Application of Queueing Model to an Outpatient Flow in a Multi-Speciality Hospital

G. Sreelekshmi[1], Gowry S Kumar[2], P.V. Ushakumari[3*]

[1,2,3*]Department of Mathematics, Amrita School of Arts & Science, Amrita Vishwa Vidyapeetham, Amritapuri Campus, Clappana P.O. Kollam, Kerala, India.

[1]lekshmi1698@gmail.com, [2]gowrysobha@gmail.com, [3*]usha@am.amrita.edu

Abstract

Waiting in line is now becoming a common problem in our daily life and long waiting in queues results in loss of customers, bad reputation etc. Hence queuing and waiting time analysis is particularly important in service systems. Management of queues plays a critical role in the health care management as it deals with the issue of treatment of patients to provide better assistance. This research aims to study the outpatient department of a multi-specialty hospital in Kerala using a single line multiple server queue. For the study, we have collected the data on arrivals, waiting time, service pattern and departures during 7 days (considering 3 servers) from 9.00 am to 3.00 pm and used them as input data in the M/M/C model. The computation of performance measures corresponding to M/M/1, M/M/2 and M/M/3 queuing models shows the average waiting time of patients in the queue and system, the average length of queue and system, the percentage of server utilization, etc. can be reduced by incrementing the number of servers. Also, the plot of delay probability with the number of servers and arrival rate shows that delay probability decreases if the number of servers increases and increases with the increase of arrival rate. The analysis of the study of hospital data using queuing models shows that instead of having three doctors at present there is a need for a minimum of seven doctors to reduce the delay time of patients and thus get the facility immediately.

Keywords: Queues in Hospital, Markovian Queues, Erlang's Formula, Waiting Time, Delay Probability.

DOI: 10.47750/pnr.2022.13.S03.054

INTRODUCTION

Queuing is a common phenomenon in our daily lives. For example, a line of people waiting for a movie ticket in a theatre, a line of customers in a bank, a line of patients waiting for service in hospitals and so on. A queue usually arises when a service's demand exceeds its capacity and queuing theory is a mathematical study that builds mathematical models of a variety of queuing systems, introduced by the Dutch Mathematician Agner Krarup Erlang in the early 20th century. Queuing theory plays a vital role in healthcare settings as the healthcare industry is plagued by delays, and overcrowding is a common feature that may affect the quality and access to health care. Healthcare management is in high demand these days because it is very useful in running a healthcare facility. Now a days, running a hospital management system is becoming extremely useful because of the increasing demand for such facilities. An ordinary healthcare centre consists of one or more counters where the patients are entertained. Due to the unavailability of enough facilities, patients' experiences delays and this results in the formation of queues and make a lot of patient discomfort. This result in bad medical conditions and that can increase subsequent treatment costs and poor health outcomes. Hence it is very important to analyse the service facilities to find a suitable solution to such problems.

