A Study on State Dependent Accessible and Second Optional Service Queue with Impatient and Feedback Customers

P. Vijaya Laxmi a, b and E. Girija Bhavani b

a, b Department of Applied Mathematics, Andhra University, Visakhapatnam 530 003, India.

Abstract. In this paper, we study a single server Markovian queue with state dependent accessible services, reneging and feedback of customers. The server can accommodate at the most $d$ customers in the service station, after which the new arrivals have to wait in the primary queue of infinite waiting space. All arrivals demand First Essential Service (FES), after completing FES, customers decide to join the second optional Service (SOS) with a probability. The services occur singly according to first come first served service discipline and service times in FES and SOS are exponentially distributed. A customer waiting in the primary queue may get impatient and renge from the queue. However, after completing SOS, if the customer is not satisfied with the service quality, he may join the queue again (feedback). The system is analyzed by a quasi birth-death process and the steady state probabilities of the model are obtained using matrix geometric method. Some performance measures and numerical illustrations are also provided. An optimization of the cost function is performed to find the optimal service rate that minimizes the total cost.

Keywords: Queue; First essential service; Second optional service; Reneging; Feedback.

AMS Subject Classification: 60K25, 65K30.

1. Introduction

Recently there have been several contributions considering queueing system in which the server may provide a second optional service. Such queueing situations...
occur in day-to-day life. For example, (a) at a barber shop every customer needs hair cut but only few of them may opt for shave or massage (b) students admitted in a particular department of a university want to complete their under graduate program of study but only some of them may join the post graduate program soon after completing the under graduate program. Such a model was first introduced in [11] by studying an $M/G/1$ queue with SOS, using the supplementary variable technique wherein general service time distribution for FES and exponential service time distribution for SOS are considered. [14] derived the transient and steady state solution for an $M/G/1$ queue with SOS using the same technique. An $M/G/1$ queue with SOS and server break downs has been analyzed in [18] and in [1] a study of $M/G/1$ queue with SOS and general service time distribution is found. An infinite capacity multi-server queue with SOS channel is analyzed in [7].

Reneging of customers due to impatience was first investigated in [2]. They studied an $M/M/1/N$ queueing system with balking and reneging. Manoharan and Sasi [12] investigated an $M/G/1$ reneging queueing system with SOS and second optional vacation. In [3] a queueing system with returning of customers and waiting line is presented. [16] studied a multi-server infinite capacity Markovian feedback queue with balking, reneging and retention of reneged customers. The queueing system with two phase service has been studied in [9]. Doshi [6] has extended [9] into case of general service times. [4] studied time dependent solution of $M[X]/G/1$ queueing model with SOS, Bernoulli $K$ — optional vacation and balking whereas in [13] analysis of an $M/G/1$ feedback queueing system with SOS and second optional vacation has been carried out. Furthermore, [17] investigated an $M/G/1$ feedback queue with two types of services having general distribution.

In this paper, we consider an $M/M/1$ queue with second optional service, reneging and feedback of customers and two types of state dependent accessible services. Arrivals occur according to a Poisson process. There is a service queue in the service station with maximum limit of $d$ customers, beyond which, the new arrivals have to wait in the primary queue. The services/ departures occur one by one and service times in FES and SOS are assumed to be exponentially distributed. We have used matrix geometric method to obtain the steady state system length distributions using rate matrices and Neuts and Rao’s [15] truncation method. Some performance measures have been discussed and a cost model is constructed to determine the optimum service rate using direct search method that minimizes the cost function. The rest of the paper is organized as follows. Sections 2 and 3 present the description and mathematical formulation of the model, respectively. Steady state solutions are presented in Section 4 followed by various performance measures in Section 5. Section 6 presents numerical results and Section 7 concludes the paper.

