Journal of Computer and Mathematical Sciences, *Vol.10(1),112-120 January 2019* *ISSN 0976-5727 (Print)*
(An International Research Journal), *www.compmath-journal.org* *ISSN 2319-8133 (Online)*

Analysis and Comparative Study of Various Performance Measures of M/G/1 and M/G/S Queuing Models

Pradeep K. Joshi[1*], Shejal Gupta[2] and K. N. Rajeshwari[3]

[1]Reader, Department of Mathematics,
IPS Academy, Indore, INDIA.
[2]Research Scholar,
School of Mathematics, DAVV, Indore, INDIA.
[3]Professor & Head,
School of Mathematics, DAVV, Indore, INDIA.

(Received on: January 11, 2019)

ABSTRACT

Queuing theory is a mathematical tool, which deals with the phenomenon of waiting lines. Keeping customers waiting for too long, affects various parameters of queuing theory. In this paper, optimization technique has been applied to calculate the performance measures like waiting time, traffic intensity, waiting cost and proportion of time the server is idle for single server M/G/1 model as well as for multiple servers M/G/S model. Also, we analysis the comparative study of various performance measures of single server M/G/1 model and multiple servers M/G/S queuing model. This investigation deals with M/G/1 and M/G/S Models in which arrival follows Poisson distribution, service time follows general distribution, and Queue discipline is FCFS. The result of analysis showed that as we increase the number of servers, the waiting time in the queue decreases, total cost of the system decreases, traffic intensity also decreases whereas idle time increases. The results are optimize and effective. The graphical representation of various performance measures and their comparison are shown by MATLAB software.

Keywords: M/G/1 Model, M/G/S Model, Queue, Traffic Intensity, waiting time, waiting cost, service cost.

1. INTRODUCTION

Queuing theory is associated with the waiting lines. It is mathematical approach to analysis of waiting lines. All of us in our day-to-day life face the problem of queues such as in banks, ATMs, hospitals, traffic signal to change, telephone call handled by an operator and many more. The queuing theory or waiting line theory owes its development to A.K. Erlang. He in 1903, took up the problem on congestion of telephone traffic. Firstly, he tried to find the delay for one operator and then he works for the several operators. After World War II, this work has been extended widely in different areas where waiting line problem was there.[1,2].

Many authors studied regarding determination of various performance measures of M/G/I queuing models.[4-6].

Authors discussed regarding unreliable server M/G/1 queque with multi-optional services and impatient customers.[7]

Authors studied on M/G/I queuing model with state dependent arrivals and vacation.[13]

Authors discussed on his paper to minimize service cost and cost of waiting time using M/M/I and M/M/S queuing model.[14]

Waiting for too long in the queue is not an ideal situation. Queuing analysis is imported because customers regards waiting as non-value added activity. Waiting time depends on number of customers in the queue, number of servers, the service time taken by individual customer as well as by service provider. To minimize the waiting time and the total cost of the system is the aim of optimization of queuing models. Mathematical optimization methods are often used as a tool to reduce service time and cost of implementation processes. Service cost and waiting cost are the two basic costs in the queuing theory. Service cost is the cost of providing service while waiting cost is the cost due to delay in the services to the customers which results customer get irritate and may switch to the competing organizations. Thus, it costs very high due to loss of business and loss of goodwill.

The most important measures is system utilization, which referrers to percentage of capacity utilized. Our aim is to increase the system utilization and minimize idle time.

2. COMPONENTS OF A QUEUING SYSTEM

The following are the main components of queuing system -

(1) **Arrival Distribution**: - It represents the pattern in which the number of customers arrive at the service facility. Arrivals may also be represented by the inter-arrival time, which is the period between two successive arrivals. The rate at which customers arrive to be serviced is called arrival rate denoted by λ. Arrivals may be random which are often best described by the Poisson distribution or it can be equal interval of time.

(2) **Service Distribution: -** It represents the pattern in which the number of customers leaves the service facility. Departures may also be represented by the service time which is the time period between two successive services. Service time may be constant or variable but known or random. The rate at which number of customers served per unit of time is called service rate denoted by μ.

(3) **Service Channels**: - The queuing system may have a single service channel or may have multi-server service channel. Arriving customers may form one line and get serviced. The system may have a number of service channels, which may be arranged in parallel or in

series or a complex combination of both. A queuing model is called one server model when the system has one server only and a multi-server model when the system has a more than one server that means the system has number of parallel channels each with one server.

