GSJ: Volume 11, Issue 10, October 2023, Online: ISSN 2320-9186
www.globalscientificjournal.com

Utilization of Queuing Models to Customer Management in a Banking System

Herbert Nnaemeka okoli *, Darlingtina Adaora
Department of Mathematics and Statistics, Enugu State University of Science and Technology,Agbani
Department of Mathematics and Statistics,Caritas University, Amorji-Nike,Enugu State, Nigeria.
Herbertokoli2016@gmail.com, ada.darlingtina@caritasuni.edu.ng

ABSTRACT

The presence of queues in banks, Specifically on Mondays and Fridays has led to the application of queuing theory in the banking sector. Queuing theory, which deals with the mathematical study of waiting lines, is well- suited for analyzing the queue and waiting line where customer have to wait for service. This research aims to determine the average time customers spend in queues and the actual time taken for service, with a focus on understanding the impact of time wastage and associated costs. To analyze the queuing model we employed the Markovian birth and death process and specifically applied the multiple servers, single queue (M/M/S) queuing model to analyze the data collected by observation from a bank from the results obtained, the arrival rate is 0.1207 and the service rate is 0.156, the probability that the servers are idle is 0.44 which shows that the servers will be 44% idle and 56% busy, the expected number in the waiting line is 0.1361,the expected number in the system is 0.9098. The expected waiting time in the queue is 1.276 and the expected total time lost waiting time in the queue is 1.276 and the expected total time lost waiting in one day is 3.2664hpurs, the average cost per day for waiting is ₦65.328 Naira and from the calculation of the comparing solution, the average cost per day from waiting Is ₦7.966 naira which means that there had been a saving in the expected cost of ₦65.328 – ₦7.966 = ₦57.362. This means that with three servers , the average cost from waiting is reduced. Hence we concluded that the aim and objectives of this paper was achieved.

Keywords: Bank, Queue, Service pattern, Markovian Birth and Death process, Poisson distribution.

INTRODUCTION

Queue is a usual sight in banks mostly on Mondays and Fridays. The word queue comes via French and the latin cauda meaning “tail”. Customers waiting in line to receive services in any service system are inevitable and that is why queue management has been where the manager faces huge challenge.

Queuing theory examines every component of waiting in line ,including the arrival process, service process, number of system places, and the number of customers, which might be people, data packets, cars and Network services. Real life application of queuing theory cover a wide range of most business.

Hence, queuing theory is suitable to be utilized in the banking system. Since it is associated with queue or waiting line where customers who cannot be served immediately have to queue (wait) for service for a long time and time being a resource ought to be managed effectively and efficiently because time is money. Queuing theory has become one of the most important, valuable and arguable one of the most universally used tool by an operational researcher. It has applications in diverse fields including telecommunications, traffic engineering, computing and design of factories, shops, offices, banks , hospitals and project management.

Queuing theory can also be utilized to a variety of operational situations such as people or vehicles awaiting there turn to be attended to or to proceed it is not possible to accurately predictthe arrival rate (or time) of customers and service rate (or time) of service facility of facilities.

Some of the analysis that can be derived using queuing theory include the expected waiting time in the queue, the average waiting time in the system, the expected queue length, the expected number of customers served at one time, the probability of balking customers, as well as the probability of the system to be in certain states, such as empty or full. [Patel, R. et al 2012]

Queuing models are used to represent the various types of queuing systems that arise in practice, the models enable in finding an appropriate balance between the cost of service and the amount of waiting. [A. Nafees, 2007]

Queuing models provide the analyst with a powerful tool for designing and evaluating the performance of queuing systems. [Bank, C. et al, 2001]

A queuing model of a system is an abstract representation whose purpose is to isolate those factors that relate to the
system's ability to meet service demands whose occurrences and durations are random. [J. Sztrik, 2010]

Any system in which arrivals place demands upon a finite capacity resource maybe termed as queuing systems, if the arrival times of these demands are unpredictable, or if the size of these demands is unpredictable, then conflicts for the use of the resource will arise and queues of waiting customers will form and the length of these queue dependon two aspects of the flow pattern: First, they depend on the average rate. Secondly, they depend on the statistical fluctuations of this rate. [Klenrock, L. 1975]

In 1909, the first study of queuing theory was done by a Danish mathematician, A.K. Erlang which resulted into the worldwide acclaimed Erlang telephone model. He examined the telephone network system and tried to determine the effect of fluctuating service demands on calls on utilization of automatic dial equipment.