The study of all varied dynamics of lines or queues, as well as, how they might be altered to run more efficiently, is the focus of queuing theory. Due to the applicability of queuing models, several researchers have studied them analytically and applied the theory in several situations. Kembe et.al. (2012) analysed the queuing behaviour at the Riverside specialist clinic of the Federal Medical Centre and found the waiting and service cost to determine the optimal number of servers. Saima Mustafa and S. un Nisa (2015) collected data from three departments (registration, pharmacy and outpatient department) of a public hospital and studied performance measures of the system using single and multiple server queuing models. They found that out of the three departments the waiting time of patients in the queue was found to be higher in the pharmacy department and also noticed that patients waiting time could be reduced by using the multiple server queuing model. Nilesh Sheth and Prashant Makwana (2016) proposed a survey on the application of queuing at the outpatient department (OPD) at Arogyam hospital, Rajkot. They have collected both primary and secondary data of thirty days from the OPD and determined the minimum number of doctors required so that the patients do not have to wait in line and can receive the medical treatment immediately**.** P.Umarani and S. Shanmugasundaram (2016) worked on the application of multi-server queues using simulation technique and analytical method and concludes that the results obtained from both analytical and simulation methods are nearly identical. Their study aims to assist the hospital administration in making decisions that will improve the satisfaction of arriving patients. Nsudeet et.al. (2017), have studied a multiple-line multiple channel system of First Bank Nig. Plc, Afikpo branch with multiple servers to analyse and predict the future behaviour of restaurant networks. In 2019, Thiyaagarajan and Mohan Kumar examined queuing system in a multi-speciality hospital via IoT such that the IoT tool calculates the number of patients waiting in the queue. Vijeta Iyer and S. Aruna Devi (2019) underwent queuing analysis in a restaurant using simulation method with various service distributions and also by analytic method. They showed that faster serving of food to the customers will improve the customer's satisfaction and thus helps the management to increase the profit. Md Obaidul Haque et.al. (2020) worked on multiple servers queuing model and conducted a case study to reduce outpatients waiting time at a public hospital without considering the cost. This research was carried out to ensure that the waiting time targets did not fall below the hospital's minimal service standard, regardless of cost. Ushakumari and Devi Krishna (2020) have studied analytically optimal service time management in a single server Markovian queue. They have analysed a finite capacity Markovian queue under a control policy for the server called the T-policy. Akhil et.al. (2021) applied different queuing models with single and multiple servers to a railway ticket window and obtained a suitable model for the data collected. The purpose of this paper is to highlight the benefits of queuing theory to hospital management systems by using the multiple server queuing models. Data for this study include the number of patients arriving at the front desk, arrival time, service time, departure time, etc., for 7 days was collected from the outpatient department of a multi-speciality hospital in Kerala. The objectives of this study are:

(i) Describe the use of multi-channel queues, in support of hospital managers to make their decisions.

(ii) Perform a case study of multi-channel queues in the context of multi-specialty hospital management.

The rest of the paper is organized as follows: - Mathematical frame work is given in section 2 and section 3 gives materials and methods. In section 4 provides results, discussion and a conclusion.

Mathematical Framework

Description of the queuing theory: A queue/waiting line is the flow of arrival of customers to a service system for service facilities, waiting in the line if service is not immediate, selection of customers by some rule called service discipline and departure of customers from the service system after service.

Kendall's notation to represent a queuing system:

A /B / C / D /E, its terms are as follows:

A: inter-arrival time or the arrival process, B: type of service time, C: the number of servers in the system

D: the system capacity, E: the queue discipline or order of service

Queuing discipline indicates the order that determines the service rendering to the customer. The standard discipline is FIFO (first in first out).

Mathematical Model and the System Characteristics: The mathematical models considered in our study are single server and multi-server queueing models. The arrival rate of patients follows a Poisson process with parameter λ and the service times follow an exponential distribution with a rate of patients per unit time as μ. The queue discipline is first in first out and the queue capacity is infinite. We use the following notations.

Steady-state queue probabilities: Let X(t) denote the number of patients in the system including those in service at time t and $P_n(t) = P(X(t) = n / X(0) = 0), n = 0, 1, 2, \ldots$. Also let P_n denote the steady-state probability of the number of patients in the system. That is, $P_n = \lim_{t\to\infty} P_n(t)$. The main performance measures of M/M/1 and M/M/C queuing systems are given in Table1.

Little's Theorem: Little's law describes the relationship between arrival rate, the time spent in the system and the average number of customers in the system which can be expressed algebraically as $L = \lambda W$.

Delay probability (P_D): The delay probability for the C server system can be calculated by using the equation

$P_D = 1 - \sum_{n=0}^{C-1}(P_n)$, where P_n is the steady-state probability of n customers in the system. The above equation is also known as Erlang's Loss Formula.

Expected Delay (k): It is the average delay experienced by a customer waiting in the queue and is given by the equation, $k = \frac{W_q}{P_D}$, where W_q is the average time, a customer spends in the queue.

