Application of Queuing Theory to Patient Satisfaction at Private HospitalsDuring the Second Wave of Covid 19

Ms ShivaliShrivastava

Sanskriti University, Mathura
shivali.mini11@gmail.com

Dr SaloniSrivastava

Raja Balwant Singh EnggTechnical Campus, Bichpuri, Agra

Salonisaloni17@gmail.com

Dr VipinSolanki

Sanskriti University, Mathura

solankivipin11@gmail.com

ABSTRACT

The second wave of COVID-19 in India has had severe consequences in the form of spiraling cases, reduced supplies of essential treatments, and increased deaths particularly in the young population. Understanding why the second wave has been more dangerous than the first could help to identify the potential areas of diagnostics to target with future control strategies. Studies have identified various circulating double-mutant and triple-mutant strains of SARS-CoV-2 across different regions of India, which are more pathogenic than the initial strains. Such altered transmissibility and pathogenicity indicates evolution of the virus. Here we study with the help of multi-server finite length queue model at what situation they are not able to get immediate treatment and have to face their deaths and also how rapidly the problem increases. We consider all the private hospitals of Agra as multiple servers and then study the whole scenario and realized that how the condition comes at its worst state.

Keywords: Queuing Theory, Multiple Servers, Finite Queue, Waiting time, Hospital operations and Patient satisfaction.

INTRODUCTION

During the first wave of the COVID-19 pandemic it emerged that the nature and magnitude of demand for mental health services was changing. Considerable increases were expected to follow initial lulls as treatment was sought for new and existing conditions following relaxation of 'lockdown' measures. For this to be managed by the various services that constitute a mental health system, it would be necessary to complement such projections with assessments of capacity, in order to understand the propagation of demand and the value of any consequent mitigations. This paper provides an account of exploratory modeling undertaken within a major Agra healthcare system during the Second wave of Covid – 19 is increasing very rapidly, when actionable insights were in short supply and decisions were made under much uncertainty. We analyze that soon the condition of outcry comes in our city Agra we are not taking this disease seriously. For this follow the following research for our consideration.**C. Lakshmi & S. A. Iyer(2013)** reviewed the contributions and applications of queuing theory in the field of health care management problems and proposed a system of classification of health care areas which are examined with the assistance of queuing models. The categories described in the literature are expanded and a detailed taxonomy for subgroups is formulated. The goal is to provide sufficient information to analysts who are interested in using queuing theory to model a health care process and who want to locate the details of relevant. **A. Sharifet. al. (2014)** considered the accumulating priority queue (APQ), a priority queue where customer priorities are a function of their waiting time. This time-dependent priority model

