ANALYSIS OF MULTIPLE-QUEUE MULTIPLE-SERVER QUEUING SYSTEM: A CASE STUDY OF FIRST BANK NIG. PLC, AFIKPO BRANCH

F. I. Nsude, Elem-Uche O. and Bassey Uwabunkonye

Department of Statistics

Akanu Ibiam Federal Polytechnic, Unwana, Ebonyi State, Nigeria

Abstract: This research paper focused on multiple-lines, multiple server systems of customers in First Bank Nigeria Plc, Afikpo Branch in Ebonyi State. The underlying mathematical concepts of queue models: arrival and service time distributions, queue disciplines and queue behavior were presented. The operating characteristic formulas for multiple-server queuing model meant to evaluate performance of practical queuing systems were also presented. Data on the arrival time of customers and service time spent by customers to receive service were collected and analyzed. Different ratios of average arrival time and service time, $\left(\frac{\lambda}{\mu}\right)$, were obtained to determine the optimal number of service facilities (servers) appropriate for the bank. The utilization factor and the average waiting time of customer in the system stood at 0.1762 and 0.1523 respectively. We suggested a reduction in the number of servers in the bank from four to three to reduce the idle time of the servers and also reduce the operation cost.

Key words: Arrival rate, Service rate, Poison distribution, Exponential distribution, performance measures.

1. 1 Introduction

Queues or waiting lines are lines of individuals or customers waiting for their turn to receive service from service operators. The customers here could be human beings or physical entities waiting to receive service in such places as petrol station, bank, mechanic workshop, airport, car park, goods for shipment, among others. Queue is everyday experience in Nigeria especially in banks and filling stations. According to Sharma, (2013), queue is formed when either the number of customers requiring service exceeds the number of facilities (servers) or the servers/facilities do not perform efficiently well such that it takes more time than necessary to serve a customer.

Customers' waiting lines create worries to managers of service providers because of the resulting consequences of balking and reneging by the customers. Long waiting lines generate stress and aggravate mistakes in addition to cost by both the waiting customers and the facility operators; see Uche and Ugah (2014). In other words, customers lose their precious time and the service providers also lose valuable customers via reneging and balking when a long queue is formed

Generally, increasing the number of servers by the operator would reduce the waiting time of customers, but increase the operating cost of the services. Yuncheng and Liang (2002) noted that the queue problem is a problem about a balance between average waiting time of customers and the idle time of the attendants in a service station. Queuing theory is an effective method used in variety of situations where it is not possible to accurately predict the arrival rate/time of customers and service rate (time) operation servers. It is used to determine the level of service that would produce effective use of time both to the customers and the service operator.

Many authors have done useful works in queue systems in different areas of human endeavors. For instance Ezeliora et al (2014) studied queuing system management of Shoprite Plaza, Enugu, using single-line multiple server analysis. They recommended a decrease in the number of attendants/servers to reduce the operating cost of the system and reduce the idle time of the servers. Adamu (2015) did a work on ECO and First Banks by simulating a single queue- single server and single queue-multi server systems. He noted that First bank should turn multi-tellers multi server system to single queue multi tellers to reduce the overall waiting time of customers from the multiple serving points and also reduce the customers' jockeying problems. For ECO bank, it was noted that the multi-tellers should improve on their turnaround in order to have further reduction of waiting time of customers. Most of the works in literature adopted single line multiple arrival rate of customers with different arrival and service rate of the system. This study focuses on multiple-line multiple channel systems of First Bank Nig. Plc, Afikpo branch with multiple average arrival rates of customers and multiple service rates of counter servers.

1.2 Objectives of the study

(i) To analyze the First Bank queue system with multiple queue, multiple server systems.

(ii) Analyze the different ratio $\left(\frac{\lambda}{\mu}\right)$, arrival rate and service rate of the bank;

(iii) Calculate the economic trade-off (cost) of the customers' waiting time and the servers.