Table 1

Performances Measures	M/M/1	M/M/C
Probability that service facility is idle (P_0)	$1-\frac{\lambda}{\mu}$	$\left[\sum_{n=0}^{C-1}\frac{\left(\frac{\lambda}{\mu}\right)^n}{n!}+\frac{1}{C!}\left(\frac{\lambda}{\mu}\right)^C\left(\frac{C\mu}{C\mu-\lambda}\right)\right]^{-1}$
Average number of customers in the queue (L_q)	$\frac{\lambda^2}{\mu(\mu-\lambda)}$	$\left[\left(\frac{1}{(C-1)!}\right)\left(\frac{\lambda}{\mu}\right)^C\frac{\lambda}{(C\mu-\lambda)^2}\right]P_0$
Average number of customers in the system (L_s)	$\frac{\lambda}{(\mu-\lambda)}$	$L_q+\frac{\lambda}{\mu}$
Average time a customer spends in the queue (W_q)	$\frac{\lambda}{\mu(\mu-\lambda)}$	$\frac{L_q}{\lambda}$
Average time a customer spends in the system (W_s)	$\frac{1}{(\mu-\lambda)}$	$W_q+\frac{1}{\mu}$
Traffic Intensity(ρ)	$\frac{\lambda}{\mu}$	$\frac{\lambda}{C\mu}$
Probability of having n customers in the system (P_n)	$P_0\,\rho^n$ [where $P_0=(1-\rho)$]	$\begin{cases}P_0\frac{(C\rho)^n}{n!}, (n\le C)\\ P_0\frac{(C\rho)^n}{C!}, (n\ge C)\end{cases}$

Materials and Methods

For the study, we have visited the outpatient department of VSM hospital on 7 days (1/12/2021-8/12/2021) from 9.00 am to 3.00 pm and conducted a survey about the arrival of patients to the reception counter and about the service rendered to the patients. The data collection was primary and the methods included direct observation and administrative questionnaires. In this hospital, patients can fix an appointment with the doctor via direct registration or booking in advance. Patients arriving at the outpatient department have to initially register at the registration counter. The entire registration process is computerized and after registering the patient will be provided an ID card, which will allow them to communicate with the hospital in the future. This card contains patient medical record numbers, which is their most important identification while in the hospital. Also, after registering, the patient will be given a token number and they will be called to the doctor's room based on the number they were given. As a result, patients had to wait in line until their number was called. They have both full-time and part-time consultants and the regular consultation times are between 9:00 a.m. to 3:30 p.m.

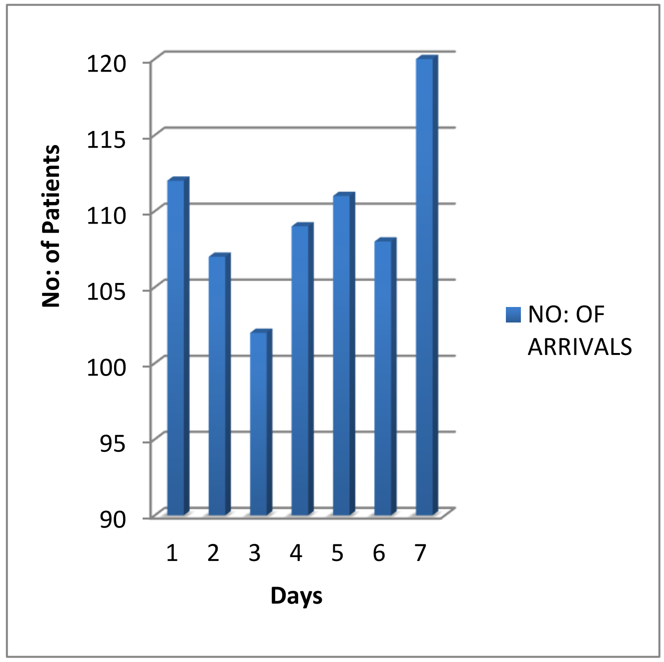

Fig: 1: Arrival data for seven days

The queue structure we have observed in the outpatient department is a single line multiple server queuing process and is labelled as M/M/C system as shown in figure 2.

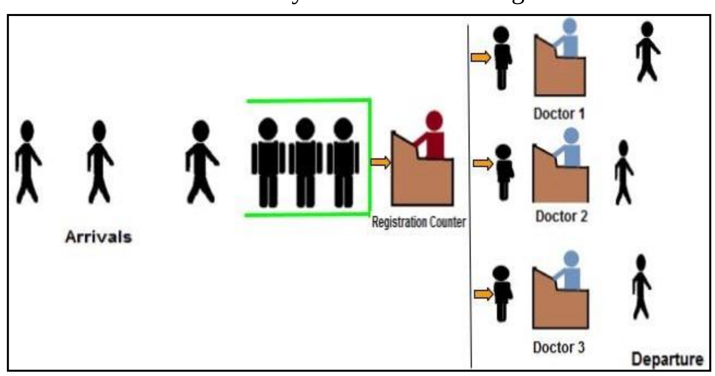

Fig: 2: Pictorial representation of hospital queue system

The following table shows sample data regarding the arrivals and departures of patients collected from the outpatient department of the hospital:

Table 2

Sl. No:	Arrival time	Inter arrival time (min)	Service begins	Service ends	Service time(min)
1	9.48	-	10.00	10.06	6
2	10.00	12	10.06	10.12	6
3	10.05	5	10.12	10.17	5
4	10.08	3	10.18	10.23	5
5	10.15	7	10.24	10.30	6
6	10.17	2	10.30	10.35	5
7	10.25	8	10.36	10.41	5
8	10.30	5	10.42	10.47	5
9	10.35	5	10.48	10.53	5
10	10.38	3	10.54	11.00	6
11	10.43	5	11.00	11.05	5
12	10.55	12	11.06	11.11	5
13	10.58	3	11.12	11.17	5
14	11.05	7	11.18	11.24	6
15	11.14	9	11.24	11.30	6
16	11.23	9	11.30	11.36	6
17	11.27	4	11.36	11.40	4
18	11.33	6	11.42	11.47	5
19	11.40	7	11.48	11.54	6
20	11.45	5	11.54	12.00	6
21	11.51	6	12.00	12.06	6
22	11.55	4	12.06	12.11	5
23	12.01	6	12.12	12.17	5
24	12.06	5	12.18	12.23	5
25	12.15	9	12.24	12.30	6
26	12.25	10	12.30	12.36	6
27	12.26	1	12.36	12.41	5
28	12.30	4	12.42	12.47	5
29	12.35	5	12.48	12.53	5
30	12.45	10	12.54	1.00	6
31	12.50	5	1.00	1.05	5
32	12.52	2	1.06	1.12	6
33	12.58	6	1.12	1.17	5
34	1.10	12	1.18	1.23	5
35	1.15	5	1.24	1.30	6
36	1.20	5	1.30	1.36	6
37	1.29	9	1.36	1.40	4
		221			**199**

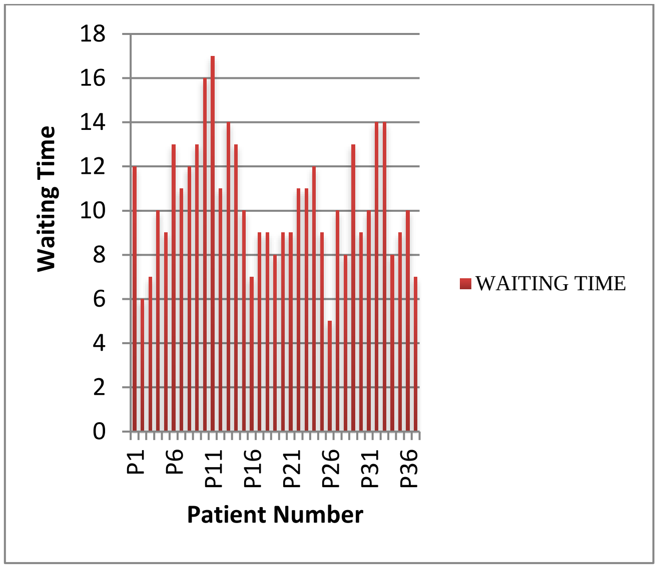

Fig: 3: Patient's waiting time (minutes) in the queue

Results and Discussions

IV(a) Analysis of data using M/M/3 queuing model

The type of queuing system that we take under consideration is the multiple server queuing system with three doctors. The assumptions made for this queuing system are:

a) It consists of a single queue with three identical servers.
b) The input process is Poisson arrival rate of λ patients per minute.
c) Service times follow an exponential distribution with the mean service rate of μ patients per minute.
d) Patients receive services from each server on a first-in, first-out basis.
e) The queue size is unlimited.

Table 3 shows the average arrival and service rates of three doctors for 7 days.