was first proposed by Kleinrock (1964), and, more recently, Stanford et al. (2013) derived the waiting time distributions for the various priority classes when the queue has a single server. In their work they derived expressions for the waiting time distributions for a multi-server APQ with Poisson arrivals for each class, and a common exponential service time distribution and also comments on how to choose feasible accumulation rates to satisfy specified performance objectives for each class. **S. R. Chakravarthy& M. F. Neuts (2014)** considered a multi-server queuing model in which two types of arrivals occur. Regular customers arrive according to a Markovian arrival process and special customers arrive according to a phase type renewal process. The service times are exponentially distributed with parameter depending on the type served. While a regular customer requires the attention of only one server, a special customer needs the attention of all servers. An arriving special customer finding the servers busy with a special customer is lost; otherwise, the special customer will enter into service immediately by possibly pre-empting existing services of regular customers. The displaced customers will enter into service once the special customer leaves. **A. PaulaIannoni et al (2015)** proposed a cutoff hypercube queuing model to analyze server-to-customer emergency services operating with server reservation. They were motivated by certain SAMU's (*Systèmed'AideMédicaleUrgente*) that give assistance to different classes of emergency requests, including specialized transfer of patients, and use the reservation strategy to improve the probability that ambulances will be available to high priority calls.
N. L. David & A. Stanford (2016) developed a multi-class multi-server queuing model with heterogeneous servers under the accumulating priority queuing discipline, where customers accumulate priority credits as a linear function of their waiting time in the queue, at rates which are distinct to the class to which they belong. At a service completion instant, the customer with the greatest accumulated priority commences service. When the system has more than one idle server, the so-called *r*-dispatch policy is implemented to determine which of the idle servers is to be selected to serve a newly-arriving customer. They established the waiting time distribution for each class of customers. They also presented a conservation law for the mean waiting time in $M/M_i/c$ systems, and study a cost function in relation to the conservation law to optimize the level of heterogeneity among the service times in $M/M_i/2$ systems. **B. Maddahet. al. (2017)** developed an analytical model to determine the optimal system configuration, in terms of the number of stations, the corresponding thresholds, and the number of servers at each station, for a given number of servers facing Poisson demand. In order to simplify the computation of the thresholds, they assumed that the load is balanced among different stations. **I. Geraint et. al. (2018)** explored deadlock in queuing networks with limited queuing capacity, presents a method of detecting deadlock in discrete event simulations, and builds Markov chain models of these deadlocking networks. The three networks for which Markov models are given include single and multi-server networks for one and two node systems. The expected times to deadlock of these models are compared to results obtained using a simulation of the stochastic process, together with the developed deadlock detection method. **R. Hassin& R. R. Green (2018)** presented a model of parallel queues in front of two servers that provide the same service. Upon arrival, each customer inspects the queue in front of one server before either joining it or inspecting the other. If both queues were inspected, the customer joins the inspected queue that minimizes sojourn time. The solution of this model is not straightforward, even when the system contains only two servers, and the equilibrium is not always a threshold strategy. They showed that, in many cases, a unique equilibrium strategy that contains cascades exists: customers choose one action (join or inspect) when they observe *i* and i+2 customers in the first observed queue, and the other action when they observe i+1 customers in the first observed queue. They find cascade equilibrium strategies even when the servers are identical with respect to service rate or inspection cost. They also show that compared to the case where all customers first inspect the same queue, symmetric inspection reduces system load but not necessarily customers' expected cost. **H. Q. Ali& S. Ghani (2020)** proposed a $M_K/Hyper_K/1/M$ queuing model for multi-sensor nodes. Queues with Poisson arrivals and hyper-exponentially distributed service time and their applications have been widely discussed in literature however hardly any work can be found that uses them for modeling WSNs. **B. J. Murchet. al. (2021)** represented that a multi-node, multi-server queuing network with reneging, is implemented in open-source software and is freely and publicly available.

RESEARCH DESIGN

The Multiple Server Model (or usually known as M/M/s server discipline) occurs in the setting of a waiting line in which there is one or more servers, the customers are supposed to arrive at a random rate that is specified as a Poisson distribution for a given time period (or the inter-arrival times are exponentially distributed), and the service times are exponentially distributed. The main parameters of a waiting line are:

- Probability of no units in the system : $P_0 = \frac{1}{\sum_{n=0}^{s-1}\frac{1}{n!}\left(\frac{\lambda}{\mu}\right)+\frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s\frac{s\mu}{s\mu-\lambda}}$
- Average Number of Units in the System : $L_s = \frac{\lambda\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2}P_0 + \left(\frac{\lambda}{\mu}\right)$
- Average Number of Units in the Queue : $L_q = L_s - \left(\frac{\lambda}{\mu}\right)$
- Average Time a unit spend in the System : $W_s = \frac{\mu\left(\frac{\lambda}{\mu}\right)^s}{(s-1)!\,(s\mu-\lambda)^2}P_0 + \left(\frac{1}{\mu}\right)$
- Average Time a unit spend in the Queue : $W_q = W_s - \frac{1}{\mu}$
- Utilization Factor : $\rho = \frac{\lambda}{\mu}$

CALCULATION

We apply this model in our city Agra. As we know that in Agra there were around 40 private hospitals. We consider all the private hospitals as parallel servers provided to the patients who are suffering from Second wave of Covid – 19

