

M/M/S Queuing Theory Model to Solve Waiting Line and to Minimize Estimated Total Cost

Dr. S. K. Tiwari¹, Dr. V. K. Gupta², Tabi Nandan Joshi³

¹Reader, School of Studies in Mathematics, Vikram University, Ujjain, India

²Professor & Head, Dept. of Mathematics Govt. Madhav Science Colleges, Vikram University, Ujjain, India

³School of Studies in Mathematics, Vikram University, Ujjain, India

Abstract: The model describe in this paper is the multiple-server queuing model with Poisson Arrival and Exponential Service Times (M/M/S). We find the total minimum expected cost. Total expected costs are studied, total costs is the sum of the cost of providing service plus the cost of waiting time. Queuing is the common activity of customers or people to avail the desired service, which could be processed or distributed one at a time. Waiting lines and service systems are important parts of the business world. In this paper we describe several common queuing situations and present mathematical models for analyzing waiting lines following certain assumptions. Those assumptions are, arrivals come from an infinite or very large population, arrivals are Poisson distributed, arrivals are treated on a FIFO basis and do not balk or renege, service times follow the negative exponential distribution or are constant, and the average service rate is faster than the average arrival rate.

Keywords: M/M/1; M/M/s; FIFO; Poisson distribution; Service; Queue; Service cost; Utilization factor; Waiting cost; Waiting time, Operating Cost.

1. Introduction

Queuing theory is the study of queue or waiting lines. Some of the analyses that can be derived using queuing theory include the expected waiting time in the queue, the average time in the system, the expected queue length, the expected number of customers served at one time. The study of waiting lines, called queuing theory [3], is one of the oldest and most widely used quantitative analysis techniques and extended to business applications of waiting lines. Waiting lines or queues are pervasive. Venture of all type, industries, colleges, schools, hospitals, banks, post offices all have queuing problems. Some examples of queues less apparent are: a traffic signal to change, waiting for a telephone call handle by operator to operate.

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. A. K. Erlang directed his first effort at finding the delay for one operator and later on the results were extended to find the delay for several operators. However, it was only after World War II that this early work was extended to other general problems involving queues or waiting lines.

2. General Features of Queuing System

Three basic components of a queuing process are arrivals, the actual waiting line and service facilities [9].

Parts of queuing system are as follows

- (1) the arrival (the entry of the customer into the system),
- (2) the queue (the waiting area)
- (3) the service (the customer get the service facility).

These three components have certain features that must be examined before mathematical queuing models can be developed.

Arrival Features

The input source that generates arrivals or customers for the service system has three major characteristics. It is important to consider the size of the calling population, the pattern of arrivals at the queuing system, and the behavior of the arrivals.

Population Size

Population sizes are considered to be either infinite or finite, for practical purpose, in our example the finite number of customers arriving at the counter for service. Most queuing models assume such an infinite calling population.

Arrival Pattern

Arrivals are considered random when they are independent of one another and their occurrence cannot be predicted exactly. Customers either arrive at a service facility according to some known schedule customers or else they arrive randomly [1]. The number of arrivals per unit of time can be estimated by a probability distribution known as the Poisson distribution. Poisson distribution can be established by using the formula:

$$P(n, t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \text{ for } n=0,1,2,3,\dots$$

Where

$P(n;t)$ = probability of n arrivals

λ = average arrival rate

n = number of arrivals per unit of time

Waiting Line Features

Queue

The waiting line itself is the main component of a queuing system. The length of a line can be either finite or infinite. A queue is finite when we made some restrictions on its length.

Analytic queuing models are treated in this article under an assumption of non finite queue length.

Queue Discipline

Most systems use a queue discipline known as the first in, first out rule (FIFO). This is obviously not appropriate in all service system, especially those dealing with emergencies, like services in hospitals, service in disaster management.

Features of Service Facility

It has two basic Features:

(1) The configuration of the service system

Service systems are usually classified in terms of their number of channels, or number of servers, and number of phases, or number of service stops, that must be made. The term FIFS (first in, first served) is often used in place of FIFO [4].