2.1 The M/M/S Model

The model discussed here is M/M/S of infinite calling population with first come, first served multiple server queue system, (∞/FCFS). In this case there is multiple but identical servers in parallel line that provide the same service to the customers. Customer on arrival joins any queue and remains there until he receives service. The queue characteristics are as follow:

(i)Arrivals of customers follow poison probability distribution with an arrival rate, λ, customers per time.
(ii) Service times are exponentially distribution with mean rate, μ, customers per unit of time. Obviously, service time varies from one customer to the other, but the same for each server.
(iii) There is no limit to the number in the queue, thus infinite calling population.
(iv) The queue discipline is first come first serve, and no balking or reneging is allowed.
(iv) If n<S, the number of customers are less than the number of servers, then, there will be no queue formed. This means that n-s, number of servers will be idle. The combine service rate is $\mu_n = n\mu$.
(v) If n≥s, then all servers are busy and the maximum number of customers in the queue is s-n. The combined service rate is $\mu_n = S\mu$, see Akinnuli and Olugbade (2014).
(vii) The total service rate must be more than the arrival rate. That is, $S\mu > \lambda$, otherwise the waiting line will be infinitely large that would give undesirable graph, (Adamu, 2015).

2.2 The performance measures

1. The probability that all the servers are idle is

$$P_0 = \left[\sum_{n=0}^{S-1} \frac{\left(\frac{\lambda}{\mu}\right)^n}{n!} + \frac{1}{S!}\left(\frac{\lambda}{\mu}\right)^S \left(\frac{S\mu}{S\mu - \lambda}\right)\right]^{-1}$$

2.The average number of customers waiting in the queue

$$L_q = \left[\frac{1}{(S-1)!}\left(\frac{\lambda}{\mu}\right)^S \frac{\lambda\mu}{(S\mu - \lambda)^2}\right] P_0$$

3. The average number of customers in the system

$$L_s = \left[\frac{1}{(S-1)!}\left(\frac{\lambda}{\mu}\right)^S \frac{\lambda\mu}{(S\mu - \lambda)^2}\right] P_0 + \frac{\lambda}{\mu}$$

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

$$W_q = \frac{\left[\frac{1}{(S-1)!}\left(\frac{\lambda}{\mu}\right)^S \frac{\lambda\mu}{(\lambda\mu - \lambda)^2}\right]}{\lambda} P_0$$

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

5. The average time a customer spends waiting in the system is

$$W_s = \frac{\left[\frac{1}{(s-1)!}\left(\frac{\lambda}{\mu}\right)^s \frac{\lambda\mu}{(\lambda\mu-\lambda)^2}\right]}{\lambda} P_0 + \frac{1}{\mu}$$

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

6. The probability that an arrival customer for services is

$$P(n \geq s) = \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s \left(\frac{s\mu}{s\mu-\lambda}\right) P_0$$

7. The probability that there are n-customers in the system is

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

Or

$$P_n = \frac{\left(\frac{\lambda}{\mu}\right)^n}{S!\,S^{n-S}} P_0; n > S$$

2.3 The economic trade-off queue management system

In developing the cost model for this queue problem, we consider the cost of both the waiting time of the customer and the cost of operating service time of the server. Thus, let the C_w be the waiting cost per hour of a customer, and C_s, the associated cost of each server hourly. Then, the total cost is given as

$$T_{ws} = C_w k + C_s S$$

Where k is the average number of customers in the system and S is the number of servers.

3.1 Presentation of data

Table 1: Arrival and Service Data from First Teller of the Bank

S/no	Arrival Time	Inter Arrival Time(min)	Service begins	Service Ends	Service time (min)
1	8.10	-	8.15	8.19	4
2	8.12	2	8.20	8.25	5
3	8.12	0	8.25	8.27	2
4	8.15	3	8.28	8.45	17
5	8.17	2	8.45	9.00	15
6	8.25	8	9.00	9.03	3
7	8.40	15	9.03	9.08	5
8	9.01	21	9.08	9.14	6
9	9.02	1	9.14	9.20	6
10	9.03	1	9.20	9.29	9
11	9.05	2	9.29	9.32	3
12	9.10	5	9.32	9.35	3
13	9.16	6	9.35	9.40	5
14	9.20	4	9.41	9.52	11
15	9.24	4	9.52	10.03	11
16	9.27	3	10.03	10.07	5
17	9.30	3	1007	10.15	8
18	9.30	0	10.15	10.19	4
19	9.40	10	10.19	10.22	3
20	9.44	4	10.22	10.30	8
21	9.45	1	10.31	10.39	8
22	9.52	7	10.40	10.46	6
23	9.55	3	10.46	10.53	7
24	9.58	3	10.53	10.58	5
25	9.59	1	10,58	11.01	3
26	10.00	1	11.01	11.03	2
27	10.02	2	11.04	11.06	2
28	10.03	1	11.06	11,15	9
29	10.05	2	11.15	11.19	4
30	10.13	8	11.19	11.23	4
31	10.18	5	11.23	11.30	7
32	10.25	7	11.30	11.38	8
33	10.25	0	11.38	11.40	2
		135			**200**

The arrival and service data from second, third, and the forth Tellers are summarized in the table below.

