Research Journal of Management Sciences ISSN 2319–1171
Vol. **9(1),** 1-10, January **(2020)** *Res. J. Management Sci.*

Performance improvements using multi server queuing model by reducing customer wait time of a bank

Tanzila Azad*, Amitsaha, Hasnain Shaikh Sithil and Rafiul Rafi
Department of Mechanical and Production Engineering, Ahsanullah University of Science and Technology, Dhaka, Bangladesh
azadshashi33@gmail.com

Available online at: www.isca.in, www.isca.me
Received 31st October 2018, revised 24th April 2019, accepted 25th October 2019

Abstract

In a bank, the customer expects to receive service quickly. The queuing in relation to the time spent by customers has a great effect, for accessing banking services which is progressively becoming a major source of concern to most of the bank service provider. Customers waiting too long in the queue could results to waiting cost in contrary not providing enough service capacity results in excessive waiting time and cost. So, these fallouts in a relation between waiting time of customer and service capacity. A study was done to a bank named Agrani Bank Limited using a Multi-server queuing model and the waiting as well as service costs determined to get optimal level of service. Using TORA optimization Software along with descriptive analysis the data was analyzed. An assessment was done to regulate the wait time of customer which was expected and on expected queue length. The outcome of the study gives recommendation of adding one more server than the previous case for improving the waiting time of customer. For discussing the cost impact of adding one more server an economic cost analysis is also provided.

Keywords: Multi-server queuing model, optimization, TORA software, customer wait time.

Introduction

In banks queue is a common sight. Via French and the Latin caudal meaning the word queue means "tail". Queue management has been the sector where the manager faces huge challenge as customers waiting in line to receive services in any service system are inevitable. The waiting time of customer is considered one of the fundamental characteristics in a bank. For that waiting time of customer directly allied to the satisfaction of customer. For any effort of recovering the waiting time of customer will develop the satisfaction of customer. Ravi Nasit applied queing theory in banking application to enhance service efficiency[1]. For determining the optimal level of service Nityangini Jhala and Pravin Bhathawala also applied queuing model[2]. Using a remote and local service could provide the QMS with better performance[3]. Using an M/M/1 queue simplifies the modeling process[4]. The round-robin method of will not be implemented, as customers need to be serviced in order, until successfully helped[5]. It is known that queuing delay could have a negative impact on the performance of a system thus different servicing models will be introduced into the system to measure performance improvements[6]. Rather than customer satisfaction Lee and Lambert point out those customer oriented efforts which focus on what management thinks and believes[7]. Relationships between the waiting time of customer, perception of service quality and customer satisfaction are very important which is also disputed by Lee and Lambert[8]. This study has the determination to provide a solution for management to decrease the waiting time of customer at a economically feasible bank.

Relating to service operations like number of servers, service time, customer arrival rate and their proportion, total customer wait time, and queue behavior is presented by a thorough analysis of a number of variables (or parameters). In this paper, a study has been conducted to Agrani Bank that adopts the M/M/S queuing model. Queuing models provide the analyst with a powerful tool for designing and evaluating the performance of queuing systems[9]. The queue discipline maintains first in first out (FIFO) and the arrival is strictly random having Poisson distributed arrival times and exponentially distributed service times. With earlier research this study maintains consistency accomplished by Chou and Liu, and Curin, Vosko, Chan, and Tsimhoni which displays that addition of one more server lessens the waiting time of customer[10-12]. The customers choose a queue of server according to some mechanism like shortest queue as well as shortest workload[13]. Modeling and managing the performances of a service industry is challenged due to the randomness of arrival and service time. This paper considers one bank in Addis Ababa and tries to model its service performances as a multiple server queue structure system. Secondary data related to arrival rate, service time and cost are collected from the same bank for a period of one month. The data was fitted in the model and tested using the enterprise dynamic simulation system. The outcomes of the study show that both arrival of customers and service time rate of servers follow a Poison exponential probability distribution respectively. Moreover, the optimum number of servers resulted from model is found to be five serves but with a low server utilization rate. So, the bank is suggested to improve its server utilization[14].

Improving the performance of service industries where arrival and service time are random and performed by human employee is a complex decision environment. The banking industry is a good example that shares the above scenario. To make a better decision and to improve the system performance decision makers assisted themselves by using intelligent system. Among the many others simulation system is one of the intelligent systems that help to model a complex environment.

Problem description: Several studies documenting customer dissatisfaction with long waiting times indicate that this is a major problem in banking practice and a common source of anxiety and dissatisfaction among customers and in many cases, bankers. In this paper, a case of Agrani Bank has been observed that uses the M/M/S queuing model. The queue discipline is first in first out (FIFO) and the arrival is random having Poisson distributed arrival times and exponentially distributed service times. For having optimum service level a study has been conducted to increase the efficiency of the models in terms of utilization and waiting length, improving the number of queues so customers will not have to wait longer. The information gathered through perception in a bank demonstrates that the framework limit is 300 customers with 3 servers, between inter arrival time for 60 customers is 75 minutes and the time taken by 60 customers to be served is 150 minutes.

Methodology

This section presents the method of data collection, data validation and the mathematical details of the M/M/S queue model employed in the study.

