An M/M/2/N Queuing System with Encouraged Arrivals, Heterogenous Service and Retention of Impatient Customers

Bhupender Kumar Som[1]
Associate Professor
Jagan Institute of Management Studies, Sec-5, Rohini,
New Delhi - 110085 (India)
bksoam@live.com

Sunny Seth
Assistant Professor
Jagan Institute of Management Studies, Sec-5, Rohini,
New Delhi - 110085 (India)
sunnyseth2005@gmail.com

Abstract

In this paper we study a finite capacity Markovian queuing system with heterogeneous service, encouraged arrivals, reneging and retention of reneged customers. The stationary system size probabilities are obtained recursively. The important measures of performance such as expected system size, the average rate of reneging and average rate of retention are obtained. Numerical Illustrations are presented. Finally, the economic analysis of the model is performed by introducing cost model.

Keywords– Heterogeneous server queues, encouraged arrivals, reneging, retention, impatient customers.

1 Introduction and Literature Survey

Globalization has introduced never ending competition in business. Customers have become more selective and as a result brand switching is more frequent. Understanding customer behavior is essential for organizations in today's competitive business environment. Generally following customer behaviors are studied in queuing theory: 1) Balking 2) Reneging 3) Jockeying 4) Collusion. First customer behavior in queues is observed by [Barrer, 1957]. He mentioned that each customer is available for service only for a fixed limited time.[Haight, 1957] considered an M/M/1 queue with balking and [Haight, 1959] also proposed an M/M/1 queue with reneging. The combined effects of balking and reneging in an M/M/1/N queue have been studied by [Ancker and Gafarian, 1963a and 1963b]. In order to stay ahead in the competition, business organizations introduce various discounts and offers to attract customers. These attracted customers are termed as encouraged arrivals in this paper. However [Jain, et. al., 2014] coined a term reverse balking. They mentioned that an arriving customer may get attracted towards a system by looking in to large customer base. While reverse balking deals with probability of joining or not joining the system, encouraged arrivals deal with percentage increase in customers' arrival due to offers and discounts. Encouraged arrivals often result in heavy rush. Due to this customers have to wait longer in queues. Long waiting times, results in customer impatience and a customer may decide to abandon the facility without completion of service, termed as reneging, as mentioned by [Ancher and Gafarian, 1963]. Reneging results in loss of goodwill and revenue both. The phenomenon of encouraged arrivals can also be understood as contrary to discouraged arrivals as discussed by [Kumar and Sharma, 2014a]. They considered a two heterogeneous server queuing system with discouraged arrivals and retention of reneged customers. They derived the steady-state solution of the model. [Raynolds, 1968] presented multi-server queuing model with discouragement. He obtained equilibrium distribution of queue length and derived other performance measures from it. In this paper we consider that customers are served through two servers with heterogeneous service rates. The pioneer work in queuing theory on heterogeneous servers is done by [Morse, 1958]. [Krishnamoorthi, 1963] studied the Poisson queue with two heterogeneous servers, where the stationary probabilities were solved via differential difference equations. [Singh, 1970] studied two-server Markovian queues with balking and heterogeneous vs. homogeneous servers. A cost model was discussed in his paper and he studied average characteristics of both the homogeneous and heterogeneous systems.

To ensure smooth functioning of the system experiencing above mentioned phenomenons, an effective strategy should be designed. If the performance of the system can be measured in advance with some probability, an effective management policy can be designed and implemented.
Hence, In this paper we develop a stochastic queuing model addressing all practically valid and contemporary challenges mentioned above. The model provides steady-state solution and economic behavior of the system. Rest of the paper is arranged as follows: in section 2, model is described, its mathematical formulation and steady state solution are presented; section 3 deals with performance measures and in section 4, numerical illustrations are presented; economic analysis of the model is performed in section 5; the paper is concluded in section 6.