Table 2. Average arrival and service time of customers for four days among four servers

Day		Server1		Server2		Server3		Server4	
		λ_1	μ_1	λ_2	μ_2	λ_3	μ_3	λ_4	μ_4
Monday	Total	135	200	165	221	201	208	157	177
	Av.	4.2188	6.0606	4.4595	7.8919	4.9024	5.7778	5.4138	6.3214
Tuesday	Av.	5.9047	7.0952	3.9631	5.6671	4.8312	6.9832	4.9123	6.5924
Wednesday	Av.	5.7143	7.9048	5.6676	7.5051	3.1671	6.5932	4.9126	6.1459
Thursday	Av.	3.8571	5.5924	4.3291	7.7149	3.6712	5.1123	4.2569	6.3218
Total		19.6949	26.6530	18.4193	28.7790	16.5719	24.4665	19.4956	25.3815
Av. (λ_i, μ_i)		4.9237	6.6633	4.6048	7.1948	4.1430	6.1166	4.8739	6.3454

Thus, λ=4.6364 and μ=6.5800

3.2 The system characteristics of this server

Number of servers 4

Arrival rate (λ) 4.6364 customers/minute

Service rate (μ) per hour 6.5800 customers/minute

Average time between arrivals 0.2157

Average service time 0.1520

The probability that the system is empty is

$$P_0 = 0.4942$$

The average number of customers waiting in the queue

$$L_q = 0.0013 \text{ Customers}$$

The average number of customers waiting in the system

$$L_s = 0.7059 \text{ Customers}$$

The average time a customer spends waiting in the queue

$$W_q = 0.0003 \text{ Minutes}$$

The average time a customer spends waiting in the system

$$W_s = 0.1523 \text{ Minutes}$$

The probability that an arrival customer must wait for service

$$P_s = 0.0062$$

The average traffic intensity is

$$\rho_0 = \frac{\lambda}{S\mu}$$
$$=0.1762$$

3.3 Estimate of the economic trade-off of the queue system

Actually, the cost of customer waiting-time may not be accurately determined by the managers of service providers, but a reasonable estimated value that might reflect the potential loss of a customer should a customer switch over to another bank because of unreasonable delays is pertinent. In this work, we make presumption the waiting cost of customer and also the attendant cost.

Assume C_w is estimated to be N100 per hour. That is N1.67 per minute waiting time. The cost of operating system could easily be determined by the management; hence it consists of the wages of teller attendant, cost of equipment and maintenance. If we assume that C_s =N150 per hour. That is N2.5 per minute service time

Then, the total cost of having four attendants is

$$T_{ws} = C_w k + C_s S$$
$$\Longrightarrow \quad T_{ws} = 1.67 \times 0.7059 + 2.5 \times 4$$

11.18 naira/minute

We now present the values of, P_0, L_s, L_q, W_q, W_s, P_s for selected values of $\frac{\lambda}{\mu}$ for four attendants. The values provided correspond to cases for which, $S\mu > \lambda$, hence the service rate is sufficient to serve all arrival customers.

Table 3. The performance characteristics of multiple-lines Multiple servers, when four servers are involved

