Research Article

Markovian Queueing System with Discouraged Arrivals and Self-Regulatory Servers

K. V. Abdul Rasheed and M. Manoharan

Department of Statistics, University of Calicut, Kerala 673 635, India

Correspondence should be addressed to K. V. Abdul Rasheed; rasheedkv@rediffmail.com

Received 28 September 2015; Revised 1 February 2016; Accepted 28 March 2016

Academic Editor: Ahmed Ghoniem

Copyright © 2016 K. V. Abdul Rasheed and M. Manoharan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider discouraged arrival of Markovian queueing systems whose service speed is regulated according to the number of customers in the system. We will reduce the congestion in two ways. First we attempt to reduce the congestion by discouraging the arrivals of customers from joining the queue. Secondly we reduce the congestion by introducing the concept of service switches. First we consider a model in which multiple servers have three service rates \(\mu_1, \mu_2, \) and \(\mu (\mu_1 \leq \mu_2 < \mu)\), say, slow, medium, and fast rates, respectively. If the number of customers in the system exceeds a particular point \(K_1\) or \(K_2\), the server switches to the medium or fast rate, respectively. For this adaptive queueing system the steady state probabilities are derived and some performance measures such as expected number in the system/queue and expected waiting time in the system/queue are obtained. Multiple server discouraged arrival model having one service switch and single server discouraged arrival model having one and two service switches are obtained as special cases. A Matlab program of the model is presented and numerical illustrations are given.

1. Introduction

Queueing theory plays an important role in modeling real life problems involving congestion in wide areas of applied sciences. A customer decides to join the queue only when a short wait is expected and if the wait has been sufficiently small he tends to remain in the queue; otherwise the customer leaves the system and then the customer is said to be impatient. When this impatience increases and customers leave before being served, some remedial actions must be taken to reduce the congestion in the system.

Balking and reneging are the forms of impatience. If a customer decides not to join the queue upon arrival by seeing a long queue, the customer is said to have balked. A customer may enter the queue but after a while loses patience and decides to leave and then the customer is said to be reneged.

In the study of queueing system the server is usually assumed to work at constant speed regardless of the amount of work existing. But in real life situation this assumption may not always be appropriate as the system size may affect the system performance. That is, the servers adapt to the system state by increasing the speed to clear the queue or decrease the speed when fatigued, which means the service rate depends on the system size; see, for example, Jonckheere and Borst [1]. Similarly we can see that the arrivals of customers into the system may also affect the level of congestion. For example, customers impatience will affect the arrivals of customers into the system. A queueing system where the arrival rate and/or service rate depends on the system size is called adaptive queueing system.

Sometimes the arrival rate of customers into the system depends on the system size instead of a constant rate. Discouraged arrival is one form of state dependence. Here the arrivals get discouraged from joining the queue when more and more people are present in the system. We can model this effect by taking the birth and death coefficients, respectively, as \(\lambda_n = \lambda/(n + 1)\), where \(n\) is the number of customers in the system and \(\lambda\) is a positive constant, and \(\mu_n = \mu\), where \(\mu\) is the constant service rate. Thus the arrival rate of the queueing system is decreased by this discouragement; consequently the congestion of the system decreases. This type of queueing systems is called discouraged arrival queueing system.
Applications of queueing with impatience can be seen in traffic modeling, business and industries, computer communication, health sectors, medical sciences, and so forth. Customers with impatience and discouragement have their own impact on the system performance of standard queueing systems. It is important to note that customers impatience has a very negative impact on the queueing system under investigation.

We can see a large number of articles which discussing the congestion control of queueing systems. For example, see, Haight [2], Ancker Jr. and Gafarian [3, 4], Boots and Tijms [5], Liu and Kulkarni [6], Kc and Terwiesch [7], Wang et al. [8], Kapodistria [9], and Kumar and Sharma [10].

In this paper we attempt to reduce the congestion in two ways. First we attempt to reduce the congestion by discouraging the arrivals of customers from joining the queue. In the second way we further reduce the congestion by introducing the concept of service switches which is discussed in Abdul Rasheed and Manoharan [11].

