\left(\mu_2 + \mu_3\right)P_{0,j,M-j} + \mu_2 P_{0,j+1,M-j-1} = 0,\, 1 \le j \le N_2 - 2 \tag{24}$$

$$-\left(\mu_2 + \mu_3\right)P_{0,j,M-j} + \mu_2' P_{0,j+1,M-j-1} = 0,\, j = N_2 - 1 \tag{25}$$

$$-\left(\mu_2' + \mu_3\right)P_{0,j,M-j} + \mu_2' P_{0,j+1,M-j-1} = 0,\, N_2 - 1 < j \le M - 1 \tag{26}$$

$$-\mu_3' P_{i,0,M-i} + \lambda P_{i-1,0,M-i} + \mu_2 P_{i,1,M-i-1} = 0,\, 1 \le i \le N_1 \tag{27}$$

$$-\mu_3' P_{i,0,M-i} + \lambda\beta P_{i-1,0,M-i} + \mu_2 P_{i,1,M-i-1} = 0,\, N_1 < i \le M - 1 \tag{28}$$

$$-\left(\lambda + \mu_3'\right)P_{i,0,k} + \lambda P_{i-1,0,k} + \mu_2 P_{i,1,K-1} + \mu_3' P_{i,0,k+1} = 0,$$
$$1 \le i \le N_1,\, 1 \le k \le M - 2,\, i + k \le M - 1 \tag{29}$$

$$-\left(\lambda\beta + \mu_3'\right)P_{i,0,k} + \lambda P_{i-1,0,k} + \mu_2 P_{i,1,K-1} + \mu_3' P_{i,0,k+1} = 0,$$
$$i = N_1 + 1,\, 1 \le k \le M - 2,\, i + k \le M - 1 \tag{30}$$

$$-\left(\lambda\beta + \mu_3'\right)P_{i,0,k} + \lambda\beta P_{i-1,0,k} + \mu_2 P_{i,1,K-1} + \mu_3' P_{i,0,k+1} = 0,$$
$$N_1 + 1 < i \le M - 2,\, 1 \le k \le M - 2,\, i + k \le M - 1 \tag{31}$$

$$-\left(\lambda + \mu_2 + \mu_3'\right)P_{i,j,k} + \lambda P_{i-1,j,k} + \mu_2 P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0,$$
$$1 \le i \le N_1,\, 1 \le j \le N_2 - 1,\, 1 \le k \le M - 3,\, i + j + k \le M - 1 \tag{32}$$

$$-\left(\lambda + \mu_2 + \mu_3'\right)P_{i,j,k} + \lambda P_{i-1,j,k} + \mu_2' P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0,$$
$$1 \le i \le N_1,\, j = N_2,\, 1 \le k \le M - 3,\, i + j + k \le M - 1 \tag{33}$$

$$-\left(\lambda + \mu_2' + \mu_3'\right)P_{i,j,k} + \lambda P_{i-1,j,k} + \mu_2' P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0, 1 \le i \le N_1,$$
$$N_2 < j \le M-3, 1 \le k \le M-3, i+j+k \le M-1 \tag{34}$$

$$-\left(\lambda\beta + \mu_2 + \mu_3'\right)P_{i,j,k} + \lambda P_{i-1,j,k} + \mu_2 P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0,$$
$$i = N_1 + 1, 1 \le j \le N_2, 1 \le k \le M-3, i+j+k \le M-1 \tag{35}$$

$$-\left(\lambda\beta + \mu_2 + \mu_3'\right)P_{i,j,k} + \lambda P_{i-1,j,k} + \mu_2' P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0,$$
$$i = N_1 + 1, \ j = N_2, 1 \le k \le M-3, i+j+k \le M-1 \tag{36}$$

$$-\left(\lambda\beta + \mu_2' + \mu_3'\right)P_{i,j,k} + \lambda P_{i-1,j,k} + \mu_2' P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0, i = N_1 + 1,$$
$$N_2 < j \le M-3, 1 \le k \le M-3, i+j+k \le M-1 \tag{37}$$

$$-\left(\lambda\beta + \mu_2 + \mu_3'\right)P_{i,j,k} + \lambda\beta P_{i-1,j,k} + \mu_2 P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0, N_1 + 1 \le M-3,$$
$$1 \le j \le N_2, 1 \le k \le M-3, i+j+k \le M-1 \tag{38}$$