- A single-channel system, with one server, is typified by the drive at bus stop that has only one open counter.
- A single-phase system is one in which the customer receives service from only one station and then exits the system.
- Multiphase implies two or more stops before leaving the system.

(2) Service time pattern

More often, service times are randomly distributed in many cases it can be assumed that random service times are described by the negative exponential probability distribution. This is a mathematically convenient assumption if arrival rates are Poisson distributed.

3. Managerial Applications of Queuing Theory

Queuing theory is very effective tool for business decision-making process. It can be applied to a wide variety of situations for scheduling. Some of these are as follows-

- 1) Aircrafts at landing and take-off from busy airports
- 2) Jobs in production control
- 3) Mechanical transport fleet
- 4) Routing sales persons
- 5) Inventory analysis and control
- 6) Scare defense equipment
- 7) Issue and return of tools from tool cribs in plants
- 8) Minimization of congestion due to traffic delays at booths.
- 9) Parts and components in assembly lines

Advantages

Following are Benefits of Queuing system.

- 1) Queuing models are helpful in creating balance between the two opportunity costs for optimization of waiting costs and service costs.
- 2) Queuing theory provides better understanding of waiting lines so as to develop adequate service with tolerable waiting.
- 3) Queuing theory provides models that are capable of determining arrival pattern of customers or most appropriate number of service stations.

Limitations of Queuing Theory

Major Limitations of Queuing Theory

- 1) The assumption of „First come first served“ is not always true and make (and this happens in several situations) queuing analysis becomes more complex
- 2) In addition to the above complications, queue discipline may also impose certain limitations. The queuing models are very complex and cannot be easily understood. The element of uncertainty is there in almost all queuing situations. Reasons for uncertainty are as follows
 - a) We may not know the form of theoretical probability distribution which applies.
 - b) We might not know the parameters of the process even when the particular distribution is known.
 - c) We would simply be known only the probability distribution of out-comes and not the distribution of actual outcomes even when (i) and (ii) are known.
- 3) In multi-channel queuing system, the departure from one queue often forms the arrival of another. This makes the analysis more difficult.
- 4) In many cases, the observed distributions of service times and time between arrivals cannot be fitted in the mathematical distributions of usually assumed in queuing models. For example, the Poisson distribution which is generally supposed to apply may not fit many business situations.

Mathematical Models (M/M/1):

Single-Channel Queuing Model with Poisson Arrivals and Exponential service times
(M/M/1)[3][9]:

Here present an analytical approach to determine important measures of performance in a typical service system. After these numerical measures have been computed, it will be possible to add in cost data and begin to make decisions that balance desirable service levels with waiting line service costs.

4. Assumptions of the Model

The single-channel, single-phase model considered here is one of the most widely used and simplest queuing models. It involves following conditions [9].

- 1) Arrivals are served on a FIFO basis.
- 2) Every arrival waits to be served regardless of the queue length of the line; i.e., there is no balking or reneging.
- 3) Arrivals are independent and the average number of arrivals (the arrival rate) does not change over time.
- 4) Arrivals are described by a Poisson probability distribution and come from an infinite or very large population.
- 5) Service time also varies from one passenger to the next and is independent of one another, but their average rate is known.
- 6) Both the number of items in queue at anytime and the waiting line experienced by a particular item are random variables.
- 7) Service times occur according to the negative exponential probability distribution. The average service rate is greater than the average arrival rate.

Above all conditions help us to develop a set of equation for development of queue system.

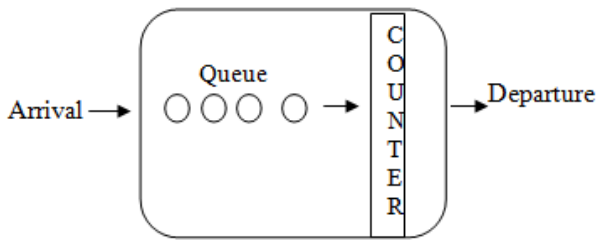


Figure 1: Single-Channel Waiting Line Queuing Equations

λ = mean number of arrivals per time period (for example, per hour)
 μ = mean number of customers served per time period.
 When determining the arrival rate (λ) and the services rate (μ), the same time period must be used. If the λ is the average number of arrivals per hour, then unit of μ must be in hour.