2. Model description

Let us consider an $M/M/1$ queueing model wherein arrivals occur according to a Poisson distribution with parameter $\lambda$ and join in the service queue whose maximum capacity is $d$. After that limit is reached, the new arrivals queue up in the primary queue. The services/departures occur singly. The service discipline, both for the service queue and primary queue is first-come first-served. The FES is needed by all arriving customers, after which a customer may proceed to SOS with probability $q_1$ or may leave the system with the complementary probability $1-q_1$. After completing SOS, a customer may again join the primary queue (feedback) with probability $q_2$ or depart from the system with probability $1-q_2$. The service times of both FES and SOS are exponentially distributed with rates $\mu_n$ and $\eta$. 
respectively, where
\[ \mu_n = n\mu \quad \text{for} \quad 1 \leq n \leq d - 1, \quad \text{and} \quad \mu_n = d\mu, \quad \text{for} \quad n \geq d. \]

A customer waiting in the primary queue may get impatient and renege from the queue if his waiting time is beyond a tolerable limit. Reneging times are also exponentially distributed with parameter \( \xi \). The average reneging rate is zero for \( n < d \).

We assume that the arrival times and the service times are all independent and identically distributed random variables. Further, the new arrivals and feedback customers are treated identical.

### 3. Mathematical formulation of the model

Let at time \( t \), \( L(t) \) and \( S(t) \) be the number of customers in the primary queue and service queue respectively. Further let \( \zeta(t) \) be the state of the server at time \( t \) which is defined as

\[ \zeta(t) = \begin{cases} 0, & \text{if the server is busy in FES,} \\ 1, & \text{if the server is rendering SOS.} \end{cases} \]

We see that \( \omega = \{L(t), S(t), \zeta(t)\} \) defines a Markov process with state space
\[ E = \{(i, j, \zeta(t)): i \geq 0, \ 0 \leq j \leq d, \ \zeta(t) = 0 \text{ or } 1\}. \]

At steady state, let \( P_{i,n}(Q_{i,n}) \), \( i \geq 0, \ 0 \leq n \leq d \), be the probability that \( i \) and \( n \) number of customers are present in the primary queue and service queue, respectively and the server is busy in FES (SOS). Using Markov theory, the set of balance equations at steady state are given by

\[
\begin{align*}
0 &= -\lambda P_{0,0} + q_2 \eta Q_{0,0} + (1 - q_1)\mu P_{0,1}, \\
0 &= -(\lambda + n\mu)P_{0,n} + \lambda P_{0,n-1} + q_2 \eta Q_{0,n} + (1 - q_1)(n + 1)\mu P_{0,n+1}, \quad 1 \leq n \leq d - 1, \\
0 &= -(\lambda + d\mu)P_{0,d} + \lambda P_{0,d-1} + (d\mu(1 - q_1) + \xi) P_{1,d} + q_2 \eta Q_{0,d}, \\
0 &= -(\lambda + d\mu + \xi)P_{n,d} + \lambda P_{n-1,d} + (d\mu(1 - q_1) + \xi)P_{n+1,d} + q_2 \eta Q_{1,d}, \quad n \geq 1, \\
0 &= -(\lambda + q_2 \eta)Q_{0,0} + q_1 \mu P_{0,1}, \\
0 &= -(\lambda + q_2 \eta)Q_{0,n} + (n + 1)q_1 \mu P_{0,n+1} + \lambda Q_{0,n-1}, \quad 1 \leq n \leq d - 1, \\
0 &= -(\lambda + q_2 \eta)Q_{0,d} + dq_1 \mu P_{1,d} + \lambda Q_{0,d-1} + \xi Q_{1,d}, \\
0 &= -(\lambda + q_2 \eta + \xi)Q_{n,d} + dq_1 \mu P_{n+1,d} + \lambda Q_{n-1,d} + \xi Q_{n+1,d}, \quad n \geq 1. 
\end{align*}
\]