(4) **Service Discipline**: - Discipline of a queuing system means the rule that a server uses to choose the next customer from the queue (if any) when the server completes the service of the current customer. Commonly used queue disciplines are: FCFS - Customers are served on a First Come First Served basis. LIFO - Customers are served in a last-in first-out manner. Priority - Customers are served in order of their importance on the basis of their service requirements, SIRO- Service in Random order.

3. MATHEMATICAL MODELS

3.1 Single Server M/G/1 Model

In single-server model, there can be only one service channels or server to provide the service to the customers. This investigation deals with a single server M/G/1 queuing model in which arrivals follows poisson distribution (with Markovian) and departures are general distribution. Firstly, we will calculate the system performance measures of the queuing system such as waiting time in the queue, traffic intensity, waiting cost and idle time of the server and then we will analyze the results.

If $\lambda > \mu$, then waiting line will be formed. If $\lambda \leq \mu$, then there will be no waiting line formed.

3.1.1 Assumptions of the Model

1. Customers are served on FCFS basis.
2. There is no balking or reneging i.e.; customers are not leaving the system whatever may be the length of the queue.
3. Arrivals are independent and average number of arrivals (arrival rate) is independent of time.
4. Arrivals may be infinite or very large population and it is described by poisson probability distribution.
5. Service time is arbitrary but it is known.

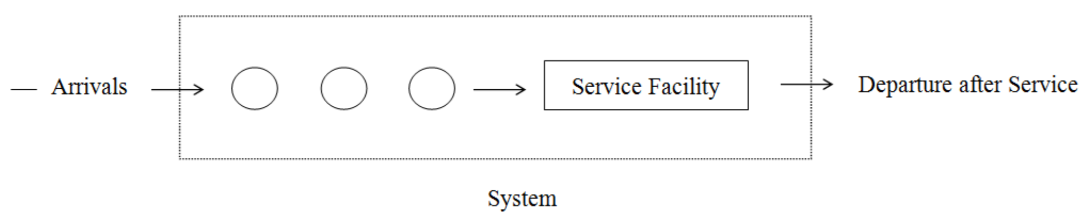

Fig. A. Single Channel Waiting line

3.1.2 Performance Measures of M/G/I Queuing Model

λ = the average number of arrivals per unit of time i.e; per hour
μ = the average number of customers served per unit of time

(Note that units of λ and μ must be same)

1. The utilization factor i.e.; the probability that the service facility is being used

$$\rho = \frac{\lambda}{\mu}$$

2. The average number of customers in the queue

$$L_q = \frac{\lambda^2 \sigma_S^2 + \rho^2}{2(1-\rho)}$$

3. The average time a customer spends in the queue

$$W_q = \frac{L_q}{\lambda}$$

4. The average time a customer spends in the system

$$W_s = W_q + \frac{1}{\mu}$$

5. The average number of customers in the system

$$L_s = \lambda W_s$$

6. The proportion of time the server is idle

$$P_0 = 1 - \rho$$

3.2 Multi-Server M/G/S Model

In multi-server model, there can be two or more than two service channels or servers to provide the service to the customers. Customers are waiting in a single line and proceed to the first available server. All these servers are independent of each other, with arrival rate λ and service rate μ. Firstly, we will calculate the system performance measures of the queuing system such as waiting time in the queue, traffic intensity, waiting cost and idle time of the server and then we will analyze the results.

3.2.1 Assumptions of the Model

The investigation of M/G/S model also assumes arrivals are poisson distribution and departures are general distribution. Other assumptions are similar as for single server M/G/1 model.

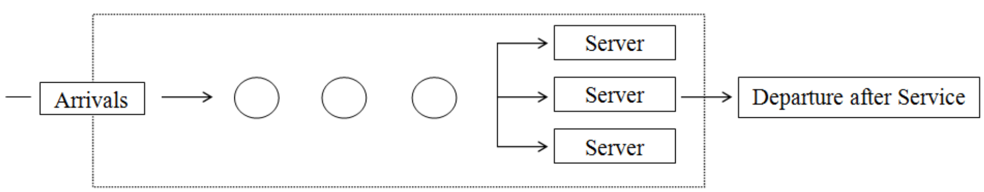

Fig.B. Multiple-Channel Waiting line

3.2.2 Performance Measures of M/G/I Queuing Model

Let S= Number of servers. Then,

1. The traffic intensity is given by

$$\rho = \frac{\lambda}{s\mu}$$

2. The average number of customers in the queue

$$L_q = \frac{\lambda^2\sigma_S^2 + (\frac{\lambda}{s\mu})^2}{2(1-\frac{\lambda}{s\mu})}$$