Let us consider the problem where there is an attack of Second wave of Covid – 19 in our city Agra. In Agra, we have around 40 Private hospitals. Now we take 40 parallel servers (= no. of private hospitals in Agra), all are having finite Queue capacity. Now let us assume three cases one by one:

Case (i) When the situation is under Control: The arrival pattern of Second wave of Covid – 19 patient in Agra is Poisson with 80 patients in every one hour and the server time is exponential with 10 patients in every 5 minutes. Then we get:

Average number of units in the system L_s	Average time a unit spends in the system (minutes) W_s	Average number of units in the queue (minutes) L_q	Average time a unit spends in the queue (minutes) W_q	Utilization factor ρ	Probability of zero units in the system P_0
2	1.33	1.25	0.83	0.75	0.4724

Table 1: Multiple Server Model by providing the arrival rate per time period $\lambda = 1.5$, the service rate per time period $\mu = 2$, and the number of servers $s = 40$

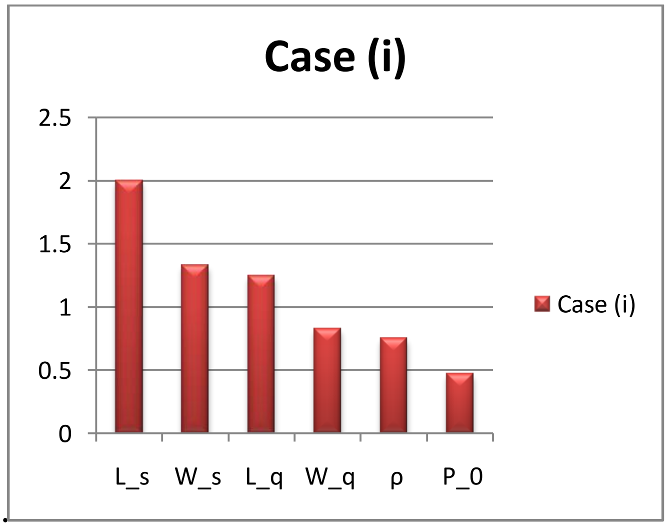

Figure 1: Multiple Server Modelby providing the arrival rate per time period $\lambda = 1.5$, the service rate pertime period $\mu = 2$, and the number of servers $s = 40$

Case (ii) When the situation is not under Control but can be justify: The arrival pattern of Second wave of Covid – 19 patient in Agra is Poisson with double rate to the previous case and also added number of patients left in the queue in the previous case which becomes arrival of patients = 180 + 2 +1 = 183 patients in every one hour and the server time is exponential with 10 patients in every 5 minutes. Then we get

Average number of units in the system L_s	Average time a unit spends in the system (minutes) W_s	Average number of units in the queue (minutes) L_q	Average time a unit spends in the queue (minutes) W_q	Utilization factor ρ	Probability of zero units in the system P_0
8	2.62	6.48	2.12	1.53	0.22

Table 2: Multiple Server Model by providing the arrival rate per time period $\lambda = 1.5$, the service rate per time period $\mu = 3.05$, and the number of servers $s = 40$

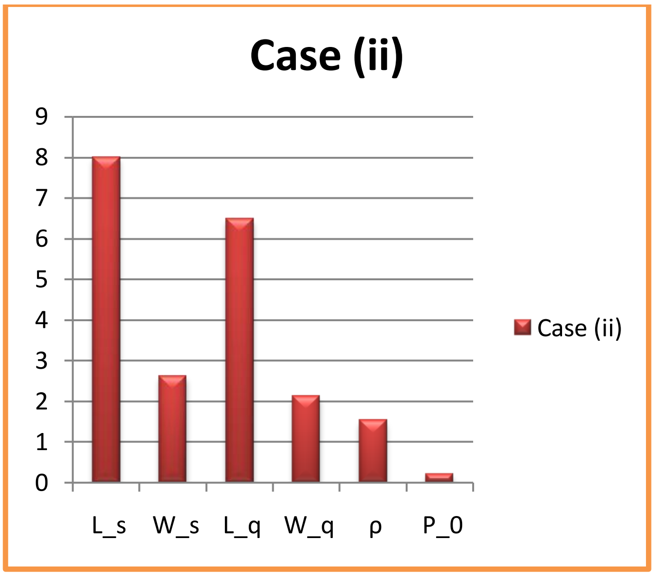

Figure 2: Multiple Server Model by providing the arrival rate per time period $\lambda = 1.5$, the service rate pertime period $\mu = 3.05$, and the number of servers $s = 40$