Table 3: Average arrivals and Service time of patients for seven days among 3 doctors

Days	Lambda (λ)	Doctor 1 (μ_1)	Doctor 2 (μ_2)	Doctor 3 (μ_3)
1	0.1495	0.16624	0.1594	0.1517
2	0.1515	0.1607	0.1484	0.1539
3	0.1531	0.1520	0.1859	0.1504
4	0.1570	0.1613	0.1843	0.1597
5	0.1703	0.1815	0.1835	0.1842
6	0.1658	0.1763	0.1799	0.1638
7	0.1622	0.1707	0.17	0.1628
Total	1.1094	1.1649	1.2114	1.1265
Average	0.1584	0.1664	0.1731	0.1609

IV(b) Queuing Characteristics of Single and Multiple server queues

The following table shows the estimated system characteristics of the M/M/1, M/M/2 and M/M/3 queues using the data

Table 4: Comparison of different queues

Performance measures	M/M/1	M/M/2	M/M/3
Arrival rate (λ)	0.1584	0.1584	0.1584
Service rate(μ)	0.1664	0.1697	0.1668
P_0	0.0481	0.3636	0.3832
L_q	18.8481	0.2599	0.0371
L_s	19.8	1.1933	0.9867
W_q	118.9904	1.6409	0.2342
W_s	125	7.5337	6.2294
ρ	0.9519	0.4667	0.3165

Table 4 shows that if we increase the number of doctors to 3 then the average time spent by the patients in the queue will decrease from 118.9904 min to 0.2342 min and that in the system will decrease from 125 min to 6.2294 min. Also, the probability that the system remains free for M/M/3 was found to be 38.32% of the time, which is higher than that of M/M/1 and M/M/2, but the ratio of arrival rate to the service rate decreases when the number of doctors is increased. When there are three doctors available, it can be found that the waiting time is considerably reduced.

IV(c) Computing delay probability, expected delay and utilization factor corresponding to the number of servers

Table 5: Delay probability, Expected delay and Server utilization factor corresponding to the various number of servers

No: of Servers (c)	Delay probability $P_D = 1 - \sum_{n=0}^{C-1}(P_n)$	Expected Delay $k = \frac{W_q}{P_D}$	Utilization Factor $\rho = \frac{\lambda}{C\mu}$
1	0.9496	119.0527	0.9496
2	0.3058	5.7067	0.4748
3	0.0801	2.9213	0.3165
4	0.01716	1.9639	0.2374
5	0.0032	1.40625	0.1899
6	0.0005121	1.1716	0.1583
7	0.00011835	0.84495	0.1357
8	0.000064935	0	0.1187
9	0.000058947	0	0.1055
10	0.000057925	0	0.09496

From Table 5, it is clear that delays in providing services to patients when all servers are busy can be reduced by increasing the number of doctors. Also, the expected delay experienced by the patients waiting in the queue decreases up to 7 servers and then remains constant in the scenario of the hospital. The higher the number of servers, the lower will be the utilization factor. The following graph shows the relation of delay probability, expected delay and utilization factor with the number of servers.