Discouraged arrival system was studied by many researchers. Raynolds [12] presented multiserver queueing model with discouragement and obtained equilibrium distribution of queue length and derived other performance measures from it. A finite capacity M/G/1 queueing model where the arrival and the service rates were arbitrary functions of the current number of customers in the system was studied by Courtois and Georges [13]. Natvig [14] studied the single server birth-death queueing process with state dependent parameters \( \lambda_n = \lambda/(n+1) \), \( n \geq 0 \), and \( \mu_n = \mu, n \geq 1 \). Van Doorn [15] obtained exact expressions for transient state probabilities of the birth-death process with parameters \( \lambda_n = \lambda/(n+1) \), \( n \geq 0 \), and \( \mu_n = \mu, n \geq 1 \). Parthasarathy and Selvaraju [16] obtained the transient solution to a state dependent birth-death queueing model in which potential customers are discouraged by queue length.

Narayanan [17] studied different linear and nonlinear state dependent Markovian queueing models, in which arrival rates and/or service rates are nonlinear and their modified forms obtain the transient/steady state probability distribution of queue length. Narayanan and Manoharan [18] considered nonlinear state dependent queueing models, in which arrival rates and/or service rates are nonlinear. Narayanan and Manoharan [19] derived the performance measures of state dependent queueing models. Ammar et al. [20] studied single server finite capacity Markovian queue with discouraged arrivals and reneging using matrix method.

Abdul Rasheed and Manoharan [11] studied a Markovian queueing system in which the arrival rate is constant and the service rate depends on the number of customers in the system. The server speed is regulated according to the system size by introducing the service switches to the model. The authors analyzed the system by calculating the performance measures such as expected number of customers in the system/queue and expected waiting time of customers in the system/queue. Some generalizations of the above models are also presented therein.

A generalization of M/M/C queueing system with service switches is considered in this paper in which the arrival rate \( \lambda_n \) and the service rate \( \mu_n \) are both functions of \( n \), the number of customers present in the system. In real life situations \( \lambda_n \) and \( \mu_n \) change whenever \( n \) changes, so that both arrival and departure have a bearing on the system state.

In many practical queueing systems, when there is a long queue, it is quite likely that a server will tend to work faster than when the queue is small. That is, the service rate \( \mu_n \) depends on \( n \), the number of customers present in the system. Similarly situations may occur where customers refuse to join the queue because of long waiting by seeing a large number of customers in the queue. That is, the arrival rate \( \lambda_n \) depends on \( n \), the number of customers present in the system. These kinds of adaptive queueing systems where the arrival rate and the service rate depend on the number of customers present in the system are discussed in this paper.

### 2. Multiserver Multirate Discouraged Arrival Queueing System

A queue is an indication of congestion which we can be seen in a system or a network consisting of many systems. Congestion arises in many areas and our interest is to control the congestion in whatever situation it arises. Abdul Rasheed and Manoharan [11] used the concept of service switches as a tool to control congestion and use multiple servers and multiple service switches if congestion is very high. They discussed congestion control aspects when the arrival rate is constant and the service rate depends on the number in the system. In this paper we discuss the congestion control using service switches when both the arrival rate and service rate are the functions of the number of customers in the system. We consider a generalized model with \( C \) servers and two service switches at the point \( K_1 \) and \( K_2 \) \( (K_2 > K_1) \) and hence the system works in three speeds, say, slow, medium, and fast. Here work is performed at the slow rates until there are \( K_1 \) customers in the system, at which point there is a switch to the medium rate and work is performed at the medium rate until there are \( K_2 \) customers in the system at which point there is a switch to the fast rate. Here the arrival rate \( \lambda_n \) is given as

\[
\lambda_n = \frac{\lambda}{n+1}, \quad n \geq 0,
\]

which means the customers will be discouraged from joining the queue and the service rate \( \mu_n \) is given by

\[
\mu_n = \begin{cases} 
\eta \mu_1, & 1 \leq n < C, \\
C \mu_1, & C \leq n < K_1, \\
C \mu_2, & K_1 \leq n < K_2, \\
C \mu_3, & n \geq K_2,
\end{cases}
\]

where \( \mu_1 \leq \mu_2 < \mu \).
The steady state probabilities are given by

\[ P_n = \begin{cases} \frac{r^n}{(n!)^2} P_0, & 0 \leq n < C, \\ \frac{r^n_{1,2}}{n!C!_n^n} P_0, & C \leq n < K_1, \\ \frac{r^n_{1,2}}{n!C!_n^n} - \frac{r^n_{1} r^n_{K_2-K_1+1}}{n!C!_n^n} P_0, & K_1 \leq n < K_2, \\ \frac{r^n_{1} r^n_{K_2-K_1+1}}{n!C!_n^n} P_0, & n \geq K_2, \end{cases} \]

where \( r_1 = \lambda/\mu_1, r_2 = \lambda/\mu_2 \), and \( r = \lambda/\mu \).