3. The average time a customer spends in the queue

$$W_q = \frac{L_q}{\lambda}$$

4. The proportion of time the server is idle

$$P_0 = 1 - \rho$$

4. CALCULATION OF VARIOUS PERFORMANCE MEASURES ANALYTICAL METHOD

Waiting time in the queue is the total time spent in the queue to get service. This is the major problem in the queuing theory and it should be reduced so that customer gets satisfied by getting quick services.

Service cost and waiting cost are the two basic costs in the queuing theory. Service cost is the cost of providing service while waiting cost is the cost due to delay in the services to the customers which results customer get irritate and may switch to the competing organizations. Thus it costs very high due to loss of business. Service cost includes the cost for equipment, materials, labour, salaries paid to employees etc. Waiting cost allows to decide the optimal number of servers to minimize the total cost including the service cost and waiting cost.

Traffic intensity is a measure of the average occupancy of a server or resource during a specified period of time, normally a busy hour. It is measured in traffic units (erlangs) and defined as the ratio of the time during which a facility is cumulatively occupied to the time this facility is available for occupancy.

The idle time or proportion of time the server is idle is the unproductive time on the part of employees or machines and during that time the employee is being paid. It could also be associated with computing, and in that case, refers to processing time.

The data used in this paper for analysis is based on observation of any real world situation.

Consider the following situation, the inter- arrival time is exponentially distributed with a mean of 10 minutes and the service time has the uniform distribution with a maximum of 9 minutes and a minimum of 7 minutes. If any customer has to wait for more than 1 hour then company has to pay Rs.10 per hour to the customer. The operating charge of one server is Rs.100 per hour.

5. DETERMINATION OF VARIOUS PARAMETERS

Here,

1. Arrival Rate λ= 1/10
2. Mean Service time= $\frac{7+9}{2} = 8$ i.e; $\mu = 1/8$
3. Variance of the service time $\sigma_s^2 = \frac{(9-7)^2}{12} = \frac{1}{3}$
4. Utilization factor for the system $\rho = \frac{\lambda}{\mu} = \frac{8}{10}$
5. Traffic intensity for the system $\Psi = \frac{\lambda}{S\mu}$
6. The average waiting time $Wt = w_q \times 100$
7. Waiting cost $Wc = Wt \times 10$
8. Operating cost Op. Ct= $S \times 100$
9. Total Cost= Wt Cost+ Op. Ct

Table I: Calculations for Traffic Intensity(in %) and Idle Time(in %)

S.No.	No. of servers	Traffic Intensity (In percentage)	Proportion of time server is idle
1.	1	80%	20%
2.	2	40%	60%
3.	3	26.6%	73.4%
4.	4	20%	80%

Table II: Calculations for waiting time in the queue

S.No.	No. of servers	L_q	$w_q = \frac{L_q}{\lambda}$ (in hours)
1.	1	193/120	16.08
2.	2	196/1440	1.36
3.	3	67/1320	0.50
4.	4	19/7680	0.024

Table III: Calculations for total cost

S.No.	No.of Servers (S)	Wt.	Wt.Ct	Op.Ct.	TC=Wt.Ct+Op.Ct
1	1	19300/12	193000/12	100	16183
2	2	19600/144	196000/144	200	1561
3	3	6700/132	67000/13	300	807
4	4	1900/768	19000/768	400	424

6. GRAPHICAL REPRESENTATION USING MATLAB

Based On various calculations and analysis, the graphical representation of these performance measures are shown in the following figures:

Fig 1. Traffic Intensity (in %) versus No. of Servers

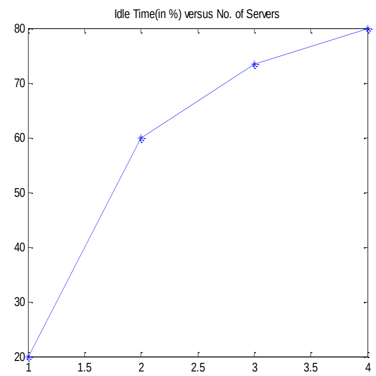

Fig 2 .Idle time (in %) versus No. of servers

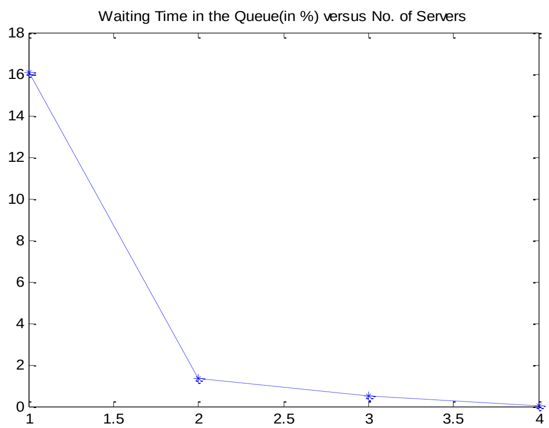

Fig. 3. Waiting Time in the Queue versus No. of Servers

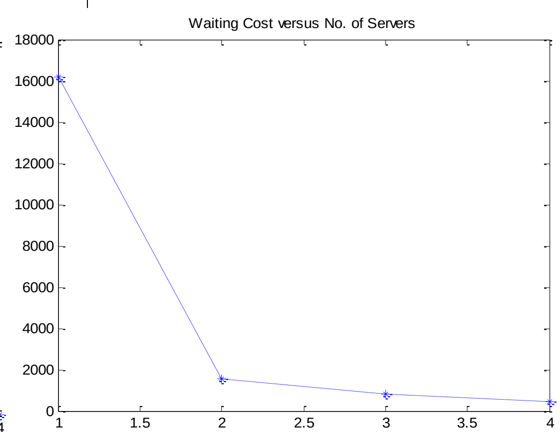

Fig. 4. Waiting Cost versus No. of servers

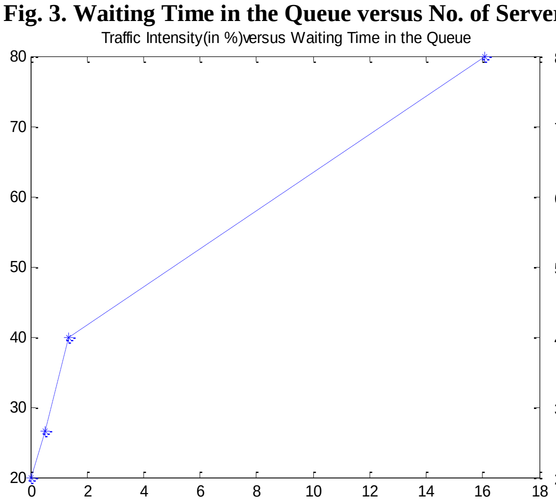

Fig. 5. Traffic Intensity versus Waiting time in the Queue

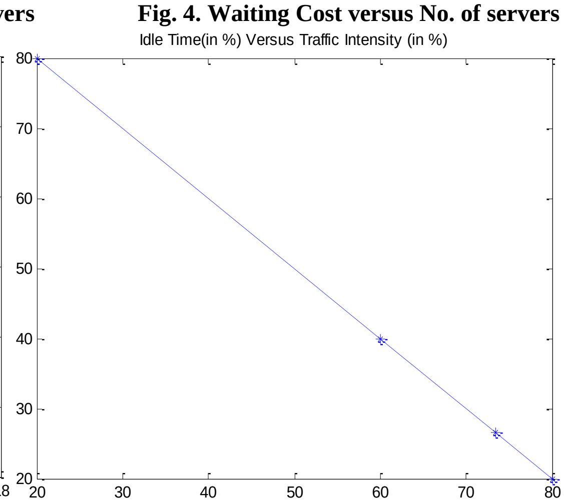

Fig. 6. Idle Time (in %) versus Traffic Intensity (in %)