S/n	$\frac{\lambda}{\mu}$	P_0	L_s	L_q	W_q	W_s	P_s	Utilization factor
1	0.1000	.9048	0.1000	0.0000	0.0000	0.1000	0.0000	0.0250
2	0.2000	.8187	0.2000	0.0000	0.0000	0.1000	0.0001	0.0500
3	0.3000	.7408	0.3000	0.0000	0.0000	0.1000	0.0003	0.0750
4	0.4000	.6703	0.4001	0.0001	0.0000	0.1000	0.0008	0.1000
5	0.5000	.6065	0.5003	0.0003	0.0001	0.1001	0.0018	0.1250
6	0.6000	.5487	.6006	0.0006	0.0001	0.1001	0.0038	0.1500
7	0.7000	.4965	0.7013	0.0013	0.0002	0.1002	0.0060	0.1750
8	0.7046	.4942	0.7059	0.0013	0.0003	0.1523	0.0062	0.1762
9	0.8000	.4491	0.8024	0.0024	0.0003	0.1003	0.0096	0.2000
10	0.9000	.4062	0.9042	0.0042	0.0005	0.1005	0.0143	0.2250
11	1.0000	.3673	1.0068	0.0068	0.0007	0.1007	0.0204	0.2500
12	1.1000	.3321	1.1106	0.0106	0.0010	0.1010	0.0279	0.2750
13	1.2000	.3002	1.2159	0.0159	0.0013	0.1013	0.0370	0.3000
14	1.3000	.2712	1.3230	0.0230	0.0018	0.1018	0.0478	0.3250
Sum								2.4512
Av.								**0.1750**

From Table 3 above, the average of the entire utilization factor is 0.1750, and hence the bench marks for the optimal performance of the system. This means that any value below the average utility factor will invariably increase the idle time of the servers, reduces the production time and increase the cost of service. Based on this, one can say that the system is working optimally as the average utility is above the calculated utility factor (0.1762) obtained from the data collected.

In the introductory part of this paper, we informed that the more the number of servers the more the operating cost of the system and invariably the less the waiting line in the system. Again, in as the operators would not like to lose customers through balking and reneging because of length of queue in the system, they would also want to use the minimum number of servers that would reduce the system length because of cost. Thus, the number of server must be optimized.

Table 4. Summary of the performance characteristics of the bank with different number of servers

No of servers	P_0	L_s	L_q	W_q	W_s	Utility factor	% of the system idle	% of the system busy
2	.4790	0.8045	0.0999	0.0215	0.1735	0.3523	47.90	35.23
3	.4928	0.7161	0.0115	0.0025	0.1545	0.2349	49.28	23.49
4	.4942	0.7059	0.0013	0.0003	0.1523	0.1762	49.42	17.62
5	.4963	0.7048	0.0001	0.0000	0.1520	0.1409	49.63	14.09
Sum								90.43
Av.								22.61

From the Table 4 above, the average waiting time of customers in the system are 0.1735, 0.1545, 0.1523, and 0.1520 hours for 2, 3, 4, and 5 servers respectively. This shows that as the number of servers increases, the average customers waiting in the system decreases and vise vase.

Again, the more servers in the system, the more idle time of the system and less utility factor of the system.

Finally, the appropriate number of servers that would optimize the system is in the neighborhood of 3 servers, hence the average above.

4.1 Conclusion

The number of four servers being in use by the bank is working optimally as shown in Table 3 above. However, the optimal number of servers for the bank is three in order to reduce the idle time and cost of the operation.

Recommendation

Queuing theory is very effective in managing waiting lines and cost in banks, we recommend the application of this instrument in other branches of the bank.

References

Adamu, M. (2015). Analysis of Single Queue-Single Server and Single Queue-Multi Server Systems using Simulation: Case Studies of ECO and First Banks. *International Journal of Scientific and Engineering Research, 6(1)*, 1284-1296.

Akinnuli,B.O & Olugbade,T.O (2014). Application of Queue Model for Performance Assessment of Multi-Channel Multi-servers Motor Spirit Filling Station. *International Journal of Scientific and Engineering Research, 3(7)*, 74-80.

Ezeliora, C.D, Ogunoh A.V, Umeh M.N, Mbeledeogu N.N (2014). Analysis of Queue System Using Single-Line Multiple Servers System: A Case Study of Shoprite Plaza Enugu State, Nigeria. *International Journal of Scientific and Engineering Research 3(3)*. 364-373

Sharma, J.K (2013). *Operations Research: Theory and Applications*. Fifth edition, India, Macmillan Publishers LTD.

Uche, P.I & Ugah, T.E (2014). Priority Assessment Based on Patronage Value. *International Journal of Scientific and Engineering Research, 5(4)*. 1-5.

Yuncheng, D.D & Liang, S (2002). *System Simulation and Application.* Beijing, Mechanism Industry Press.