### 4. Matrix-geometric solution

The matrix geometric method allows us to deal with the models with rapid growth of the state space introduced by the need to explicitly construct the generator matrix of the underlying Markov process. The method can only be applied if the system can be decomposed into two parts: the initial portion and the repetitive portion. For example, in the present model, customers queued up has a structure that is possibly unbounded. Typically, there exists an integer \( d + 1 \) beyond which the behavior of system for all \( i \geq d + 1 \) is expected to be possibly same as the behavior of the system for \( d + 1 \). Such similarity need not hold for \( 0, 1, \ldots, d \). Therefore, we can represent the system by storing the information for the initial portion \( 0, 1, \ldots, d \) and the repeating portion \( d + 1, d + 2, \ldots \). It is tedious to obtain a closed form solution for the QBD process presented in section 3. In order to obtain an
efficient and numerical stable solution, we employ matrix geometric method to obtain the probabilities for the Markov chain. As there are repetitive block sub-matrices in transition rate matrix, we can easily employ the matrix geometric method to evaluate the stationary probability vector $P_n$.

The transition rate matrix $Q$ of the Markov chain corresponding to the coefficients of equations (1) to (8) has the block tri-diagonal form given by:

$$Q = \begin{pmatrix}
A_0 & C \\
B_1 & A_1 & C \\
& B_2 & A_2 & C \\
& & & \ddots & \ddots \\
B_d & A_d & C \\
B_{d+1} & A_{d+1} & C \\
B_{d+1} & A_{d+1} & C \\
& & & \ddots & \ddots 
\end{pmatrix}$$

The rate matrix $Q$ of the QBD process has the sub-matrices given as:

$$A_i = \begin{pmatrix}
-(\lambda + i\mu) \\
q_2\eta \\
-(\lambda + q_2\eta + \xi)
\end{pmatrix}, \quad 0 \leq i \leq d,$$

$$= \begin{pmatrix}
-(\lambda + d\mu + \xi) \\
q_2\eta \\
-(\lambda + q_2\eta + \xi)
\end{pmatrix}, \quad i > d,$$

$$B_i = \begin{pmatrix}
(1 - q_1)i\mu & iq_1\mu \\
0 & 0
\end{pmatrix}, \quad 1 \leq i \leq d,$$

$$= \begin{pmatrix}
(1 - q_1)d\mu + \xi & dq_1\mu \\
0 & \xi
\end{pmatrix}, \quad i > d,$$

$$C = \begin{pmatrix}
\lambda & 0 \\
0 & \lambda
\end{pmatrix}.$$  

Let $P$ be the corresponding steady state probability vector of $Q$. By partitioning the vector $P$ as $P = \{P_0, P_1, P_2, \ldots\}$ where $P_n = [P_{ij}, Q_{ij}], \quad i \geq 0, \quad 0 \leq j \leq d$.

The vector $P$ satisfies $PQ = 0$ and $Pe = 1$ where $e$ is a $2 \times 1$ vector having each element as unity. According to [15], the system is stable and the steady state probability vector exists if and only if $yCe < yB_{d+1}e$ where $y$ is an invariant probability of the matrix $M = A_{d+1} + B_{d+1} + C$ such that $y$ satisfies the equations $yM = 0$ and $ye = 1$. Apparently, when the stability condition is satisfied, the sub vectors of $P$, corresponding to different levels satisfy

$$P_n = P_{d+1}R^{n-(d+1)}, \quad n \geq d + 2,$$  \hspace{1cm} (9)

where matrix $R$ is the minimal non-negative solution of the matrix quadratic equation

$$C + RA_{d+1} + R^2B_{d+1} = 0,$$  \hspace{1cm} (10)

which can be obtained by using the following iterative procedure.

**Computational algorithm for $R$:**
Step 1: Set \( k = 1 \).
Step 2: Set \( U = A_{d+1} \) and calculate \( G = (I - U)^{-1}B_{d+1} \).
Step 3: Increment \( k \) by 1.
Step 4: Replace \( U = A_{d+1} + CG \) and \( G = (I - U)^{-1}B_{d+1} \).
Step 5: Repeat steps 3 and 4 until \( ||e - Ge||_\infty < \epsilon \), where \( \epsilon \) is a stopping tolerance.
Step 6: Calculate \( R = C(I - U)^{-1} \).