2 Model Formulation and Steady State Solution

The arrivals in the model are termed as encouraged arrivals. The arrivals occur one by one in accordance to Poisson process with parameter $\lambda(1+\eta)$, where η represents the percentage increase in number of customers calculated from past or observed data. For instance, if in past an organization offered discounts and the percentage increase in number of customers was observed 50 percent or 150 percent then η=0.5 or η=1.5 respectively. Customers are serviced by two servers with different service rates μ_1 and μ_2 respectively. The service times follow exponential distribution. An arriving customer joins the system with probability π_1 in front of server 1 and with $\pi_2 = 1 - \pi_1$ in front of server 2. The capacity of the system is finite, say N. The queue discipline is first come, first served. A customer waiting for his service in the queue may get impatient after some time and reneges. The reneging times are exponentially distributed with parameter ξ. The probability of retention of a reneged customer is q and the probability that customer is not retained is $p = 1 - q$. In this section we present the mathematical model of the system. First of all we define some notations:
Let $P_n(t)$ = the probability that there are n customers in the system at time t.
$P_{10}(t)$ = the probability that the first server is engaged and the second server is free with no waiting line at time t.
$P_{01}(t)$ =the probability that the first server is free and the second server is engaged with no waiting line at time t.
$P_{00}(t)$ = the probability that the system is empty at time t.
Also, $P_2(t) = P_{11}(t)$ and $P_1(t) = P_{10}(t) + P_{01}(t)$

The differential-difference equations of the model are:

$$\frac{d}{dt}P_{00}(t) = -\lambda(1+\eta)P_{00}(t) + \mu_1 P_{10}(t) + \mu_2 P_{01}(t) \tag{1}$$

$$\frac{d}{dt}P_{10}(t) = -(\lambda(1+\eta)+\mu_1)P_{10}(t) + \mu_2 P_{11}(t) + \lambda(1+\eta)\pi_1 P_{00}(t) \tag{2}$$

$$\frac{d}{dt}P_{01}(t) = -(\lambda(1+\eta)+\mu_2)P_{01}(t) + \mu_1 P_{11}(t) + \lambda(1+\eta)\pi_2 P_{00}(t) \tag{3}$$

$$\frac{d}{dt}P_{2}(t) = -(\lambda(1+\eta)+\mu_1+\mu_2)P_{2}(t) + (\mu_1+\mu_2+\xi p)P_3(t) + \lambda(1+\eta)P_1(t) \tag{4}$$

$$\frac{d}{dt}P_{n}(t) = -(\lambda(1+\eta)+\mu_1+\mu_2+(n-2)\xi p)P_{n}(t) + (\mu_1+\mu_2+(n-1)\xi p)P_{n+1}(t) + \lambda(1+\eta)P_{n-1}(t) \tag{5}$$

$$\frac{d}{dt}P_{N}(t) = \lambda(1+\eta)P_{N-1}(t) - (\mu_1+\mu_2+(N-2)\xi p)P_N(t) \tag{6}$$

Steady-state solution of the model
In steady-state equations (1) to (6) become:

$$0 = -\lambda(1+\eta)P_{00} + \mu_1 P_{10} + \mu_2 P_{01} \tag{7}$$

$$0 = -(\lambda(1+\eta)+\mu_1)P_{10} + \mu_2 P_{11} + \lambda(1+\eta)\pi_1 P_{00} \tag{8}$$

$$0 = -(\lambda(1+\eta)+\mu_2)P_{01} + \mu_1 P_{11} + \lambda(1+\eta)\pi_2 P_{00} \tag{9}$$

$$0 = -(\lambda(1+\eta)+\mu_1+\mu_2)P_{2} + (\mu_1+\mu_2+\xi p)P_3 + \lambda(1+\eta)P_1 \tag{10}$$

$$0 = -(\lambda(1+\eta)+\mu_1+\mu_2+(n-2)\xi p)P_{n} + (\mu_1+\mu_2+(n-1)\xi p)P_{n+1} + \lambda(1+\eta)P_{n-1} \tag{11}$$

$$0 = \lambda(1+\eta)P_{N-1} - (\mu_1+\mu_2+(N-2)\xi p)P_N \tag{12}$$

Solving equation (7) to (12), we obtain

$$P_{10} = \left(\frac{(\lambda(1+\eta)) + (\mu_1+\mu_2)\pi_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{\lambda(1+\eta)}{\mu_1}P_{00} \tag{13}$$

$$P_{01} = \left(\frac{(\lambda(1+\eta)) + (\mu_1+\mu_2)\pi_2}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{\lambda(1+\eta)}{\mu_2}P_{00} \tag{14}$$

Adding (13) and (14), we get

$$P_{1} = \left(\frac{\lambda(1+\eta) + \pi_1\mu_2 + \pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{\lambda(1+\eta)(\mu_1+\mu_2)}{\mu_1\mu_2}P_{00} \tag{15}$$