The idle probability \( P_0 \) can be obtained as

\[ P_0 = \left[ \sum_{n=0}^{C-1} \frac{r^n_{1}}{(n!)^2} + \frac{K_1-1}{C!} \sum_{n=C}^{K_2-1} \frac{r^n_{1,K_2-K_1+1}}{n!C!_n^n} + \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C!_n^n} \right]^{-1} + \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C!_n^n} \]

After a careful manipulation of the infinite series on the right hand side of the above expression and further simplification, we get

\[ P_0 = \left( \frac{r^n_{1,K_2-K_1+1}}{C!C^n} \right) \left[ \sum_{n=0}^{C-1} \frac{r^n_{1}}{(n!)^2} + \frac{K_1-1}{C!} \sum_{n=C}^{K_2-1} \frac{r^n_{1,K_2-K_1+1}}{n!C!_n^n} + \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C!_n^n} \right]^{-1} + \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C!_n^n} \]

The expected queue size \( \langle L \rangle \) is given as

\[ L_q = \sum_{n=C+1}^{K_1-1} (n-C) P_n + \sum_{n=K_1}^{K_2-1} (n-C) P_n + \sum_{n=K_2}^{\infty} (n-C) P_n = L_{q1} + L_{q2} + L_{q3}, \]

where

\[ L_{q1} = \sum_{n=C+1}^{K_1-1} \frac{r^n_{1}}{(n!)^2} P_0 - \sum_{n=C+1}^{\infty} \frac{r^n_{1}}{n!C^n} P_0, \]

\[ L_{q2} = \sum_{n=K_1}^{K_2-1} \frac{r^n_{1,K_2-K_1+1}}{n!C^n} P_0 - \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C^n} P_0, \]

\[ L_{q3} = \sum_{n=K_2}^{\infty} \frac{n P_n - C \sum_{n=K_2}^{\infty} P_n}{n!C^n} = L_{q31} - L_{q32}, \]

\[ L_{q31} = \frac{r^n_{1,K_2-K_1+1}}{C!C^n} P_0 \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C^n} P_0, \]

\[ L_{q32} = C \left[ \sum_{n=0}^{C-1} P_n - \sum_{n=1}^{C} P_n - P_0 - \sum_{n=C+1}^{K_1-1} P_n \right] \]

Therefore \( L_q \) becomes

\[ L_q = \left[ \frac{r^n_{1,K_2-K_1+1}}{C!C^n} P_0 \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C^n} P_0 \right] - C \left[ \sum_{n=0}^{C-1} P_n - \sum_{n=1}^{C} P_n - P_0 - \sum_{n=C+1}^{K_1-1} P_n \right] - \sum_{n=K_1}^{\infty} \frac{r^n_{1,K_2-K_1+1}}{n!C^n} P_0. \]
After some steps we get expected number of customers in the queue by using \( L_{q_1}, L_{q_2}, \) and \( L_{q_3} \) as

\[
L_q = \left[ \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{C!C^{n-C}} P_0 \right] + \sum_{n=1}^{C} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + \sum_{n=C+1}^{K_1-1} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + \sum_{n=K_1}^{K_2-1} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + \sum_{n=K_2}^{\infty} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + C \right] \sum_{n=0}^{\infty} \left( (n-C) P_n \right),
\]

The expected system size \( L \) is given as

\[
L = \sum_{n=0}^{C} n P_n + \sum_{n=1}^{K_1} n P_n + \sum_{n=K_1+1}^{K_2} n P_n + \sum_{n=K_2+1}^{\infty} n P_n
\]

\[
= \left[ \sum_{n=0}^{C} n P_n + \sum_{n=K_1+1}^{K_2} n P_n + \sum_{n=K_2+1}^{\infty} n P_n \right] + \sum_{n=K_1+1}^{\infty} (n-C) P_n + C \left[ \sum_{n=K_1+1}^{K_2} P_n + \sum_{n=K_2+1}^{\infty} P_n \right].
\]