$$-\left(\lambda\beta + \mu_2 + \mu_3'\right)P_{i,j,k} + \lambda\beta P_{i-1,j,k} + \mu_2' P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0,$$
$$N_1 + 1 < i \le M-3, \ j = N_2, 1 \le k \le M-3, i+j+k \le M-1 \tag{39}$$

$$-\left(\lambda\beta + \mu_2' + \mu_3'\right)P_{i,j,k} + \lambda\beta P_{i-1,j,k} + \mu_2' P_{i,j+1,k-1} + \mu_3' P_{i,j,k+1} = 0,$$
$$N_1 + 1 < i \le M-3, N_2 < j \le M-3, 1 \le k \le M-3, i+j+k \le M-1 \tag{40}$$

$$-\left(\mu_2 + \mu_3'\right)P_{i,j,M-i-j} + \lambda P_{i-1,j,M-i-j} + \mu_2 P_{i,j+1,M-i-j-1} = 0,$$
$$1 \le i \le N_1 + 1, 1 \le j \le N_2 - 2, i+j \le M-1 \tag{41}$$

$$-\left(\mu_2 + \mu_3'\right)P_{i,j,M-i-j} + \lambda P_{i-1,j,M-i-j} + \mu_2' P_{i,j+1,M-i-j-1} = 0,$$
$$1 \le i \le N_1 + 1, \ j = N_2 - 1, i+j \le M-1 \tag{42}$$

$$-\left(\mu_2' + \mu_3'\right)P_{i,j,M-i-j} + \lambda P_{i-1,j,M-i-j} + \mu_2' P_{i,j+1,M-i-j-1} = 0,$$
$$1 \le i \le N_1 + 1, N_2 - 1 < j \le M-2, i+j \le M-1 \tag{43}$$

$$-\left(\mu_2 + \mu_3'\right)P_{i,j,M-i-j} + \lambda\beta P_{i-1,j,M-i-j} + \mu_2 P_{i,j+1,M-i-j-1} = 0,$$
$$N_1 + 1 < i \le M-2, 1 \le j \le N_2 - 2, i+j \le M-1 \tag{44}$$

$$-\left(\mu_2 + \mu_3'\right)P_{i,j,M-i-j} + \lambda\beta P_{i-1,j,M-i-j} + \mu_2' P_{i,j+1,M-i-j-1} = 0,$$
$$N_1 + 1 < i \le M-2, \ j = N_2 - 1, i+j \le M-1 \tag{45}$$

$$-\left(\mu_2' + \mu_3'\right)P_{i,j,M-i-j} + \lambda\beta P_{i-1,j,M-i-j} + \mu_2' P_{i,j+1,M-i-j-1} = 0,$$
$$N_1 + 1 < i \le M-2, N_2 - 1 < j \le M-2, i+j \le M-1 \tag{46}$$

The above system of linear equations governing the present model in terms of steady-state probabilities in matrix form can be written as

$$AP = 0 \tag{47}$$

where **A** is square matrix of order (M+3)(M+2)(M+1)/6 whose elements are the coefficients of state probabilities, **P** is column vector of state probabilities and **0** is null column vector of order (M+3)(M+2)(M+1)/6.

Using the normalizing condition

$$\sum\sum\sum P_{i,j,k} = 1 \tag{48}$$

The system of linear equations in (47) can be expressed as

$$\mathbf{A}^*\mathbf{P} = \mathbf{B} \tag{49}$$

where $\mathbf{A}^*$ is the matrix **A** replacing the last row with a row vector having all unit element and **B** is the column vector of the form $[0, 0, \ldots 0, 1]^T$ of order (M+3)(M+2)(M+1)/6. Equation (49) can be solved using matrix-inverse method to obtain steady-state probabilities that can be further used to evaluate various system performance measures.

## 4. The System Performance Measures

For queueing model under consideration, using the steady-state probabilities derived in the previous section, we compute following system performance measures:

(i) The mean number of jobs in the system (L) is

$$L = \sum_{\substack{i,j,k \\ i+j+k=1}} P_{i,j,k} + 2\sum_{\substack{i,j,k \\ i+j+k=2}} P_{i,j,k} + \ldots + M\sum_{\substack{i,j,k \\ i+j+k=M}} P_{i,j,k} \tag{50}$$

(ii) The blocking probability at counter I is obtained by

$$P_I = \sum_{i=0}^{M} \sum_{j=0}^{M} P_{i,j,0}, \qquad i+j \le M \qquad (51)$$

(iii) The blocking probability at counter III is obtained by

$$P_{III} = \sum_{i=0}^{M} \sum_{j=0}^{M} \sum_{k=1}^{M} P_{i,j,k}, \quad i+j+k \le M \qquad (52)$$

(iv) Now the effective arrival rate is

$$\bar{\lambda} = \sum_{i=0}^{M} \lambda_i P_{i,j,k}, 0 \le j \le M, 0 \le k \le M \ and\ i+j+k \le M \ . (53)$$

Here $\lambda_i$ is obtained as $\lambda_i = \begin{cases} \lambda & ; i < N_1 \\ \lambda\beta & ; i \ge N_1 \end{cases}$ .