Case (iii) When the situation is not under Control : The arrival pattern of Second wave of Covid – 19 patient in Agra is Poisson with three times the rate to the case (i) and also added number of patients left in the queue in the case (ii) which becomes arrival of patients = 270 + 8 + 6 = 284 patients in every one hour and the server time is exponential with 10 patients in every 5 minutes. Then we get

Average number of units in the system L_s	Average time a unit spends in the system (minutes) W_s	Average number of units in the queue (minutes) L_q	Average time a unit spends in the queue (minutes) W_q	Utilization factor ρ	Probability of zero units in the system P_0
20	4.22	17.63	3.73	2.37	0.09

Table 3: Multiple Server Model by providing the arrival rate per time period $\lambda = 1.5$, the service rate per time period $\mu = 4.73$, and the number of servers $s = 40$

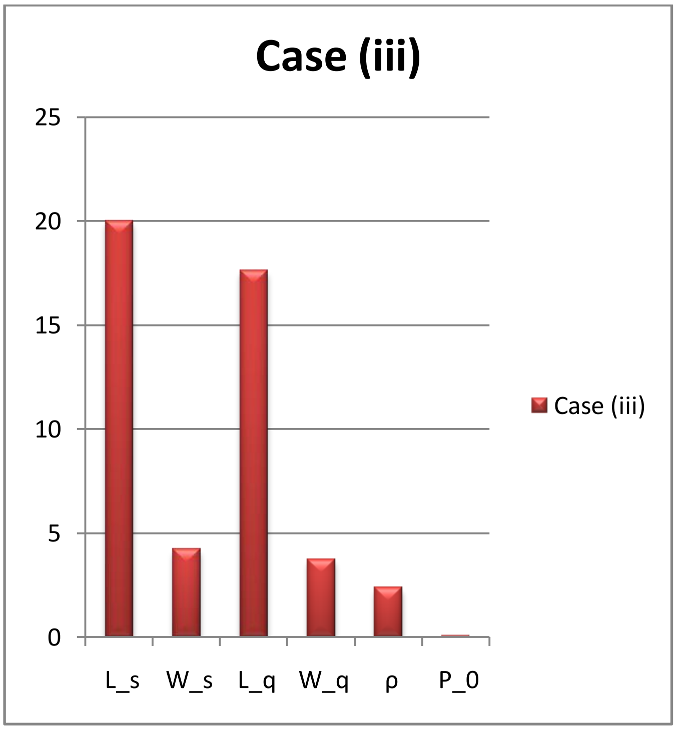

Figure 3: Multiple Server Model Calculator, by providing the arrival rate per time period $\lambda = 1.5$, the service rate pertime period $\mu = 4.73$, and the number of servers $s = 40$

Case (iv): In the case (ii) we can see that the patients left without treatment are 38, which is a big number if it is attached to this death stage. Now in the last fourth case we assume that the patients of Second wave of Covid – 19 are now four times of the case (i). Then the arrival of patients = 360 + 20 + 18 = 284 patient in every one hour and the server time is exponential with 10 patients in every 5 minutes and hence we get:

Average number of units in the system L_s	Average time a unit spends in the system (minutes) W_s	Average number of units in the queue (minutes) L_q	Average time a unit spends in the queue (minutes) W_q	Utilization factor ρ	Probability of zero units in the system P_0
40	6.03	36.69	5.53	3.32	0.04

Table 4: Multiple Server Model by providing the arrival rate per time period $\lambda = 1.5$, the service rate per time period $\mu = 6.63$, and the number of servers $s = 40$