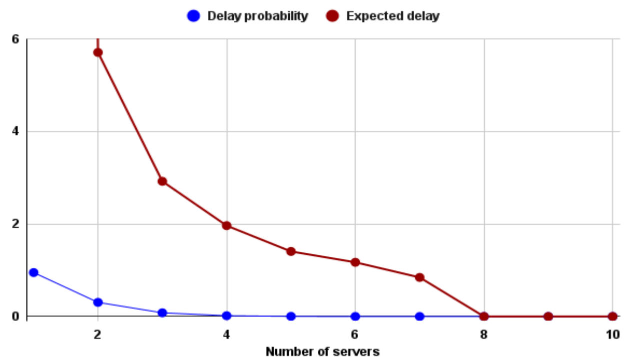

Fig 4: Variation of Delay probability and Expected delay with number of servers

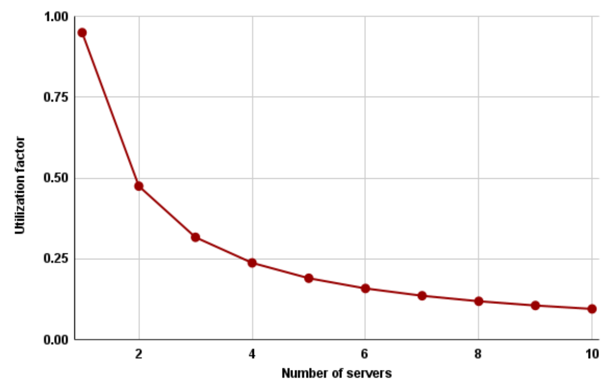

Fig: 5: Utilization factor v/s Number of server

IV(d) Computing delay probability corresponding to the arrival rate of each day

Table 6: Delay probability corresponds to an increase in arrival rates

Days	Lambda (λ)	Delay probability $P_D = 1 - \sum_{n=0}^{C-1}(P_n)$
1	0.1495	0.0695
2	0.1515	0.0739
3	0.1531	0.0734
4	0.1570	0.0782
5	0.1622	0.0849
6	0.1658	0.0896
7	0.1703	0.0957

The above table shows that if the number of patients arriving per minute increases, then the probability of delay in getting service also increases.

Conclusion

In the present study, we have described the application of queuing theory in the health care system by surveying a multispecialty hospital in Kerala, using single and multiple server queuing models. We have performed a comparative study of M/M/1 and M/M/C queues using the data collected from a multi-specialty hospital in Kerala. The outcome of the research shows that the waiting time of the patients could be reduced when we increase the number of doctors to seven instead of three. On comparing single server and multiple server queuing models, the results show that when the number of doctors is increased, the average number of patients waiting in the system and the queue and the average time the patients spend in the system and the queue decreases. Further, using graphical study, it is noticed that the utilization factor, the probability of delay in providing services to patients when all servers are busy, and the average delay experienced by patients waiting in the queue decrease as the number of servers increases, and the rate of arrival of patients is directly proportional to the probability of delay. Based on the study, for customer satisfaction, the management should increase the number of doctors up to seven to take care of the patients.

References

Akhil M Nair, Sreelatha. K.S, Ushakumari P.V. Application of Queuing Theory to a Railway ticket window, International Conference on Innovative Particles in Technology and Management (ICIPTM), 2021, 154-158.

Kembe, M.M, Onah, E.S, lorkegh, S. A Study of Waiting and Service Costs of a Multi-Server Queuing Model in A Specialist Hospital, International Journals on Scientific & Technology Research, 2012, September 1(8), 19-23.

Md Obaidul Haque, Sharmin Akter, Anwar Hossen, Zillur Rahman, A Case Study on Outpatient Waiting Time for Treatment with Multiple Servers Queuing Model at Public Eye Hospital in Bangladesh, Quest Journals, Journal of Research in Applied Mathematics, 2020, 6(1), 04-09.

Nilesh Sheth, Dr. Prashant Makwana, Application of queuing model on hospital OPD, International Journal for Scientific Research and Development, 2016, 3(11), 124-126.

F.I. Nsude, Elem-Uche O and Bassey Uwabunkonye Analysis of Multiple-Queue Multiple-Server Queuing System: A Case Study of First Bank Nig.plc, Afikpo Branch, International Journal of Scientific & Engineering Research, 2017, 8(1), 1700-1709.

Saima Mustafa, S. un Nisa, A Comparison of Single Server and Multiple Server Queuing Model in Different Departments of Hospitals, Journal of Mathematics, 2015, 47(1), 01-08.

K. Thiyagarajan, Dr.K. Mohan Kumar (2019), Optimized Server Utilisation in Multi-Speciality Hospital By Using The Queuing Model Through IoT, International Journals on Scientific & Technology Research, 2019, 8(10), 1038-1042.

P. Umarani and S. Shanmugasundaram, A study on M/M/C Queuing Model under Monte Carlo Simulation in a Hospital, International Journal of Pure and Applied Mathematical Sciences, 2016, 9(2), 109-121.

P.V. Ushakumari and S. Devi Krishna, Optimal Service Time Management in a Single server Queue, International Journal of Advanced Research in Engineering and Technology, 2020, 11(6), 50-58.

Vijeta Iyer, S. Aruna Devi, M/M/C Queue under Monte Carlo Simulation for a Restaurant Model, International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019, 8(10), 1818-1823.

Ibrahim, S., & Koksal, M. E. (2021). Realization of a fourth-order linear time-varying differential system with nonzero initial conditions by cascaded two second-order commutative pairs. Circuits, Systems, and Signal Processing, 40(6), 3107-3123.