Hence the expected number of customers in the system \( L \) is obtained as

\[
L = \left[ \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{C!C^{n-C}} P_0 \right] + \sum_{n=1}^{C} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + \sum_{n=C+1}^{K_1-1} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + \sum_{n=K_1}^{K_2-1} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + \sum_{n=K_2}^{\infty} \left( \frac{r_{K_1-1} r_{K_2-K_1} r_{K_2} r_{/C}}{(n!)^2} \right) r_{K_1-1} P_0 + C \right] \sum_{n=0}^{\infty} \left( (n-C) P_n \right) + C \left[ \sum_{n=K_1+1}^{K_2} P_n + \sum_{n=K_2+1}^{\infty} P_n \right].
\]

The effective arrival rate \( \lambda^* \) can be obtained by the following summation schemes:

\[
\lambda^* = \sum_{n=0}^{C-1} \lambda_n P_n + \sum_{n=0}^{C-1} \lambda_n P_n + \sum_{n=K_1}^{K_2-1} \lambda_n P_n + \sum_{n=K_2}^{\infty} \lambda_n P_n
\]

\[
= \left[ \sum_{n=0}^{C-1} \lambda_n P_n + \sum_{n=K_1}^{K_2-1} \lambda_n P_n + \sum_{n=K_2}^{\infty} \lambda_n P_n \right] + \sum_{n=0}^{C-1} \lambda_n P_n + \sum_{n=K_1}^{K_2-1} \lambda_n P_n + \sum_{n=K_2}^{\infty} \lambda_n P_n.
\]
The expected waiting times in the system are

$$W = \left( \frac{r_2^{K_2-K_1+2} - r_2 r_1^{K_2-K_1+1}}{C! (C^n-C)} \right)$$

$$+ \sum_{n=1}^{C-1} \left( \frac{(r_2^{n-K_1+1} - r_2) K_2 r_1^{n-K_1+1}}{(n+1)! C! (C^n-C)} \right)$$

$$+ \sum_{n=K_1}^{K_2-1} \left( \frac{(r_2^{n-K_1+1} - r_2) K_2 r_1^{K_2-K_1+1}}{(n+1)! C! (C^n-C)} \right)$$

$$+ \sum_{n=K_1}^{C} \left( \frac{(r_2^{n-K_1+1} - r_2) K_2 r_1^{n-K_1+1}}{(n+1)! C! (C^n-C)} \right)$$

$$- C (1 - P_0) \left( \frac{r_2^{K_2-K_1+2} - r_2 r_1^{K_2-K_1+1}}{C! (C^n-C)} \right)$$

(14)

Similarly the expected waiting times in the queue are

$$W_q = \left( \frac{r_2^{K_2-K_1+2} - r_2 r_1^{K_2-K_1+1}}{C! (C^n-C)} \right)$$

$$+ \sum_{n=1}^{C-1} \left( \frac{(r_2^{n-K_1+1} - r_2) K_2 r_1^{n-K_1+1}}{(n+1)! C! (C^n-C)} \right)$$

$$+ \sum_{n=K_1}^{K_2-1} \left( \frac{(r_2^{n-K_1+1} - r_2) K_2 r_1^{K_2-K_1+1}}{(n+1)! C! (C^n-C)} \right)$$

$$+ \sum_{n=K_1}^{C} \left( \frac{(r_2^{n-K_1+1} - r_2) K_2 r_1^{n-K_1+1}}{(n+1)! C! (C^n-C)} \right)$$

$$- C (1 - P_0) \left( \frac{r_2^{K_2-K_1+2} - r_2 r_1^{K_2-K_1+1}}{C! (C^n-C)} \right)$$

(15)

(16)

For illustrating the analytical feasibility of the methods proposed we consider the following hypothetical example situation.