## 5. Numerical Results

In order to validate the computational tractability, the performance measures established above have been used to generate numerical results. The computer program is developed in software package MATLAB by writing code to compute performance measures. The average number of jobs in the system and blocking probabilities are obtained numerically by varying various parameters and are summarized in Tables 1, 2, and 3.

**Table 1.**  Performance measures by varying λ and M (for β=0.8, μ₁=1.1, μ₂=0.8, μ₃=1.2, N₁=5, N₂=5)

| M | λ | L | $P_I$ | PIII |
|---|---|---|---|---|
|   | 0.1 | 0.6722 | 0.1171 | 0.1230 |
| 5 | 0.5 | 2.7856 | 0.3991 | 0.3385 |
|   | 1.0 | 3.9844 | 0.4902 | 0.3817 |
|   | 0.1 | 1.0446 | 0.1437 | 0.1422 |
| 7 | 0.5 | 3.8779 | 0.4468 | 0.3591 |
|   | 1.0 | 5.6880 | 0.5390 | 0.3986 |
|   | 0.1 | 1.4926 | 0.1673 | 0.1591 |
| 9 | 0.5 | 4.9358 | 0.4729 | 0.3730 |
|   | 1.0 | 7.4062 | 0.5574 | 0.4101 |

In Table 1, average number of jobs in the system and blocking probabilities are displayed with the variation of arrival rate λ from 0.1 to 1.0 for different room capacity (M). The service rates at different counter are fixed as $\mu_1 = 1.1$, $\mu_2 = 0.8$, $\mu_3 = 1.2$. The threshold parameters $N_1$ and $N_2$ for balking are fixed at 5 and the balking rate (β) is 0.8. The value of faster service rates at counter II and III are $\mu'_2 = 1.0, \mu'_3 = 1.4$ respectively. A significant increasing trend are observed in all performance measures L, $P_I$, $P_{III}$ with respect to room capacity (M) and arrival rate (λ) which is true in realistic situation.

**Table 2.**  Performance measures by varying μ and M (for λ=0.5, β=0.8, μ₁=1.1μ, μ₂=0.8 μ, μ₃=1.2 μ, N₁=5, N₂=5)

| M | μ | L | $P_I$ | $P_{III}$ |
|---|---|---|---|---|
|   | 1.0 | 2.7856 | 0.3991 | 0.3385 |
| 5 | 1.6 | 1.9445 | 0.2997 | 0.2846 |
|   | 2.0 | 1.6065 | 0.2544 | 0.2521 |
|   | 1.0 | 3.8779 | 0.4468 | 0.3591 |
| 7 | 1.6 | 2.7706 | 0.3425 | 0.3098 |
|   | 2.0 | 2.3349 | 0.2949 | 0.2784 |
|   | 1.0 | 4.9358 | 0.4729 | 0.3730 |
| 9 | 1.6 | 3.6529 | 0.3724 | 0.3284 |
|   | 2.0 | 3.1405 | 0.3255 | 0.2989 |

In Table 2, various performance measures are evaluated for different service rate where service rate are $\mu_1=1.1\mu$, $\mu_2=0.8\mu$, $\mu_3=1.2\mu$. The value of μ is varying from 1 to 2. The changed service rate for counter II and III are $\mu'_2 = \mu_2 + 0.2$ and $\mu'_3 = \mu_3 + 0.2$ . The arrival rate chosen is λ=0.5. All others parameters are same as in Table 1. An average number of jobs in the system and blocking probabilities are exhibiting decreasing trend with respect to service rate.