From the equation \( PQ = 0 \) the governing system of difference equations can be given as

\[
P_0A_0 + P_1B_1 = 0, \quad (11)
\]
\[
P_{n-1}C + P_nA_n + P_{n+1}B_{n+1} = 0, \quad 1 \leq n \leq d, \quad (12)
\]
\[
P_{n-1}C + P_nA_{d+1} + P_{n+1}B_{d+1} = 0, \quad n \geq d + 1, \quad (13)
\]

and normalizing condition

\[
\sum_{n=0}^{\infty} P_n e = 1. \quad (14)
\]

From equations (11) to (13), after some mathematical manipulations, we get

\[
P_0 = -P_1B_1(A_0^{-1}) = P_1\phi_1, \quad (15)
\]
\[
P_{n-1} = -P_nB_n(A_{n-1} + \phi_{n-1}C)^{-1} = P_n\phi_n, \quad 2 \leq n \leq d + 1, \quad (16)
\]
\[
P_{d+1}\phi_{d+1}C + P_{d+1}A_{d+1} + P_{d+1}RB_{d+1} = 0. \quad (17)
\]

Thus, \( P_n \) \((0 \leq n \leq d)\) in equation (15) and (16) can be written as product form in terms of \( P_{d+1} \). To find \( P_{d+1} \), we use the normalization condition and equation (17).

\[
\sum_{n=0}^{\infty} P_n e = P_{d+1}\left(\sum_{j=1}^{d+1} \prod_{i=d+1} m \phi_i + (I - R)^{-1}\right)e = 1 \quad (18)
\]

Solving equations (17) and (18), we obtain \( P_{d+1} \). We use equations (9),(15) and (16), to get \( P_n \) for \( n \geq 0 \).

5. Performance measures

Performance measures show the behaviour of the system from various perspectives. Expected number of customers in the system and in the queue, respectively, when the server is busy in FES are given by

\[
E[F] = \sum_{n=1}^{d}nP_{0,n} + \sum_{n=1}^{\infty}(n + d)P_{n,d};
\]
\[
E[QF] = \sum_{n=1}^{d}nP_{0,n} + \sum_{n=1}^{\infty}nP_{n,d}.
\]
Expected number of customers in the system and the queue, respectively, when the server is rendering SOS are given by

\[ E[S] = d \sum_{n=1}^{\infty} nQ_{0,n} + \sum_{n=1}^{\infty} (n + d)Q_{n,d}; \]

\[ E[QS] = d \sum_{n=1}^{\infty} nQ_{0,n} + \sum_{n=1}^{\infty} nQ_{n,d}. \]

Expected reneging rate of the customer is given by

\[ E[RC] = \sum_{n=1}^{\infty} \xi P_{n,d} + \sum_{n=1}^{\infty} \xi Q_{n,d}. \]

Probability that the server is busy in FES is given by

\[ P[F] = d \sum_{n=1}^{\infty} P_{0,n} + \sum_{n=1}^{\infty} P_{n,d}. \]

Probability that the server is busy in SOS is given by

\[ P[S] = d \sum_{n=1}^{\infty} Q_{0,n} + \sum_{n=1}^{\infty} Q_{n,d}. \]

6. Numerical investigation

To demonstrate the applicability of the formulae obtained in the previous sections, numerical computations have been carried out using Mathematica software. Some of the numerical computations have been presented below in the form of tables and graphs. The parameters of the model are assumed to be \( \lambda = 0.26, \mu = 0.4, \eta = 0.6, q_1 = 0.7, q_2 = 0.5, \xi = 2.0, d = 5, n = 25. \)