Adding equation (8) and (9) and using (7) and (15), we get

$$P_{2} = \left(\frac{\lambda(1+\eta) + \pi_1\mu_2 + \pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{\lambda(1+\eta)}{\mu_1}\frac{(\lambda(1+\eta))}{\mu_2}P_{00} \tag{16}$$

Using equation (11) and generalizing, we obtain

$$P_{n} = \left(\frac{\lambda(1+\eta) + \pi_1\mu_2 + \pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{\lambda(1+\eta)}{\mu_1}\frac{\lambda(1+\eta)}{\mu_2}\prod_{k=3}^{n}\left(\frac{\lambda(1+\eta)}{\mu_1+\mu_2+(k-2)\xi p}\right)P_{00} \tag{17}$$

Using (12) and (17), we obtain

$$P_N = \left(\frac{\lambda(1+\eta)+\pi_1\mu_2+\pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{\lambda(1+\eta)}{\mu_1}\frac{\lambda(1+\eta)}{\mu_2}\prod_{k=3}^{N}\left(\frac{\lambda(1+\eta)}{\mu_1+\mu_2+(k-2)\xi p}\right)P_{00} \quad (18)$$

Using condition of normality $\sum_{n=1}^{N} P_n = 1$, we get

$$P_0 = \Bigg[1 + \left(\frac{\lambda(1+\eta)+\pi_1\mu_2+\pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{\lambda(1+\eta)(\mu_1+\mu_2)}{\mu_1\mu_2} + \left(\frac{\lambda(1+\eta)+\pi_1\mu_2+\pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{[\lambda(1+\eta)]^2}{\mu_1\mu_2}$$
$$+ \sum_{n=3}^{N-1}\left(\frac{\lambda(1+\eta)+\pi_1\mu_2+\pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{[\lambda(1+\eta)]^2}{\mu_1\mu_2}\prod_{k=3}^{n}\left(\frac{\lambda(1+\eta)}{\mu_1+\mu_2+(k-2)\xi p}\right)$$
$$+ \left(\frac{\lambda(1+\eta)+\pi_1\mu_2+\pi_2\mu_1}{2\lambda(1+\eta)+\mu_1+\mu_2}\right)\frac{[\lambda(1+\eta)]^2}{\mu_1\mu_2}\prod_{k=3}^{N}\left(\frac{\lambda(1+\eta)}{\mu_1+\mu_2+(k-2)\xi p}\right)\Bigg]^{-1} \quad (19)$$

3 Measures of Performance

In this section we derive some important measures of performance:

3.1 Expected System Size ($\mathbf{L}_s$):

$$L_s = \sum_{n=1}^{N} nP_n = P_1 + 2P_2 + \sum_{n=3}^{N-1} nP_n + NP_N$$

3.2 Expected queue length ($\mathbf{L}_q$):

$$L_q = \sum_{n=3}^{N}(n-2)P_n = \sum_{n=3}^{N-1}(n-2)P_n + (N-2)P_N$$

3.3 Average rate of reneging ($\mathbf{R}_r$):

$$R_r = \sum_{n=3}^{N}(n-2)\xi pP_n = \sum_{n=3}^{N-1}(n-2)\xi pP_n + (N-2)\xi pP_N$$

3.4 Average rate of retention ($\mathbf{R}_R$):

$$R_R = \sum_{n=3}^{N}(n-2)\xi qP_n = \sum_{n=3}^{N-1}(n-2)\xi qP_n + (N-2)\xi qP_N$$

4 Numerical Illustrations

In this section we present numerical illustrations of the model:

Variation in measures of performance with respect to arrival rate.
Taking $\mu_1 = 2, \mu_2 = 3, \pi_1 = 0.4, q = 0.4, \xi = 0.2$ and $N = 10$.

Table - 1

λ	L_s	L_q	R_r	R_R
1	0.65808	0.05573	0.00669	0.00446
1.2	0.82004	0.09892	0.01187	0.00791
1.4	1.00085	0.16219	0.01946	0.01298
1.6	1.20577	0.25115	0.03014	0.02009
1.8	1.44072	0.37219	0.04466	0.02978
2	1.71191	0.53220	0.06386	0.04258
2.2	2.02488	0.73764	0.08852	0.05901
2.4	2.38322	0.99330	0.11920	0.07946
2.6	2.78702	1.30072	0.15609	0.10406
2.8	3.23168	1.65680	0.19882	0.13254
3	3.70763	2.05333	0.24640	0.16427
3.2	4.20125	2.47766	0.29732	0.19821
3.4	4.69696	2.91454	0.34975	0.23316
3.6	5.17965	3.34863	0.40184	0.26789
3.8	5.63681	3.76661	0.45199	0.30133
4	6.05970	4.15857	0.49903	0.33269
4.2	6.44351	4.51838	0.54221	0.36147
4.4	6.78684	4.84334	0.58120	0.38747
4.6	7.09078	5.13336	0.61600	0.41067

From table 1 we can see that as the arrival rate increases, the expected system size, expected length of the queue, average rate of reneging as well as average retention rate all increases.