**Table 3.**  Performance measures by varying β and M (for λ=0.5, μ₁=1.1, μ₂=0.8, μ₃=1.2, N₁=5, N₂=5)

| M | β | L | PI | $P_{III}$ |
|---|---|---|---|---|
|   | 0.5 | 2.8576 | 0.4079 | 0.3551 |
| 5 | 0.7 | 2.8263 | 0.4029 | 0.3448 |
|   | 0.9 | 2.7237 | 0.3945 | 0.3310 |
|   | 0.5 | 4.0910 | 0.4661 | 0.3687 |
| 7 | 0.7 | 3.9918 | 0.4551 | 0.3634 |
|   | 0.9 | 3.6933 | 0.4361 | 0.3526 |
|   | 0.5 | 5.2785 | 0.4951 | 0.3795 |
| 9 | 0.7 | 5.1226 | 0.4826 | 0.3765 |
|   | 0.9 | 4.6022 | 0.4589 | 0.3664 |

In Table 3, we measure the performance by varying the rate of balking from 0.5 to 0.9 with an increment of 0.2 for arrival rate λ=0.5. All other parameters are same as in Table 1. A prominent decreasing effect is observed in the value of performance measures for higher room capacity with respect to balking probability.

In Figures 2(a)-2(d) the variation of average number of jobs in system is displayed by varying λ, μ, M, and β respectively. The average number of jobs increases with the increase in arrival rate (λ) and room capacity (M) which matches with physical situation. By increasing service rate (μ) and balking probability (β), the average number of jobs reduces but the effect of β is negligible for small room capacity.

These figures exhibit realistic result. Hence, the system designer should provide the efficient servers in service system to check reluctance behavior among customers which directly or indirectly affect on average number of jobs in the system. The present study shows that system capacity is important parameter for timeliness and efficient service. The decision maker should analyze the system's size efficiently since it directly affect service cost and customer's satisfactions.
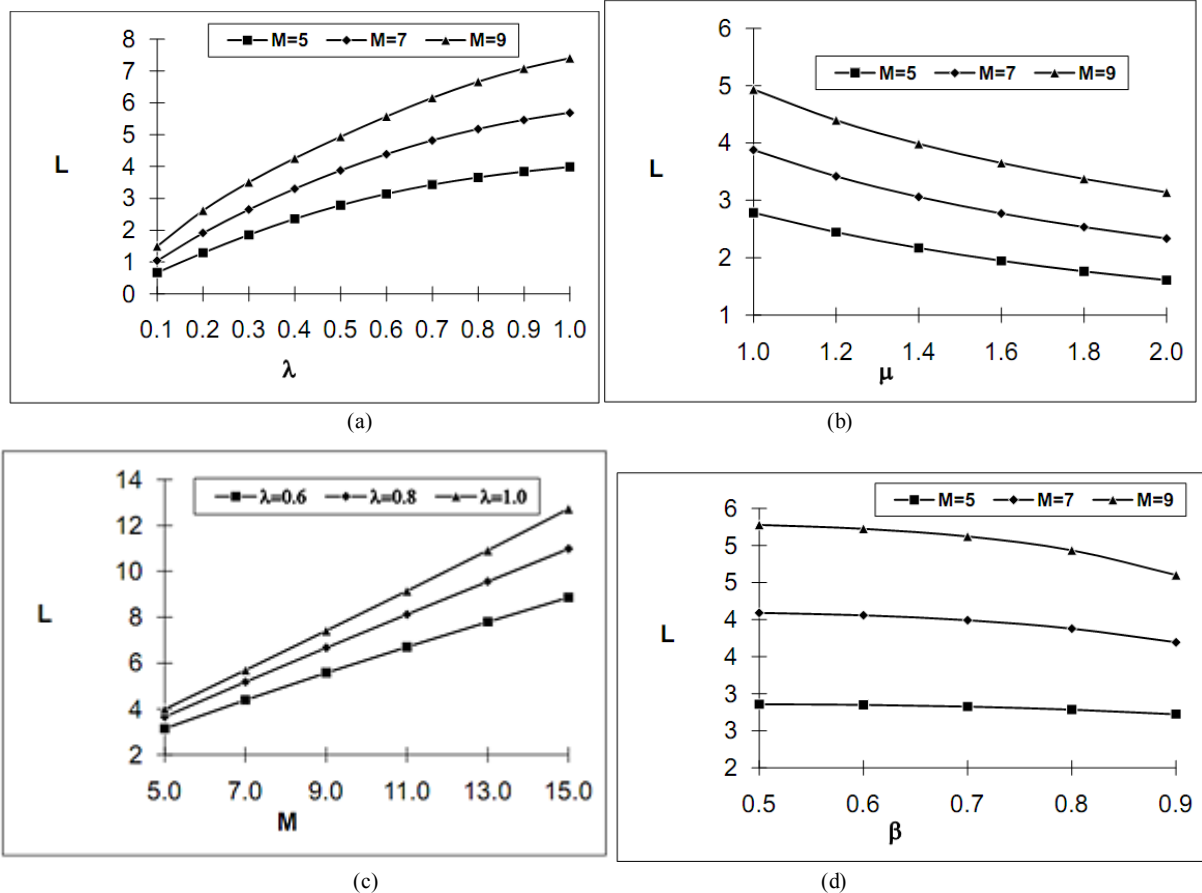
**Figure 2.** Number of jobs in the system