This study is required to investigate the expected waiting time of customers and the actual waiting time in banks, where the gap between the actual and expected waiting time can be analyzed to know how to improve on the efficiency and effectiveness of their bank. Such problems are:

- How poor service facilities has affected the overall bank performance.
- How poor service pattern affects queue discipline.
- How service facilities has affected the time of customers
- How poor service delivery impacts on time.
- How poor service's delivery has affected customers behaviour.

The aim of this study is to determine the amount of average time customers spend on a queue and actual time of service delivery. Therefore the objectives of this study are as follows:-

- To examine the impact of time wasting on the weak performance.

- To improve on the efficiency and effectiveness of their operations.
- To help bank mangers improve customer's satisfaction through queue management.
- To improve on time management.

This paper when completed will be significant to many people and organisations especially banks in Nigeria. First ofall, it will add to the literature on queuing theory and management which will be accessed by lecturers and scholars.
Most importantly, bank managers will benefit a lot from this study as they will apply this theory in their various banks, thereby reducing the amount of time spent on queues which might lead to customer's satisfaction and improve on their overall efficiency and effectiveness

In terms of the analysis of queuing situations, the types of questions in which we are interested in are typically concerned with measures of system performance which includes:

- To what extent does the service time differ from the actual time that customers have to wait before being served?
- To what extent does poor service pattern affect queue discipline?
- To what extent do service facilities affect customer's service?
- To what extent does the average service time affect the overall performance of the bank?
- How does poor service delivery affect customer's behaviour?

In this paper, we will be studying one bank that adopts the M/M/S queuing model. The queue discipline is first in first out (FIFO) and the arrival is strictly random having Poisson distributed arrival times and exponentially distributed service times.

1.1. Queuing Models and Kendall's Notation

In most cases, queuing models can be characterized by the following factors:

- Arrival time distribution: Inter-arrival times most commonly fall into one of the following distribution patterns a Poisson distribution, a Deterministic distribution, or a General distribution. How-ever, inter-arrival times are most often assumed to be independent and memory less, which is the attributes of a Poisson distribution. However inter-arrival times are most often assumed to be independent and memory less which is the attribute of a Poisson distribution.
- Number of server: The way queuing calculations are performed varies depending on whether there is a single server or multiple servers for the queue. In a single server queue, there is only one server assigned to serve the entire queue.This is often seen in grocery stores,where each cashier has there own dedicated line. On the other hand , in a multiple server queue,there are multiple servers available but customers form a single line and wait for the first available server to assist them. This scenerio occurs in banks ,where there maybe several tellers but a centralized queue system.
- Queue length: The queue in a system can be characterized as either having an infinite or finite queue length.
- System capacity: The maximum number of customers in a system can be from 1 upto infinity. This includes the customers waiting in the queue.
- Queuing Discipline: There are several possibilities in terms of the sequence of customers to be served such as FIFO (first in first out, I.e, in order of arrival) , random order, LIFO (last in first out,I.e the last one to come will be first to be served) or priorities.

Kendall, in 1953, proposed a notation system to represent the six characteristics discussed above. The notation of a queue is written as: A/B/P/Q/R/Z where A, B, P, Q, R and Z describes the queuing system properties.

- A describes the distribution type of the inter arrival.
- B describes the distribution type of the service times.
- P describes the number of servers in the system.
- Q describes the maximum length of the queue.
- R describes the size of the system population.
- Z describes the queuing discipline.

1.2. Introducing Notations and Their Terminology

State of the system = number of customers in queuing system.

λ = arrival rate

μ=service rate.

λ_n =mean arrival rate (expected number of arrivals per unit time) of new customers when n customers are in the system.

μ_n= mean service rate for overall system (expected number of servers completing service per unit time) when n customers are in system.

N(t)= number of customers in queuing system at time t.

$P_n(t)$= probability that exactly n customers are in queuing

system at time t. P_0 (t)= probability that there are no customers

in queuing system in time t. S= number of servers.

M/M/S= inter arrival time and inter departure times are exponentially distributed in a queuing system with S servers.