Variation in measures of performance with respect to reneging rate.
Taking $\mu_1 = 2, \mu_2 = 3, \pi_1 = 0.4, q = 0.4, \lambda = 3$ and $N = 10$.

Table - 2

ξ	L_s	R_r
0.1	3.97104	0.13758
0.15	3.83311	0.19506
0.2	3.70763	0.24640
0.25	3.59327	0.29247
0.3	3.48883	0.33400
0.35	3.39324	0.37161
0.4	3.30556	0.40583
0.45	3.22493	0.43711
0.5	3.15061	0.46583
0.55	3.08194	0.49231
0.6	3.01834	0.51681
0.65	2.95929	0.53959
0.7	2.90436	0.56083
0.75	2.85313	0.58071
0.8	2.80525	0.59938
0.85	2.76042	0.61695
0.9	2.71835	0.63356
0.95	2.67880	0.64928
1	2.64154	0.66421

From table 2 we can see that the expected system size decreases while the average rate of reneging increases with the increase in reneging rate.

Variation in measures of performance w.r.t. the prob. of retention.
Taking $\mu_1 = 2, \mu_2 = 3, \pi_1 = 0.4, \xi = 0.2, \lambda = 3$ and $N = 10$.

Table - 3

q	L_s	R_R
0.1	3.48883	0.03711
0.15	3.52261	0.05658
0.2	3.55741	0.07670
0.25	3.59327	0.09749
0.3	3.63023	0.11899
0.35	3.66833	0.14124
0.4	3.70763	0.16427
0.45	3.74816	0.18811
0.5	3.78997	0.21281
0.55	3.83311	0.23841
0.6	3.87763	0.26494
0.65	3.92359	0.29247
0.7	3.97104	0.32103
0.75	4.02003	0.35067
0.8	4.07063	0.38145
0.85	4.12288	0.41342
0.9	4.17685	0.44664
0.95	4.23261	0.48117
1	4.29021	0.51707

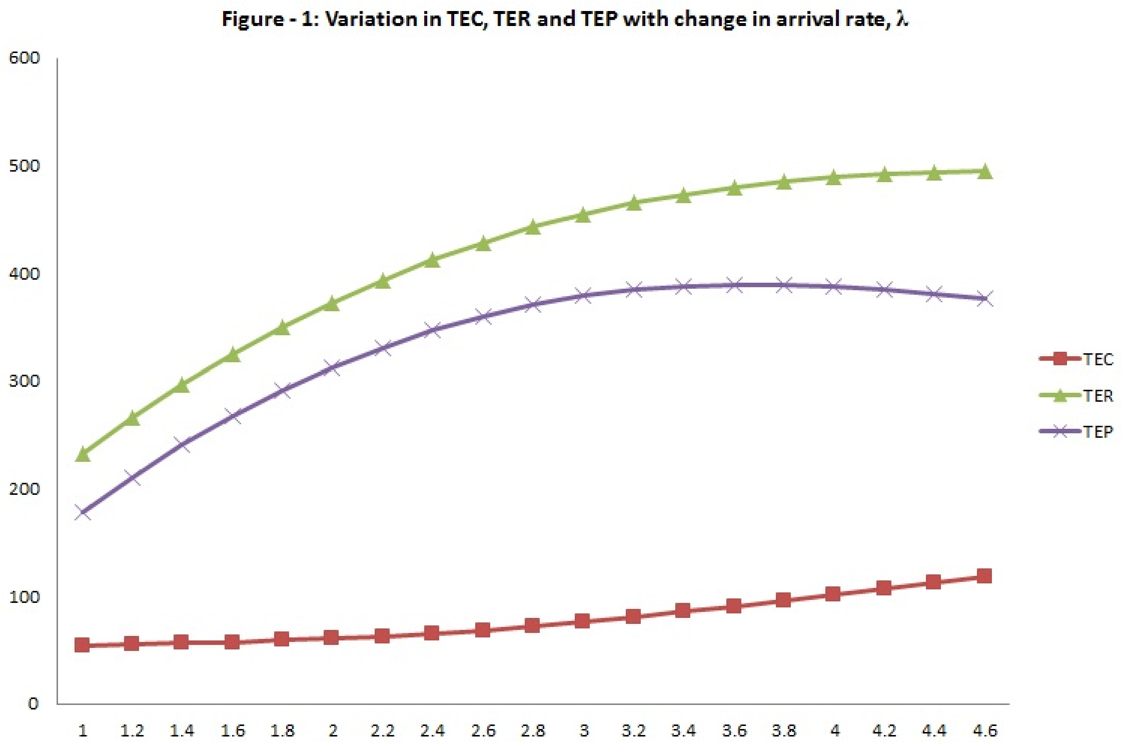
From table 3 we can see that as the probability of retention increases, the expected system size as well as average rate of retention increases.