Table 1 shows the effect of service rate during FES (\( \mu \)) on performance measures. We can see that as (\( \mu \)) increases, \( E[F] \) and \( P[F] \) and \( E[RC] \) decrease. Also the percent variation, P.V. (which is the difference between the last value and the first value expressed as percentage) in the last row indicates the decreasing trend. This is because, some customers may opt for SOS after completing FES resulting in the decrease in the above performance measures. On the other hand, increasing trend is observed for \( E[S] \) and \( P[S] \) as \( \mu \) increases. Also the percent variation in the last row indicates the increasing trend. This is due to feedback of customers from SOS to primary queue.
Table 1. Effect of $\mu$ on performance measures.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>1.29375</td>
<td>1.53028</td>
<td>0.399469</td>
<td>0.545636</td>
<td>0.216025</td>
</tr>
<tr>
<td>0.6</td>
<td>0.981279</td>
<td>1.76571</td>
<td>0.229791</td>
<td>0.621838</td>
<td>0.182103</td>
</tr>
<tr>
<td>0.8</td>
<td>0.789019</td>
<td>1.90928</td>
<td>0.182436</td>
<td>0.667222</td>
<td>0.158631</td>
</tr>
<tr>
<td>1.0</td>
<td>0.659311</td>
<td>2.00573</td>
<td>0.15117</td>
<td>0.697222</td>
<td>0.14153</td>
</tr>
<tr>
<td>1.2</td>
<td>0.566051</td>
<td>2.07492</td>
<td>0.129017</td>
<td>0.71851</td>
<td>0.128543</td>
</tr>
</tbody>
</table>

P.V. -72.76 54.46 -18.04 17.28 -8.7

Table 2. Effect of $\eta$ on performance measures.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>1.24775</td>
<td>1.56647</td>
<td>0.294485</td>
<td>0.560017</td>
<td>0.213552</td>
</tr>
<tr>
<td>0.7</td>
<td>1.3376</td>
<td>1.49557</td>
<td>0.323724</td>
<td>0.531931</td>
<td>0.218371</td>
</tr>
<tr>
<td>0.9</td>
<td>1.4194</td>
<td>1.43026</td>
<td>0.350235</td>
<td>0.506383</td>
<td>0.222715</td>
</tr>
<tr>
<td>1.1</td>
<td>1.49412</td>
<td>1.36999</td>
<td>0.374364</td>
<td>0.483067</td>
<td>0.22665</td>
</tr>
<tr>
<td>1.3</td>
<td>1.56258</td>
<td>1.31428</td>
<td>0.396404</td>
<td>0.46172</td>
<td>0.230229</td>
</tr>
</tbody>
</table>

P.V. 31.48 -25.21 10.19 -9.82 1.66

Table 3. Effect of $q_1$ on performance measures.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>1.94091</td>
<td>0.927128</td>
<td>0.506299</td>
<td>0.347771</td>
<td>0.244504</td>
</tr>
<tr>
<td>0.5</td>
<td>1.55415</td>
<td>1.27705</td>
<td>0.387072</td>
<td>0.46659</td>
<td>0.226434</td>
</tr>
<tr>
<td>0.7</td>
<td>1.29375</td>
<td>1.53028</td>
<td>0.309469</td>
<td>0.545636</td>
<td>0.216025</td>
</tr>
<tr>
<td>0.9</td>
<td>1.10679</td>
<td>1.72431</td>
<td>0.255838</td>
<td>0.601257</td>
<td>0.209926</td>
</tr>
</tbody>
</table>

P.V. -83.412 79.718 -25.0461 25.3486 -3.4578

Table 4. Effect of $q_2$ on performance measures.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>1.17945</td>
<td>1.61975</td>
<td>0.272179</td>
<td>0.581374</td>
<td>0.209855</td>
</tr>
<tr>
<td>0.5</td>
<td>1.29375</td>
<td>1.53028</td>
<td>0.309469</td>
<td>0.545636</td>
<td>0.216025</td>
</tr>
<tr>
<td>0.7</td>
<td>1.39565</td>
<td>1.44929</td>
<td>0.342546</td>
<td>0.5138</td>
<td>0.22665</td>
</tr>
<tr>
<td>0.9</td>
<td>1.48694</td>
<td>1.37581</td>
<td>0.396404</td>
<td>0.46172</td>
<td>0.230229</td>
</tr>
</tbody>
</table>