A birth and death process is one that is appropriate for modelling changes in the size of a population. Thus, most queuing models especially the single ones can be analyzed as birth and death processes. In the context of queuing theory, we define the following terms:-

- Birth refers to the arrival or entrance of a new customer into the queuing system.
- Death refers to the departure of a served customer.
- The state of the system at time t, (t>0) is given by N(t)

Thus, the birth and death process describes probabilistically how N (t) changes as t changes.

Generally speaking, the birth and death process states that individual birth and death occurs randomly, where their mean occurrence rates depends only upon the current state of the system.

1.3. Multiple Servers Queue with M|M|S Model

An M|M|S system is a queuing process having Poisson arrival patterns, S server, with S independent, identically distributed, exponential service times (which does not depend on the state of the system); infinite capacity, and a FIFO queue discipline. The arrival pattern being stated independent, $\lambda_n = \lambda$ for all n. The service times associated with each server are also independent, but since the number of servers that actually attend to customers (i.e. are not idle) does depend on the number of customers in the system, the effective time it takes the system to process customer through the service time facility is state dependent. In particular, if $\frac{1}{\mu}$ is the mean service time for one server to handle one customer, then the mean rate of service completion when there are customers in the system is

$$\lambda_n = \lambda,\ \mu_n = \begin{cases} n\mu & \text{f } n = 0,1,\dots,c. \quad \text{.e.} \quad n \le c \\ c\mu & \text{f } n = c+1, c+2. \quad \text{.e.}\ n \ge c \end{cases} \tag{1.3.1}$$

The probability of zero customers in the system (P_o) and the probability of n customer in the system (P_n) are given by:

$$P_0 = \left\{ \frac{(\lambda/\mu)^c}{C!\left[1 - \frac{\lambda/\mu}{c}\right]} + \frac{(\lambda/\mu)^1}{1!} + \frac{(\lambda/\mu)^2}{2!} + \dots + \frac{(\lambda/\mu)^{c-1}}{(c-1)!} \right\}^{-1} \tag{1.3.2}$$

$$P_n = P_0 \frac{(\lambda/\mu)^n}{n!} \quad \text{if } n \le c \tag{1.3.3}$$

$$P_n = P_0 \frac{(\lambda/\mu)^n}{C!\,C^{n-c}} \quad \text{if } n > c \tag{1.3.4}$$

The capacity utilization in this system is $\lambda/_{c\mu}$

We can use the above equation of $\lambda/_{c\mu} < 1$

If $\frac{\lambda}{c\mu} > 1$, then the waiting line grows larger and larger i.e. becomes infinite if the process runs long enough.

when C = 1 (there is one service facility), equations (1.3.3) and (1.3.4) reduces to

$$P_n = \left(\frac{\lambda}{\mu}\right)^n P_0 \tag{1.3.5}$$

From equations (1.3.3) and (1.3.4), we have

$$P_n = P_0 \frac{(\frac{\lambda}{\mu})^n}{n!} \text{ if } n \le C$$

But n can only take on values of 0 or 1 if $n \le C = 1$. Thus

$$P_n = P_0 \left(\frac{\lambda}{\mu}\right)^n$$

If C = 1, equation 4 also reduces to equation 5

With C service facilities, the average number of customers in the queue is

$$N_q = \frac{(\lambda/\mu)^{c+1} P_0}{C.C![1-\frac{\lambda/\mu}{C}]^2} \qquad (1.3.6)$$

The average number in the system (waiting plus service) is

$$N_s = N_q + \frac{\lambda}{\mu} \qquad (1.3.7)$$

The expected waiting time in the queue for an arrival i

$$T_q = \frac{N_q}{\lambda} \qquad (1.3.8)$$

The expected total time spent in the system (waiting plus service) is

$$T = \frac{N_s}{\lambda} \qquad (1.3.9)$$

The method adopted in this paper is the Birth and Death process which is a special class of Markov process.

RESULT AND DISCUSSION

The methodology can now be illustrated using the data below: if the data collected by observation in a bank shows that the system capacity is 238 customers with 2 servers, inter-arrival time for 238 customers is 1972 minutes and the time taken by 239 customers to be served is 1529 minutes.

We first of all estimate the parameters,N= 238 customers

T= 1972 minutes

S= 1529 minutes

C= 2

Where N= system capacityT= time interval

S= time taken by 239 customers to be servedC= number of servers.