The Queuing Equations Follow

1. The average number of customers or units in the system, L_s , that is, the number in line plus the number being served :

$$L_s = \frac{\rho}{1-\rho} \text{ or } L_s = \frac{\lambda}{\mu-\lambda}$$

2. The average time a customer spends in the system, W_s , that is, the time spent in line plus the time spent being served:

$$W_s = \frac{1}{\mu-\lambda}$$

3. The average number of customers in the queue, L_q :

$$L_q = \frac{\lambda^2}{\mu(\mu-\lambda)}$$

4. The average time a customer spends waiting in the queue, W_q :

$$W_q = \frac{\lambda}{\mu(\mu-\lambda)}$$

5. The utilization factor for the system ρ , that is, the probability that the service facility is being used :

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

6. The present idle time, P_o , that is, the probability that no one is in the system:

$$P_o = 1 - \frac{\lambda}{\mu}$$

Multiple-Channel Queuing Model with Poisson Arrivals and Exponential service Times (M/M/S)[9]

The multiple channel queuing system, in which two or more servers or channels are available to handle arriving customer. Let us still assume that customer waiting service from one single line and then proceed to the first available server. Each of these channels has an independent and identical exponential service time distribution with mean $1/\mu$. The arrival process is Poisson with rate λ . Arrivals will join a single queue and enter the first available service channel [7]. The multiple-channel system presented here again assumes that arrivals follow a Poisson probability

distribution and that service times are distributed exponentially. Service is first come, first served, and all servers are assumed to perform at the same rate. Other assumptions listed earlier for the single-channel model apply as well.

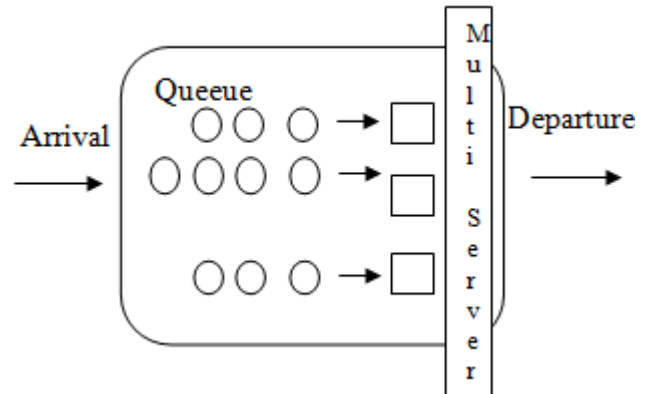


Figure 2: Multiple-Channel Waiting Line

Equations for the Multi-channel queuing Model

If we suppose

S = number of channels open,

λ = average arrival rate, and

μ = average service rate at each channel.

The following formulas may be used in the waiting line analysis[2]:

1. Utilisation rate: $\rho = \frac{\lambda}{s\mu}$

2. The average time a customer spend in the queue waiting for service

$$W_q = \frac{\lambda}{\mu(\mu-\lambda)}$$

3. The average waiting time
 $W_t = W_q \times 100$

4. Waiting Cost
 $W_c = W_t \times 10$

5. Operating Cost
 $Op.Ct = S \times 50$

6. Total Cost = $W_t.Ct + Op.Ct$

Waiting Line Cost

There are two basic types of costs associated with waiting-line problems [10]. First, there are the fairly „tangible“ costs involve in operating each service facility like the costs for equipment, materials, labor, etc. these cost of course , rise as the number of service facilities put into operation increase. On the other hand, there are the relatively „intangible“ costs associated with causing customers to have to wait in line for some period of time prior to being waiting upon- physical discomfort, adverse emotional reactions, reduced or lost sales and so on. Of course, as the number of service facilities in operation increases, the time the customer has to wait in line, on the average, decreases, and hence so too do these costs.

Let us consider an example of a system, where 100 customers are arriving on an average per hour for a particular service and operator service rate is 110 customers per hour. If any customer has to wait for more than an hour then company has to pay Rs. 10 per hour to the customer. The operating charge of one counter is Rs. 50/- per hour. We have to estimate the expected total cost of the system and

advice how many counter company will open to minimize the total expected cost.