Example 1. A young hard worker started a beauty parlour. Customers are taken on a first come first serve basis. Inside the beauty parlour, sitting facility is available for waiting customers and in front of the beauty parlour there is a vast parking area, so no limitation on the number of customers who can wait for service. But the number of arrivals depends on the number of customers already present in the beauty parlour. If arriving customers see a large number in the system he may not join the queue. Since the congestion is very high the young man appointed one more worker and he decides to run the beauty parlour at three speeds, say, slow, medium, and fast. At the slow speed, it takes 40 minutes, on the average; at the medium speed, it takes 30 minutes; and at the fast speed, it takes 20 minutes to cut the hair with service switches at 5 and 7. That is, up to 4 customers in the system the beauty parlour runs at the slow speed. If the number of customers is more than 4 but less than 7, the beauty parlour runs at the medium speed. If the number of customers is more than 6, the beauty parlour runs at the fast speed. That is, $K_1 = 5$ and $K_2 = 7$. The interarrival time of customers is 35 minutes.

Now we can calculate the measures of effectiveness.

If $K_1 = 5$, $K_2 = 7$, $\lambda = 1/35$, $\mu_1 = 1/40$, $\mu_2 = 1/30$, $\mu = 1/20$, and $C = 2$. Then we get $\lambda^* = 0.0190$, $P_0 = 0.3935$, $L = 0.795$, $L_q = 0.032$, $W = 41.68$ minutes, and $W_q = 1.7$ minutes.

Figure 1 gives the graph of steady state probability of number of customers in the system.

If $K_1 = 3$ and $K_2 = 5$ and using the same parameters we get $\lambda^* = 0.0190$, $P_0 = 0.3966$, $L = 0.776$, $L_q = 0.022$, $W = 40.46$ minutes, and $W_q = 1.18$ minutes.

If $K_1 = 5$, $K_2 = 7$, $\lambda = 1/35$, $\mu_1 = 1/40$, $\mu_2 = 1/30$, $\mu = 1/20$, and $C = 1$. Then we get $\lambda^* = 0.0170$, $P_0 = 0.3195$, $L = 1.134$, $L_q = 0.4544$, $W = 66.56$ minutes, and $W_q = 26.65$ minutes.

Figure 2 gives the graph of steady state probability of number of customers in the system.

From this example we can observe that waiting time of the customers decreases if the values of the switch point decrease and also by increasing the number of servers.

Some special cases of the generalized C server model and two service switches are discussed in the following section.
The steady state probabilities are given by

\[
P_n = \begin{cases} 
    \frac{r_1^n}{(n!)^2} P_0, & 0 \leq n < C, \\
    \frac{r_1^n}{n! C^{n-C}} P_0, & C \leq n < K, \\
    \frac{r_1^{K-1} (n-K+1^n-K+1)}{n! C^{n-C}} P_0, & n \geq K.
\end{cases}
\]  

The idle probability \(P_0\) is obtained as

\[
P_0 = \left[ \sum_{n=0}^{C-1} \left( \frac{r_1^n}{(n!)^2} - \frac{r_1^{K-1} (n-K+1^n-K+1)}{n! C^{n-C}} \right) \right] P_0.
\]  

The expected number of customers in the system is

\[
L = \left[ \frac{r_1^{K-1} r_{-2K} e^{r/C}}{C! C^{1-C}} P_0 
+ \sum_{n=1}^{C} \left( \frac{r_1^n}{(n-1)!} - \frac{r_1^n (n-K+1^n-K+1)}{(n-1)! C^{n-C}} \right) r_1^{K-1} P_0 
+ \sum_{n=C+1}^{K-1} \left( \frac{r_1^n - r_1^{n-K+1}}{(n-1)!} \right) r_1^{K-1} P_0 - C (1 - P_0) \right].
\]

The effective arrival rate can be calculated as follows:

\[
\lambda^* = \sum_{n=0}^{C-1} \lambda_n P_n + \sum_{n=C}^{K-1} \lambda_n P_n + \sum_{n=K}^{\infty} \lambda_n P_n 
= \left[ \sum_{n=0}^{C-1} \left( \frac{r_1^n}{(n+1)!} - \frac{r_1^n}{C! (n+1)! C^{n-C}} \right) \right] r_1^{K-1} P_0
+ \sum_{n=C+1}^{K-1} \left( \frac{r_1^n - r_1^{n-K+1}}{(n+1)! C^{n-C}} \right) P_0.
\]
The expected waiting time in the system $W$ is given by

$$W = \left( \frac{r_{K-1}^{C} e^{r/C}}{C!C^{C-1}} \right) + 
\sum_{n=1}^{C} \left( \frac{r_{K-1}^{n} e^{r/C}}{(n-1)!n! - (n-1)!C^{n-C}} \right) + 
\sum_{n=C+1}^{K-1} \left( \frac{r_{K-1}^{n} - r_{K-1}^{n+1}}{(n-1)!C^{n-C}} \right) \cdot \lambda P_{0}^{(n)} \right)^{-1},$$