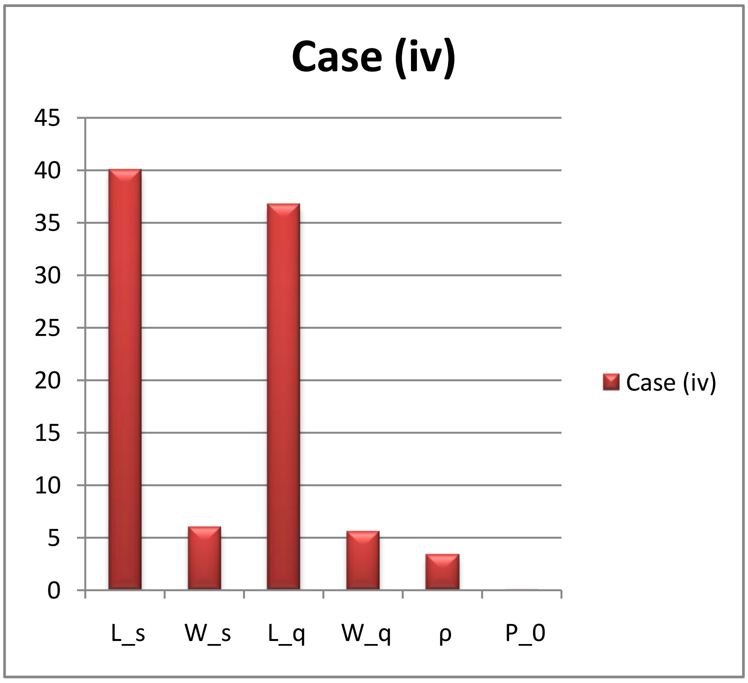

Figure 4: Multiple Server Model by providing the arrival rate per time period $\lambda = 1.5$, the service rate pertime period $\mu = 6.63$, and the number of servers $s = 40$

Now as we can see this is the outburst situation which create outcry in our city Agra.
In the below figure we analysis all the four cases together

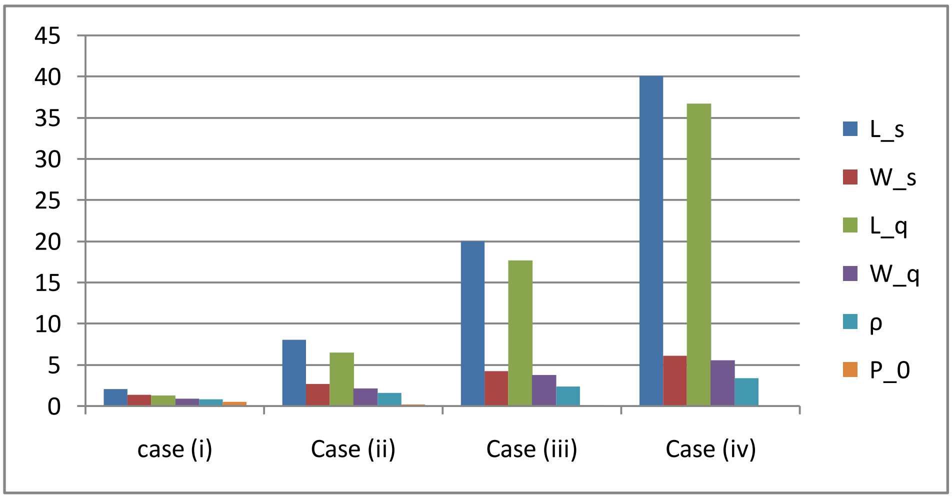

Figure 5: Summarize form of all the four cases

RESULTS

The health care industry worldwide is plagued with delays. We are used to waiting for an appointment to see a physician or schedule a procedure, and when we arrive at the physician's office, we wait again to be seen. In hospitals, we often find patients waiting for beds in hallways, and delays for surgery or diagnostic tests are common.There are numerous consequences to waiting — delayed care, poor patient satisfaction, financial implications and more.Queuing theory deals with delays caused by the disparity between demand for a service and the capacity to meet the demand. The purpose of a queuing model is to balance customer service (shorter wait time) and resource limitation (number of servers).