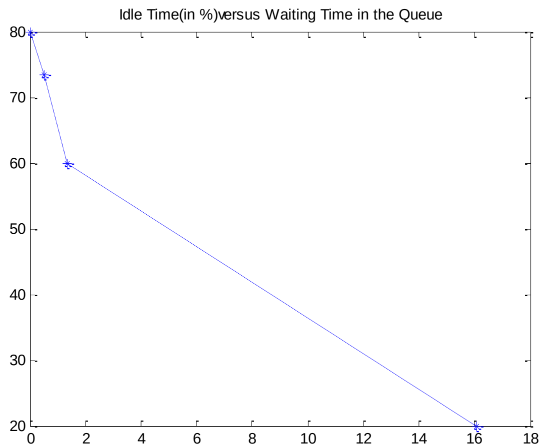

Fig.7. Idle Time (in %) versus Waiting Time in the Queue

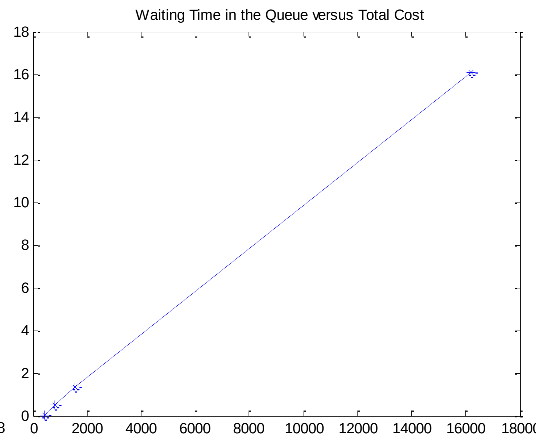

Fig.8. Waiting Time in the Queue versus Total Cost

7. CONCLUSION

From the above example, we have analyzed the comparative study of the various performance measures of M/G/1 and M/G/S queuing models. . This paper concludes the basic logics of queuing theory, how mathematical analysis can be used to calculate system performance measures. Further, we found that as we increase the number of servers, the waiting time in the queue, total cost of the system and traffic intensity decreases whereas proportion of time the server is idle increases. So, this analysis suggest that as we are increasing the number of servers, the proper decision required regarding trade – off between the cost of providing good service and cost of customers waiting time in optimal manner. The graphical representation of various performance measures and their comparison is also discussed. This model can also be used by decision and other policy makers to solve other multi-server queuing problems.

REFERENCES

1. Fredrick S. Hillier, Gerald J. Lieberman, Bodhibrata Nag, Preetam Basu, Introduction to Operations Research, McGraw Hill Education, Ninth Edition, (2012).
2. Hamdy A.Taha, Operations Research: An Introduction, Pearson Publication, Ninth Edition, (2011).
3. Davis, M. M., Aquilano, J. N., Chase, B. R., Fundamentals of Operations Management. Boston, McGraw- Hill Irwin, Fourth Edition, (2003).
4. J.W. Cohen, The Single Server Queue, North-Holland, Amsterdam, (1969). 6. M.L. Chaudhry and J.G.C. Templeton, The queueing system M/GB/1 and its ramifications, *European J. Oper. Res*. 6, 56-60, (1981).
5. K.C. Madan, An M/G/1 queue with second optional service, *Queueing Systems* 34, 37-46, (2000).

6. J. Cao and K. Cheng, Analysis of M/G/1 queueing system with repairable service station, *Acta Math. Appl. Sinica* 5, 113-127, (1982).
7. Madhu Jain , G.C. Sharma , Richa Sharma , Unreliable server M/G/1 queue with multi-optional services and multi-optional vacations, *International Journal of Mathematics in Operational Research*, Volume 5, Issue 2 page 25-36, (2013).
8. A. K. Erlang's, The Theory of Probabilities and Telephone Conversations (Nyt tidsskrift for Mathematic, B, 20 , p. 33, (1909).
9. J. Medhi: Stochastic Processes; Second Edition, New Age International(p) LTD, (2001).
10. Mohammad Shyfur Rahman Chowdhury, "Queuing Theory Model used to solve the waiting line of a Bank- A Study on Islami Bank Bangladesh Limited, Chawkbazar Branch, Chittagong" *Asian Journal of Social Sciences & Humanities* vol.2 No.3 August (2013).
11. Morse, P., Methods of operations research. London: Chapan and Hall (1971).
12. Hui Liand Tao Yang, "Theory and methodology queues with a variable number of servers" *European Journal of operational Research*, Vol.124, pp.615-628, (2000).
13. Charan Jeet singh, Madhu Jain, Binay Kumar, Analysis of M/G/1 Queuing model with state dependent arrival and vacation, *Journal of Industrial Engineering International*, Springer, 19-28 (2012).
14. Mohammad Shyfur Rahman Chowdhury, " Queuing Theory Model used to solve the waiting line of a Bank- A Study on Islami Bank Bangladesh Limited, Chawkbazar Branch, Chittagong"s *Asian Journal of Social Sciences & Humanities* vol.2 No.3 August (2013).