and expected waiting time in the queue $W_q$ is

$$W_q = \left( \frac{r_{K-1}^{C} e^{r/C}}{C!C^{C-1}} \right) + 
\sum_{n=1}^{C} \left( \frac{C r_{K-1}^{n} e^{r/C}}{(n-1)!n!} - \frac{r_{K-1}^{n} e^{r/C}}{(n-1)!C^{n-C}} \right) \cdot \lambda P_{0}^{(n)} \right)^{-1},$$

3.2. Multiple Server Adaptive Queueing System. If $\mu_1 = \mu_2 = \mu$ and hence $r_1 = r_2 = r$ (which means no switch), the model with $C$ server and two service switches reduces to the $C$ server model with no service switch. The following results can be obtained from the $C$ server model and two service switches. The service rate $\mu_n$ is given by

$$\mu_n = \begin{cases} n\mu, & 1 \leq n < C, \\ \mu, & n \geq C. \end{cases}$$

The steady state probability of $n$ customers in the system is

$$P_n = \begin{cases} \frac{r^n}{(nl)^2} P_0, & 0 \leq n < C, \\ \frac{r^n}{n!C^{n-C}} P_0, & n \geq C. \end{cases}$$

The idle probability can be obtained as

$$P_0 = \left[ \sum_{n=0}^{\infty} \frac{r^n}{(nl)^2} + \sum_{n=C}^{\infty} \frac{r^n}{n!C^{n-C}} \right]^{-1}.$$
3.3. Model with Single Server and Two Service Switches. If \( C = 1 \), the model with one server and two service switches reduces to the single server model with two service switches at the points \( K_1 \) and \( K_2 \) (\( K_1 < K_2 \)). The following results can be obtained from the multiserver model with two switches. The service rate \( \mu_n \) is given by

\[
\mu_n = \begin{cases} 
\mu_1, & 1 \leq n < K_1, \\
\mu_2, & K_1 \leq n < K_2, \\
\mu, & n \geq K_2.
\end{cases}
\]

The expected number of customers in the system \( L \) is given by

\[
L = \sum_{n=1}^{K_1-1} \left( \frac{r_1^{K_1-1} n^{K_1-1} - r_2^{K_1-1} n^{K_1-1}}{(n-1)!} \right) + \sum_{n=K_1}^{K_2-1} \left( \frac{r_1^{K_1-1} K_2-K_1 r^{n-K_1} n^{K_1-1}}{(n-1)!} \right) + r_1^{K_1-1} K_2-K_1 r^{K_2-K_1} e^{-r} P_0.
\]

The steady state probability of \( n \) customers in the system is

\[
P_n = \begin{cases} 
\frac{r_1^n}{n!} P_0, & 0 \leq n < K_1, \\
\frac{r_1^{K_1-1} n^{K_1-1} - r_2^{K_1-1} n^{K_1-1}}{n!} P_0, & K_1 \leq n < K_2, \\
\frac{r_1^{K_1-1} K_2-K_1 r^{n-K_1} n^{K_1-1}}{n!} P_0, & n \geq K_2.
\end{cases}
\]

We have the idle probability

\[
P_0 = \sum_{n=0}^{K_1-1} \left( \frac{r_1^{K_1-1} n^{K_1-1} - r_2^{K_1-1} n^{K_1-1}}{n!} \right) + \sum_{n=K_1}^{K_2-1} \left( \frac{r_1^{K_1-1} K_2-K_1 r^{n-K_1} n^{K_1-1}}{(n-1)!} \right) + r_1^{K_1-1} K_2-K_1 r^{K_2-K_1} e^{-r}.
\]

The expected waiting time in the system is

\[
W = \left[ \left( \frac{r_1^{K_1-1} K_2-K_1 r^{K_2-K_1}}{e^{-r}} \right) \left( \frac{r_1^{K_1-1} n^{K_1-1} - r_2^{K_1-1} n^{K_1-1}}{(n-1)!} \right) + \sum_{n=1}^{K_1-1} \left( \frac{r_1^{K_1-1} K_2-K_1 r^{n-K_1} n^{K_1-1}}{(n-1)!} \right) \right]
\]