5 Economic Analysis of the Model

In this section we develop cost model of the above queuing model to perform economic analysis by developing the functions of Total Expected Cost (TEC), Total Expected Revenue (TER) and Total Expected Profit (TEP). Following notations are used in this cost model:
C_s =Cost per service per unit time.
C_h =Holding cost per unit per unit time.
C_L =Cost associated with each lost unit per unit time.
C_r =Cost associated with each reneged unit per unit time.
C_R =Cost associated with each retained unit per unit time.
R = Revenue earned per unit per unit time.

Cost model functions are as follows:

$$TEC = C_h L_s + C_L \lambda (1+\eta) P_N + C_r R_r + C_R R_R + C_s(\mu_1 + \mu_2)$$
$$TER = R \times (\mu_1 + \mu_2) \times (1 - P_0)$$
$$TEP = TER - TEC$$



From figure 1 we can see that as the arrival rate increases, Total Expected revenue increases and as a result total expected profit also increases but after certain level total expected cost increases to a greater extent than revenue which results in slight decrease in total expected profit after certain rate of arrival.

6 Conclusions and Future Scope

The results are of immense use for any business organization in order to manage operations effectively. The model gives a clear insight of expected system size, length of the queue, probability of any number of customers in the system, probability that the system is vacant and any other related operational measures of performance under the assumptions. Expected system size gives average load on the facility that has to be managed per unit time. Average length of queue enables the system to keep queue under control as a longer queue may result in to panic and the firm may lose their potential customers, which inturn hampers the goodwill. Average rate of reneging shall be reduced once it is known. Total expected cost, revenue and profit provides economical insight of the system. The model can be adopted and implemented to measure overall performance of the system numerically.
By implementing this model, efficiency of any business organization can be improved and better service can be provided to the customers.

The model analysis is limited to finite capacity. The infinite capacity case of the model can also be studied. The paper may be extended to more than two heterogeneous servers. Further, the model can be solved in transient state to get time-dependent results.

References

1. Barrer, D. Y. "Queuing with Impatient Customers and Indifferent Clerks" Operations Research, 5(5), 644-649 (1957).
2. Haight, F. A. "Queueing with balking" Biometrika, 44, 360-369 (1957).
3. Morse, P.M., Queues, inventories and maintenance, Willey, New York (1958).

4. Haight, F. A., Queuing with reneging, Metrika, 2, 186-197 (1959).
5. Ancker Jr., C. J. and Gafarian, A. V., Some queuing problems with balking and reneging I, Operations Research, 11, 88-100 (1963a).
6. Ancker Jr., C. J. and Gafarian, A. V., Some queuing problems with balking and reneging II, Operations Research, 11, 928-937 (1963b).
7. Krishanamoorthy, B., On Poisson queues with heterogeneous servers, Operations Research, 11, 321-330 (1963).
8. Reynolds, John F. "The Stationary Solution of a Multiserver Queuing Model with Discouragement." Operations Research, 16(1), 64-71 (1968).
9. Singh, V.P., Two-server Markovian queues with balking: heterogeneous vs homogeneous servers, Operations Research. 18, 145-159 (1970).
10. Jain, N. K., Kumar, R. and Som, B. K., An M/M/1/N queuing system with Reverse balking, American Journal of Operational Research, 4(2), 17-20 (2014).
11. Kumar, R. and Sharma, S. K., Two-Heterogeneous Server Markovian Queueing Model with Discouraged Arrivals, Reneging and Retention of Reneged Customers, International Journal of Operations Research, 11, 64-68 (2014a).