In case (i) we see that the condition is in our control but when the number of patients of Second wave of Covid – 19 comes daily in the ratio when our private hospital of Agra can treat them in a proper way. In case (ii) we see that what happened when the number of patient of Second wave of Covid – 19 comes in twice of the case (i). In this case we can see from the table that utilization factor value comes out to be more than 1 which indicates that demand > max demand, it is the situation when the beds are not allotted to the patients they have to do their treatment in the corridor or where else they get the space, although doctors can control this situation. In case (iii), we can see that the patient left for the treatment is very large and creates and outcry situation in the city also ulilization factor is higher than the 1 value. In case (iv) we have outburst situation which creates an curse to the whole city. We also noticed that after case (i) the value of utilization factor increases more than one which shows that the number of arrivals into system exceeds the number of customer being served. In order words, as time goes on, the number of customer in the system will increase by time. In queue theory, we call this condition as transient-state condition; otherwise it is called as steady-state condition. By looking all the cases we refer that the server to be infinite and for that we need extra beds and doctors on demand which we get from the city where people from the Second wave of Covid – 19 are very few in number then only we can outcome from this situation and this should be done within two or three days as we get many patients (specially children) of Second wave of Covid – 19.

Our main propose of this paper is to get aware from the outcry situation developed the Second wave of Covid – 19 which we already shown with the help of data, table and charts practically in the upper section. Now the scenario has been cleared in front of us. And for this we use queuing model. It is only a theoretical model that statistically can predict the future operational state based on implementing certain physical, behavioral and work process changes. The model provides only direction and confidence to the improvement team; the team members still must do the heavy lifting of implementation and associated change management. Application of queuing model in isolation, without a strong implementation team, does not guarantee project success. Nevertheless, queuing theory promotes team alignment, team focus and will continue to drive us toward our patient experience goal: approaching zero wait time for our patients.

REFERENCES

1. **Ali H. Q. &Ghani S. (2020)**: “Multi-Sensor Based M_K/Hyper$_K$/1/M Queuing Model for Heterogeneous Traffic”, Computer Networks, 181:107512.
2. **Chakravarthy S. R. &Neuts M. F. (2014)**: “Analysis of a multi-server queueing model with MAP arrivals of regular customers and phase type arrivals of special customers”, Simulation Modelling Practice and Theory, 43: 79-95.
3. **David N. L. & Stanford A. (2016)**: “Multi-server accumulating priority queues with heterogeneous servers”, European Journal of Operational Research, 252(3): 866-878.
4. **Geraint I., Paul P. R., Vincent H., Knight A. (2018)**: “Modelling deadlock in open restricted queueing networks”, European Journal of Operational Research, 266(2): 609-621.
5. **Hassin R. &Green R. R. (2018)**: “Cascade equilibrium strategies in a two-server queueing system with inspection cost”, European Journal of Operational Research, 267(3):1014-1026.
6. **Lakshmi C. &Iyer S. A. (2013)**: “Application of queueing theory in health care: A literature review”, Operations Research for Health Care, 2(1-2): 25-39.
7. **Maddah B., Nasr W. W., Charanek A. (2017)**: “A multi-station system for reducing congestion in high-variability queues”, European Journal of Operational Research, 262(2): 602-619.
8. **Murch B. J., Cooper J. A., Hodgett T. J., Gara E. L., Walker J. S., Wood R. M. (2021)**: “Modelling the effect of first-wave COVID-19 on mental health services”, Operations Research for Health Care, 30: 100311.
9. **PaulaIannoni A., Chiyoshi F., Morabito R. (2015): “A spatially distributed queuing model considering dispatching policies with server reservation”, Transportation Research Part E: Logistics and Transportation Review, 75: 49-66.**
10. **Sharif A. B., Stanford D. A., Taylor P., Ziedins I. (2014)**: “A multi-class multi-server accumulating priority queue with application to health care”, Operations Research for Health Care, 3(2):73-79.