The expected number of customers in the queue \( L_q \) is given by

\[
L_q = \left[ \left( \frac{r_1^{K_1-1} K_2-K_1 r^{K_2-K_1}}{e^{-r}} \right) \left( \frac{r_1^{K_1-1} n^{K_1-1} - r_2^{K_1-1} n^{K_1-1}}{(n-1)!} \right) + \sum_{n=0}^{K_1-1} \left( \frac{r_1^{K_1-1} K_2-K_1 r^{n-K_1} n^{K_1-1}}{(n-1)!} \right) \right] P_0.
\]

The effective arrival rate now can be obtained as

\[
\lambda^* = \left[ \left( \frac{r_1^{K_1-1} K_2-K_1 r^{K_2-K_1}}{e^{-r}} \right) \left( \frac{r_1^{K_1-1} n^{K_1-1} - r_2^{K_1-1} n^{K_1-1}}{(n-1)!} \right) + \sum_{n=0}^{K_1-1} \left( \frac{r_1^{K_1-1} K_2-K_1 r^{n-K_1} n^{K_1-1}}{(n-1)!} \right) \right] \lambda P_0.
\]

We can see that \( L_q = L - (1 - P_0) \) from the above equation.
The expected waiting time in the queue is

\[
W_q = \left[ \frac{K_1}{r_1} r_2 - K_1, r_2 - K_2, r - K_2, e^r P_0 \right. \\
+ \sum_{n=1}^{K_1-1} \left( \frac{(r^n-K_1-1) - r^n-K_2, r^n-K_2}{(n-1)!} \right) r_1^{K_1-1} P_0 \\
+ \sum_{n=K_1}^{K_2-1} \left( \frac{(r^n-K_2-1) - r^n-K_2, r^n-K_2}{(n-1)!} \right) r_1^{K_2-K_1} P_0 - (1 - P_0) \\
- P_0 \right) \left( r_1^{K_1-1} r_2 - K_1, r_2 - K_2, (e^r - 1) \right) \\
+ \sum_{n=0}^{K_1} \left( \frac{r_1^{n+1} - r_2^{n-K_1+1}}{(n+1)!} \right) r_1^{K_1-1} \\
+ \sum_{n=K_1}^{K_2-1} \left( \frac{r_2^{n-K_2-1} - r^n-K_2}{(n+1)!} \right) r_1^{K_2-K_1} P_0 - (1 - P_0) \\
\cdot \lambda P_0^{-1} \right] .
\]

Also we can establish the relationship \( W = W_q + (1 - P_0) / \lambda^* \), so the expected service time is \((1 - P_0) / \lambda^*\).

3.4. Model with Single Server and One Service Switch. If \( C = 1 \), \( \mu_1 = \mu \), and hence \( r_1 = r \), the model with \( C \) server and two service switches reduces to the single server model with one service switch at the point \( K \). We obtained the following results from the \( C \) server two switch models. The service rate \( \mu_n \) is given as

\[
\mu_n = \begin{cases} 
\mu_1, & 1 \leq n < K, \\
\mu, & n \geq K.
\end{cases}
\]

The steady state probabilities are given by

\[
P_n = \begin{cases} 
\frac{\rho^n}{n!} P_0, & 0 \leq n < K, \\
\frac{1}{r^2 - K_2} P_0, & n \geq K.
\end{cases}
\]

The idle probability can be obtained from the result \( \sum_{n=0}^{\infty} P_n = 1 \), as

\[
P_0 = \left[ \frac{1}{r^2 - K_2} - \sum_{n=0}^{K-1} \frac{(r^n-K+1) - r^n-K_1}{(n+1)!} \right]^{-1}.
\]

The expected number of customers in the system \((L)\) is

\[
L = \left[ \frac{1}{r^2 - K_2} \right] \left( \frac{r^n-K+1 - r^n-K_1}{(n+1)!} \right) P_0.
\]

The expected number of customers in the queue \((L_q)\) is

\[
L_q = \left[ \frac{e^{r^2-K_2} - (r^2-K_2 - r^2-K_1)}{\lambda P_0} - \sum_{n=2}^{K-1} \frac{(r^n-K+1 - r^n-K_1)}{(n+1)!} \right]^{-1}.
\]