Now the

Arrival rate, $\lambda = \frac{N}{T} = \frac{238}{1972} = 0.1207$

Service rate, $\mu = \frac{N}{\ } = \frac{238}{\ } = 0.156$

S 1529

Traffic intensity (ℓ) = $\frac{\lambda}{\mu} = \frac{0.1207}{0.156} = 0.7737$

This implies that $\mu = 0.1207$

$\lambda = 0.156$

C = 2

$\frac{\lambda}{\mu} = 0.7737$

We now calculate the probability values

Probability of n customers in the system is from equ. (1.3.4)

$$P^n = \frac{(\frac{\lambda}{\mu})^n}{c!\, c^{n-c}} = \frac{(0.7737)^n\ 0.4421}{2! \times 2^{n-2}} = \frac{(0.7737)^n\ 0.4421}{2^{n-1}}$$

Now we calculate the Queue measure

The expected number in the waiting line, from equ (1.3.6) we have that

$$N_q = \frac{(\frac{\lambda}{\mu})^{c+1}}{C\,C!\,[1-\frac{\lambda/\mu}{C}]^2} P_0$$

$$= \frac{0.7737^3 (0.4421)}{2 \times 2! [1 - \frac{0.7737}{2}]^2} = \frac{(0.4631) \times 0.4421}{4 \times 0.3760} = 0.1361$$

The expected number in the system (waiting plus in service) is by equ. (1.3.7)

$$N_s = N_q + \frac{\lambda}{\mu}$$

$$= 0.1361 + 0.7737 = 0.9098$$

The expected waiting time in the queue is by equation (1.3.8)

$$T_q = \frac{N_q}{\lambda} = \frac{0.136\,1}{0.1207} = 1.1276$$

The expected total time lost in one day

$T_L = \lambda \times 24 \text{ hours} \times T_q$ (hint 1 day = 24 hours)

$= 0.1207 \times 24 \times 1.1276 = 3.2664$ hours.

Assuming a cost is associated with this hour (i.e. ₦20 for each hour lost by customer waiting).The average cost per day from waiting is:

= 3.2664 × ₦20 = ₦65.328.

Now, we want to find out if increasing the number of servers can help to reduce the amount of time spent on queue and the hence minimize the cost incurred by waiting. Hence we compare solutions:

Let $\lambda = 0.1207$

$\mu = 0.156$

$\frac{\lambda}{\mu} = 0.7737$ which is same in the data above but let C = 3

Now we have

$$P_o = \frac{1}{\frac{(0.7737)^3}{3! \times [1 - \frac{0.7737}{3}]} + 1 + \frac{(0.7737)^1}{1!} +}$$

GSJ: Volume 11, Issue 10, October 2023
ISSN 2320-9186

(3 2
0 7 !
.)
7 2
7

$= \frac{1}{0.1040 + 1 + 0.7737 + 0.2993} = \frac{1}{2.177} = 0.4593$

The expected number in the waiting line is

$N_q = \frac{(0.7737)^4}{3 \times 3!(1 - 0.7737/3)^2}(0.4593)$

$= \frac{0.1646}{9.9128} = 0.0166$

Expected number in the system is

$N_s = N_q + \frac{\lambda}{\mu}$

= 0.0166 + 0.7737 = 0.7903

Then expected waiting time in the queue is

$T_q = \frac{Nq}{\lambda} = \frac{0.0166}{0.1207}$

The expected total time lost waiting on one day $T_L = \lambda \times 24 \text{ hours} \times T_q$

= 0.1207 × 24 × 0.1375 = 0.3983 hours

The average lost per day from waiting is

= 0.3983 × ₦20 = ₦7.966

Considering the analytical solution, the capacity of the system under study is 238 customers and the arrival rate is 0.1207 while the service rate is 0.156. This shows that the service rate of the system is greater than the arrival rate, this does not necessarily imply that there is no queue but that queue may not be long. Considering the ratio of the two rates that is arrival rate over the service rate called traffic intensity which is less than one. Probability that the servers are idle is 0.44 which shows that the servers will be 44% idle and 56% busy, while the probability that there are n customers in the system is

$$\frac{(0.4421)(0.7737)^n}{2^{n-1}}$$

The expected number in the waiting line is 0.1361. The expected number in the system is 0.9098. The expected waiting time in the queue is 1.276 and the expected total time lost waiting in one day is 3.2664 hours.