Application of the queuing system configurations and discussion of waiting cost and total expected cost of system to serve the customers [6]. The Analysis is summarized in Table - A below;

Table A

S.No	No. of Counter Open (S)	Wq/hour	Wt	Wt.Ct	Op.Ct	TC = Wt.Ct + Op.Ct
1	1	1/11	100/11	1000/11	50	140.90
2	2	1/264	100/264	1000/264	100	103.78
3	3	1/759	100/759	1000/759	150	151.3175
4	4	1/1496	100/1496	1000/1496	200	200.668
5	5	1/2475	100/2475	1000/2475	250	250.40

The Estimated Characteristic of the System:

- 1) Average number of customers arriving at the system, $\lambda = 100$ customer/ hour
- 2) The mean number of customers served per hour, $\mu = 110$ customer/ hour
- 3) The utilization factor for the system, $\rho = \lambda/\mu = 0.9090$
- 4) The traffic intensity for the system, $\psi = \lambda/S\mu = 0.4545 < 1$
- 5) The average number of customers in line waiting for service (check in) $Lq = 9.09$
- 6) The average number of customers in the system $Ls = 10$
- 7) The average time a customer spends in the queue waiting for service $Wq = 0.0037$
- 8) The average time a customers spends in the waiting line or being serviced (namely, in the
- 9) System): $Ws = 0.10$

Expected Total Cost

The Objective Function

Min Exp (Total Cost) = Wt.Ct + Op.Ct

Where

Wt.Ct : Total Waiting Cost;

Op.Ct : Total Operating Cost depend on the number of open counters.

5. Conclusion

For the above example we see that if number of counter increase then cost of system changes, but the minimum expected cost is found if $S = 2$ i.e. two counters are to be operative. This paper has introduced the basic concepts of queuing models, and shown how linear programming, and in some cases a mathematical analysis, can be used to estimate the performance measures of system. The main operating features for a system are shown to be

- 1) Utilization rate,
- 2) Percent idle time,
- 3) Average time spent waiting in the system and in the queue,
- 4) Average number of customers in the system and in the queue, and
- 5) Probabilities of various numbers of customers in the system. The article presents especially the total minimum expectation cost of the system. Unlike the discrete and continuous probability distribution used in the analysis of queuing models.

References

- [1] D.M. Brann and B.C. Kulick, "Simulation of restaurant operations using the Restaurant Modeling Studio," Proceedings of the 2002 Winter Simulation Conference, IEEE Press, Dec. 2002, pp. 1448-1453.
- [2] Hui, L, & Tao, Y. (2000). Theory and methodology queues with a variable number of servers. European journal of operational research, 124, 615-628.
- [3] Kapoor, V. K. (1999). Operations Research. Fourth Edition, July.
- [4] Lee, A. M. (1966). Applied queuing theory. New York: St Marints Press,
- [5] M. Laguna and J. Marklund, Business Process Modeling, Simulation and Design. ISBN 0-13-091519-X. Pearson Prentice Hall, 2005.
- [6] Mohammad Shyfur Rahman Chowdhury Asian Journal Of Social Sciences & Humanities vol. 2 No. 3 August 2013"Queuing Theory Model Used To Solve The Waiting Line Of A Bank -A Study On Islami Bank Bangladesh Limited, Chawkbazar Branch, Chittagong"
- [7] Morse, P. (1971). Methods of operations research. London : Chapan and Hall.
- [8] T. Ahtiok and B. Melamed, Simulation Modeling and Analysis with Arena. ISBN 0-12-370523-Academic Press, 2007. Dec. 2005. Dec. 2008, doi: 10.1109/WSC.2008.4736323.
- [9] Taha, H. (1987). Operations research. Fourth Edition, Collier Mac Willam
- [10] Wagner, H. M. (1975). Principles of Operations Research with Applications to Managerial Decisions (International Edition). Prentice Hall.
- [11] Winston, W. L. (1991).Operations research: applications and algorithms (2nd Edition). Boston: Pws-Kent Publishing Company.