The effective arrival rate \( \lambda^* \) can be obtained as

\[
\lambda^* = \frac{r_1^{K-1} - r^2-K_1}{\lambda P_0}.
\]

The expected waiting time in the system can be obtained as

\[
W = \left[ \frac{e^{r^2-K_2} - (r^2-K_2 - r^2-K_1)}{\lambda P_0} + \sum_{n=1}^{K-1} \frac{(r^n-K+1 - r^n-K_1)}{(n+1)!} \right]^{-1}.
\]

Similarly the expected waiting time of customers in the queue can be obtained as

\[
W_q = \left[ \frac{e^{r^2-K_2} - (r^2-K_2 - r^2-K_1)}{\lambda P_0} \right]^{-1}.
\]
3.5. **Single Server Adaptive Queueing System.** If $C = 1, \mu_1 = \mu_2 = \mu$, and hence $r_1 = r_2 = r$, the model with $C$ server and two service switches reduce to the single server model with no service switch and hence the following results can be obtained from the $C$ server model and two service switches. The service rate $\mu_n$ is given by $\mu_n = \mu, n \geq 1$.

The steady state probability of $n$ customers in the system is

$$P_n = \frac{r^n}{n!} P_0, \quad (49)$$

where

$$P_0 = e^{-r}. \quad (50)$$

Hence steady state probability of $n$ customers in the system becomes

$$P_n = \frac{e^{-r} r^n}{n!}, \quad (51)$$

which is a Poisson distribution with parameter $r = \lambda/\mu$.

The expected number of customers in the system ($L$) is

$$L = \sum_{n=1}^{\infty} n P_n = r, \quad (52)$$

since $P_n$ follows Poisson distribution.

Similarly the expected number of customers in the queue ($L_q$) is

$$L_q = \sum_{n=1}^{\infty} (n-1) P_n = e^{-r} + r - 1. \quad (53)$$

Expected waiting times of customers in the system ($W$) are given as

$$W = \frac{r}{\mu (1 - e^{-r})}. \quad (54)$$

Expected waiting times of customers in the queue ($W_q$) are given as

$$W_q = \frac{e^{-r} + r - 1}{\mu (1 - e^{-r})}. \quad (55)$$

**Sensitivity Analysis.** Now we investigate the nature of the expected number of customers in the system and expected waiting time of customers in the system on the basis of the various values of the switch point $K$ by the following sensitivity analysis.

From Table 1 we can observe that as the value of the switch point $K$ increases, the results of the single server discouraged arrival model with one service switch tending to the results of the single server discouraged arrival model with no service switch. That is, if $K \geq 9$, $P_0 = 0.3189$ which is the $P_0$ of the single server discouraged arrival model with no service switch. Similarly if $K \geq 11$, $L = 1.142857$ which is the $L$ of the single server discouraged arrival model with no service switch and if $K \geq 12$, $W = 67.118963$ which is the $W$ of the single server discouraged arrival model with no service switch. Hence we conclude that there is no effect by the switch point if its value increases.

**4. Summary and Concluding Remarks**

In this paper we study a discouraged arrival Markovian queueing systems. To this system we introduce self-regulatory servers and analyzed the model by deriving steady state characteristics. A generalized multiple server discouraged arrival model with two service switches are discussed. By introducing service switches we could speed up the service based on the switch point if the number of arrivals increases. That is, the speed of the server can be slow, medium, and fast. Thus the congestion of customers can be reduced by the two features mentioned above. The steady state probabilities and all the performance measures such as expected number of customers in the queue/system and expected waiting time of customers in the queue/system are derived.

From this general model we derived $C$ server discouraged arrival model with one service switch, single server discouraged arrival model with two service switches, single server discouraged arrival model with one service switch, multiple server adaptive queueing system, and single server adaptive queueing system as a special case.

Numerical illustration of the model is given. A Matlab program was developed to help the numerical illustration. A sensitivity analysis is conducted to obtain the optimum value of the switch point.

From the numerical illustration, we can observe that as the value of the switch points decreases the waiting time of customers also decreases.

For implementing the above models in real life situations we need to include the cost of the various service rate of the system. An optimization problem can be a future work of the model by finding the optimal choice of the service switches and the number of servers by considering the costs involved in the queueing system.

**Competing Interests**

The authors declare that they have no competing interests.

**References**