For illustration purpose we give the following example. Consider the repair shop wherein a machine arriving for repair, is first checked by primary repairman. He recommends the jobs to specialized repairman according to the requirement. Finally primary repairman checks the machines and handover it to the customer and takes the payment. Another similar example occurs at bank where the account holder comes to withdraw his money from his account. First he has to go to the cashier counter with cheque and pass book. At this stage cashier checks the account number, name and amount in his account and gives the token to the customer. Cashier then forwards his cheque to senior officer to verify the signature of the customer. After verifying the signature from the specimen copy, the officer returns the cheque back to the cashier. Cashier then calls the customer and takes the token before giving the cash. From the above examples we see that, there may be situations where the job has to be processed by two servers in three stages where the primary server works at counter I and III whereas secondary server provides service at counter II. Each job has to go through all the three counters in a sequential manner. Primary server gives the priority to the customers at counter III as they have already got some service.

## 6. Conclusions

In this paper we have studied the multi-tasking server that oscillates between counters with balking and state-dependent service rate. The service of job is done through three counters. Numerical solution is obtained using matrix method. The performance of the system is measured in terms of average waiting time in the system, average number of jobs in the system and blocking probabilities for different arrival rate, service rate and balking rate. The state dependent rate incorporated for modelling multi-counters system make our results more closer to realistic situation. The system analyst can analyze such model by incorporating associated cost factors and optimize the time fraction between counters and service rates of the servers. However there is scope of extension of this work by considering bulk arrivals and/or bulk service in which direction the attention should be paid. We should go for the transient solution to make our model more realistic. We can analyze present model under shadow of various type of server's policies and customer's policies. We can extend our model for other different kind of architectural restrictions.

## ACKNOWLEDGEMENTS

improving the quality and clarity of the paper.

# REFERENCES

[1] Ding, J., Greenberg, B.S. (1989) Optimal order of server in series with no queue capacity Technical Report, University of Texas, Austin, TX.

[2] Altiok, T. (1982) Approximation analysis of exponential tandem queue with blocking European Journal Operational Research 11, 390-398.

[3] Altiok, T., Peros, H.G. (1986) Open networks of queue with blocking-split and merge configuration IIE Transaction 1986 86, 251-261.

[4] Brandwajn, A., Jow, Y.L. (1988) An approximation method for tandem queues with blocking Operations Research 36, 73-83.

[5] Morris, T.D., Perros, H.G. (1993) Approximate analysis of a discrete-time tandem network of cut-through queues with blocking and bursty traffic Performance Evaluation 17(3), 207-223.

[6] Charles, A., LaPadula, H.L. (1996) Customer delay in very large multi-queue single-server systems Performance Evaluation. 26(3), 201-218.

[7] Haghighi, A.M. (1998) An analysis of the number of tasks in a parallel multi-processor system with task-splitting and feedback Computers and Operations Research 25(11), 941-956.

[8] Zhang, Z.G., Tian, N. (2004) An analysis of queuing systems with multitask servers European Journal of Operational Research. 156(2), 375-389.

[9] Awan, I., Ahmad, B., Ahmad, S. (2007) Performance analysis of networks of queues under active queue management scheme Simulation Modelling Practice and Theory 15(4), 416-425.

[10] Bhaskar, V. and Lallement, P. (2010) Modeling a supply chain using a network of queues Applied Mathematical Modelling 34(8), 2074-2088.

[11] Yang, D.Y., Wang, K.H., Kuo, Y.T. (2011) Economic application in a finite capacity multi-channel queue with second optional channel Applied Mathematics and Computation 217(18), 7412-7419.

[12] Kavusturucu, A., Gupta, S.M.A (1998) A methodology for analyzing finite buffer tandem manufacturing systems with N-policy Computer and Industrial Engineering. 34(4), 837-848.

[13] Vidalis M.I., Papadopoulos, H.T. (2001) A recursive algorithm for generating the transition matrices of multistation multiserver exponential reliable queueing networks Computer and Operations Research 28(9), 853-883.