Figure 1 shows the effect of arrival rate $\lambda$ on $E[F]$ for different values of $\mu$. We can see from the graph that as $\lambda$ increases, $E[F]$ decreases for a constant service rate in FES($\mu$). But as $\mu$ increases, the number of customers served per unit time increases and hence $E[F]$ decreases. Figure 2 shows the effect of $\lambda$ on $E[S]$ for different values of $\eta$. Generally as $\lambda$ increases, $E[S]$ also increases. But when the service rate in SOS ($q_1$) increases, it is obvious that $E[S]$ decreases.

The effect of $\mu$ on $E[F]$ and $E[S]$ is plotted in Figure 3. It is clear from the figure that as $\mu$ increases $E[F]$ decreases and $E[S]$ increases. This is due to the fact that after getting FES some customers tend to SOS. Moreover, for fixed value of $\mu$, the intersecting point of the curves of $E[F]$ and $E[S]$ gives the value $q_2 = 0.59$ for which $E[F]$ is minimum and $E[S]$ is maximum.

The effect of $\eta$ on $E[S]$ and $E[F]$ for different values of $q_2$ is plotted in Figure 4 and Figure 5, respectively. Generally, as $\eta$ increases, $E[S]$ decreases and $E[F]$ increases. Also as $q_2$ increases, number of customers again join the primary queue increases. Hence we observe decreasing trend for $E[S]$ in Figure 4 and increasing trend for $E[F]$ in Figure 5 as $q_2$ increases.
Figure 1. Effect of $\lambda$ on $E[F]$ for different $\mu$.

Figure 2. Effect of $\lambda$ on $E[S]$ for different $\eta$.

Now, we develop a total expected cost function per unit time with an objective to determine the optimum value of $\mu$ that minimizes the expected cost function. Let us define:

- $C_{ef}$ Cost per unit time when the server is busy in FES,
- $C_{es}$ Cost per unit time when the server is busy in SOS,
- $C_f$ Cost per unit time when the server is serving in FES,
- $C_s$ Cost per unit time when the server is serving in SOS,
- $C_{rc}$ Cost per unit time per lost customer due to reneging.

The cost minimization problem is expressed mathematically as:

$$F[\mu] = C_{ef}E[F] + C_{es}E[S] + C_f \mu + C_s \eta + C_{rc}E[RC].$$

Assuming the coefficients of the cost function as $C_{ef} = 4, C_{es} = 2.5, C_f = 3, C_s = 1.5, C_{rc} = 10$ and the model parameters as $\lambda = 0.26, \mu = 0.4, \eta = 0.6, q_1 = 0.7, q_2 = 0.4, \xi = 2.0, d = 5, n = 25$. Varying the service rate in FES ($\mu$) such that $0.4 \leq \mu \leq 1.2$. Using direct search method, it is found that the minimum value of the total expected cost function is $F[\mu] = 12.8063$ at $\mu = 0.7$ which is shown graphically in Figure 6.

7. Conclusion

This paper presents a single server infinite buffer queue with second optional service in which customers may renge due to impatience and feedback of customers
may exists due to dissatisfaction with the service. Such models have applications in hospital services, production systems, bank services, computer and communication networks, etc. We have used matrix geometric method to find steady state probabilities and presented some performance measures along with the numerical results. Optimization of cost function with respect to the service rate $\mu$ is also discussed using direct search method.
Acknowledgements

The authors are thankful to the referees for their valuable comments and suggestions which have helped in improving the quality of the presentation of this paper.

References