From the foregoing Queue measures, the average cost per day for waiting is ₦65.328 and from the calculation of the comparing solutions, the average cost per day from waiting is ₦7.966. There had been a saving in the expected cost of ₦65.328 - ₦7.966 = ₦57.362. This means that with three servers, the average cost from waiting is reduced.

CONCLUSION

We now conclude that adding one more server will help reduce time spent on queue which can improve customer's satisfaction. Hence the objective of this paper is achieved. The advantage of using this single system with multiple servers is that a slow server does not affect the movement of the queue i.e if a server is slow it does not affect the movement of the queue because next customer can go to the next available server instead of waiting for the slow server.

ACKNOWLEDGEMENT

I acknowledge the Almighty God for His steadfast love and grace upon my life throughout my stay in the university.I wish to appreciate my wonderful supervisor Dr Everestus. O. E for his constructive criticisms and his fatherly love. May the good Lord bless you. I want to say a big thank you to the Head of Department of Industrial Mathematics and Applied statistics in the person of Dr. Ikechukwu Godwin Ezugorie for his encouragement. Also a big thank you goes to all the lecturers of our beloved department for impacting the knowledge of mathematics on me; I am very grateful. My special gratitude goes to the authors cited during the research in the persons of Olosore.O, Klenrock, L. Keilson, J and Patel R et al for the knowledge pass across to me through their researches works, I am very grateful.

REFERENCES

[1] Abate ,J. Whitt, W.(1988). Simple spectral representations for the M/M/1 queue".Queueing Systems.

[2] Dr Janos Sztrik (2011). Basic Queuing theory. University of Debrecen, Faculty of informatics.

[3] Kasumu R.B (1994). Introduction to probability theory ;Fatol Ventures Publishers. Lagos Nigeria.

[4] Keilson, J. Kooharian, A.(1960). On Time Dependent Queuing Processes". The Annals of Mathematical Statistics

[5] Stewart, Williams J. (2009). Probability Markov chains, queues and simulation:the mathematical basis of performance modeling. Princeton University Press. ISBN 978-0-691-14062-9.

[6] Asmussen, S.R (2003). Queuing Theory at the Markovian Level". Applied probability and Queues. Stochastic Modeling and Applied Probability

[7] Kasumu R.B (2000); Introduction to Stochastic Process; Fatol Ventures Publishers Lagos Nigeria.

[8] Kleinrock, L (1975): Queuing systems. Vol. I. Theory. John Wiley & Sons, New York.

[9] O. Olasore (1992); Nigeria Banking and Economic Management "Journal of the Nigeria Institute of Bankers" Vol.3 page 8-13. Nigeria.

[10] Patel, J.J, Rajeshkumar, M.C, Pragnesh, A.P, Makwana, P (2012) "Minimise the waiting time of customer and gain more profit in restaurant using queuing model" International conference on management, humanity and economics(ICMHE 2012) p77-80

[11] A.k Sharma, G.K Sharma (2013): Queuing theory approach with queuing model: A study" ISSN: 231966726 volume 2issue 2, February 2013 PP: 01-11

[12] Banks, J. Carson, J.S. Nelson, B.L, Nicol, D.M (2001): Discrete – Event System

GSJ: Volume 11, Issue 10, October 2023
ISSN 2320-9186

Simulation, Prentice Hall international series, 3rd edition, P.24-37.

[13] Kingman, J. F. C .; Atiyah (October 1961). The single server queue in heavy traffic" . Mathematical proceedings of the Cambridge Philosophical Society.

[14] Robertazzi, Thomas G. (2000). Computer Network and Systems. New York, NY: Springer New York. P.72 ISBN 978-1-4612-7029-4.

[15] Morrison, J. A (1985). Response time Distributing for a processor-sharing system" SIAM Journal on Applied Mathematics.

[16] Harrison, P,G (1993). Response time distributions in queuing network models performance evaluation of computer and communication systems. Lecture notes in Computer science. Vol 729. Pp. 147-164.

[17] Adam, Ivo "Course QUE: Queueing theory, Fall 2003: The M/M/1 System

[18] Gross, Donald; Shortle, John F.; Thompson James M.; Harris, Carl M.(2011-09-23) "Busy period analysis" . Fundermental of queueing Theory

[19] Sturgul, John R. (2000). Mine design: examples using simulation. SME. P. Vi. ISBN 0-87335-181-9

[20] Harrison, Peter: Patel , Naresh M. (1922) Performance modeling of communication Networks and Computer Architectures.