[14] Aweya, J., Ouellette, M., Montuno, D.Y. (2002) Multilevel active queue management with dynamic thresholds Computer and Communication 25(8), 756-771.

[15] Sharma, V., Virtamo J.T. (2002) A finite buffer queue with priorities Performance Evaluation 47(1), 1-22.

[16] Chydzinski, A., Winiarczyk, R. (2008) On the blocking probability in batch Markovian arrival queues Microprocessors and Microsystems - Embedded Hardware Design 32(1), 45-52.

[17] Goswami, V., Samanta, S.K., Vijaya Laxmi, P., Gupta, U.C. (2008) Analyzing a multiserver bulk service finite buffer queue Applied Mathematical Modelling 32(9),1797-1812.

[18] Diamantidis, C.A., Papadopoulos, T.C. (2009) Exact analysis of a two workstation one buffer flow line with parallel unreliable machines European Journal of Operational Research 197(2), 572-580.

[19] Wang, K.H., Ke, J.B., Ke, J.C. (2007) Profit analysis of the M/M/R machine repair problem with balking, reneging, and standby switching failures Computers and Operations Research 34(3), 835-847.

[20] Yang, W.S. and Choo, T.S. (2009) M/M/s queue with impatient customers and retrials Applied Mathematical Modelling 33(6), 2596-2606.

[21] Al-Seedy, R.O., EI-Sherbiny, A.A., El-Shehawy, S.A., Ammar, S.I. (2009) Transient solution of the M/M/c queue with balking and reneging Computers and Mathematics with Applications 57(8), 1280-1285.

[22] Perel, N., Yechiali, U. (2010) Queues with slow servers and impatient customers European Journal of Operational Research 201(1), 247-258.

[23] Shagon, A.W. (1979) A single server queue with arrival rate dependent on server breakdown Naval Research Logistics 26, 487-497.

[24] Gong, W.B., Yan, A., Gassandras, C.G. (1992) The M/G/1 queue with queue length dependent arrival rate Communications in Statistics, Stochastic Models 8(4), 733-741.

[25] Wang, P.P. (1996) Queueing models with delayed state-dependent service times European Journal of Operational Research. 88, 614-621.

[26] Jain, M. (1997) Optimal N-policy for single server Markovian queue with breakdown, repair and state dependent arrival rate International Journal of Management System 13(3), 245-260.

[27] Hwang, R.H., Kurok, J.F., Towsley, D. (2000) MDP routing for multi-rate loss networks Computer networks 34, 241-261.

[28] Jain, M., Sharma, G.C., Shekhar, C. (2005) Processor-shared service systems with queue-dependent processors Computers and Operations Research 32(3), 629-645.

[29] Lee, J., Kim, J. (2006) A workload-dependent M/G/1 queue under a two-stage service policy Operations Research Letters 34(5), 531-538.

[30] Parthasarathy, P.R., Sudesh, R. (2007) Time dependent analysis of a single server retrial queue with state-dependent rates Operations Research Letters 35(5), 601-611.

[31] Wall, A.D., Worthington, D.J. (2007) Time-dependent analysis of virtual waiting time behavior in discrete time queues European Journal of Operational Research 178(2), 482-499.

[32] Soares, A.D., Latouche, G. (2009) Fluid queues with level dependent evolution European Journal of Operational

Research 196(3), 1041-1048.

[33] Banik, A.D. (2011) Analyzing state-dependent arrival in G1 / BMSP / 1/queues Mathematical and Computer Modelling 53 (5-6), 1229-1246.

[34] Lee, C. (2011) On moment stability properties for a class of state-dependent stochastic networks Journal of the Korean Statistical Society 40(3), 325-336.

[35] Banerjee, A., Gupta, U.C. (2012) Reducing congestion in bulk-service finite - buffer queueing system using batch – size - dependent service Performance Evaluation 69(1), 53-70.

[36] Louvel, M., Plantec, A., Babau, J.P. (2013) Resource management for multimedia applications, distributed in open and heterogeneous home networks Journal of System Architecture 59(3), 121-134.

[37] Zhou, W., Lian, Z., Xu, W., Huang, W. (2013) A two-stage queueing network with MAP inputs and buffer sharing Applied Mathematical Modelling 37(6), 3736-3747.

[38] Chakravarthy, S.R., Karatza, H.D. (2013) Two servers parallel system with pure space sharing and Markovian arrivals Computers and Operations Research 40(1), 510-519.