Data collection: The documents used in the Queuing model is gathered for an arrival time of each customer in two days. The records for number of customers in a queue, their arrival-time and departure-time were taken without disturbing the employees. The whole process of the service unit was observed and recorded using a time-watch during the same time period for each day. The bank of our concern consists of a multiple server (counter) for customers and the bank adopts First in First out (FIFO) scheduling for serving the customers.

Table-1: Collected data for server 1.

Customer no	Time of arrival (min) (am)	Time of inter arrival (min)	Time of service (min)	Time of service starts (min) (am)	Time of service ends (min) (am)	Time of customer waits in queue (min)	Idle time of the server (min)	Time of customer spend in the system (min)
1	11:09	0	3	11:09	11:12	0	0	3
2	11:11	2	1	11:12	11:13	1	0	2
3	11:11	0	3	11:13	11:16	2	0	5
4	11:13	2	4	11:16	11:20	3	0	7
5	11:14	1	2	11:20	11:22	6	0	8
6	11:14	0	1	11:22	11:23	8	0	9
7	11:14	0	2	11:23	11:25	9	0	11
8	11:15	1	4	11:25	11:29	10	0	14
9	11:17	2	2	11:29	11:31	12	0	14
10	11:19	2	3	11:31	11:34	12	0	15
11	11:19	0	4	11:34	11:38	15	0	19
12	11:21	2	1	11:38	11:39	17	0	18
13	11:23	2	2	11:39	11:41	16	0	18
14	11:23	0	1	11:41	11:42	18	0	19
15	11:27	4	3	11:42	11:45	15	0	18
16	11:27	0	1	11:45	11:46	18	0	19
17	11:29	2	2	11:46	11:48	17	0	19
18	11:29	0	4	11:48	11:52	19	0	23
19	11:30	1	2	11:52	11:54	22	0	24
20	11:32	2	4	11:54	11:58	22	0	24
		∑23	∑49					

Table-2: Collected data for server 2.

Customer no	Time of arrival (min) (am)	Time of inter arrival (min)	Time of service (min)	Time of service starts (min) (am)	Time of service ends (min) (am)	Time of customer waits in queue (min)	Idle time of the server (min)	Time of customer spend in the system (min)
1	11:10	0	3	11:10	11:13	0	0	3
2	11:11	1	1	11:13	11:14	2	0	3
3	11:11	0	3	11:14	11:17	3	0	6
4	11:12	1	2	11:17	11:19	5	0	7
5	11:14	2	1	11:19	11:20	5	0	6
6	11:15	1	3	11:20	11:23	5	0	8
7	11:18	3	1	11:23	11:24	5	0	6
8	11:19	1	3	11:24	11:27	5	0	8
9	11:23	4	4	11:27	11:31	4	0	8
10	11:23	0	2	11:31	11:33	8	0	10
11	11:24	1	2	11:33	11:35	9	0	11
12	11:26	2	1	11:35	11:36	9	0	10
13	11:27	1	5	11:36	11:41	9	0	14
14	11:29	2	2	11:41	11:43	12	0	14
15	11:29	0	3	11:43	11:46	14	0	17
16	11:30	1	4	11:46	11:50	16	0	20
17	11:30	0	3	11:50	11:53	20	0	23
18	11:32	2	2	11:53	11:55	21	0	23
19	11:32	0	2	11:55	11:57	23	0	25
20	11:33	1	1	11:57	11:58	24	0	25
		∑23	∑48					

Table-3: Collected data for server 3.

Customer no	Arrival time (min) (am)	Inter arrival time (min)	Service time (min)	Service time starts (min)	Service time ends (min) (am)	Customer waits in queue (min)	Idle time of the server (min)	Time customer spend in the system (min)
1	11:32	0	3	11:32	11:35	0	0	3
2	11:34	2	2	11:35	11:37	1	0	3
3	11:35	1	1	11:37	11:38	2	0	3
4	11:36	1	3	11:39	11:42	3	1	6
5	11:37	1	3	11:42	11:45	5	0	8
6	11:38	1	2	11:45	11:47	7	0	9
7	11:40	2	3	11:51	11:54	11	4	14
8	11:42	2	3	11:56	11:59	14	2	17
9	11:44	2	2	12:01 pm	12:03 pm	17	2	19
10	11:45	1	3	12:05 pm	12:08 pm	20	2	23
11	11:47	2	4	12:12 pm	12:16 pm	25	4	29
12	11:49	2	1	12:16 pm	12:17 pm	27	0	28
13	11:51	2	4	12:22 pm	12:26 pm	31	5	25
14	11:52	1	4	12:29 pm	12:33 pm	37	3	41
15	11:54	2	1	12:37 pm	12:38 pm	43	4	44
16	11:55	1	2	12:41 pm	12:43 pm	46	3	48
17	11:57	2	4	12:43 pm	12:47 pm	46	0	50
18	11:59	2	4	12:47 pm	12:51 pm	48	0	52
19	12:01 pm	2	3	12:55 pm	12:58 pm	54	4	57
20	12:01 pm	0	1	1:00 pm	1:01 pm	59	2	60
		∑29	∑53					

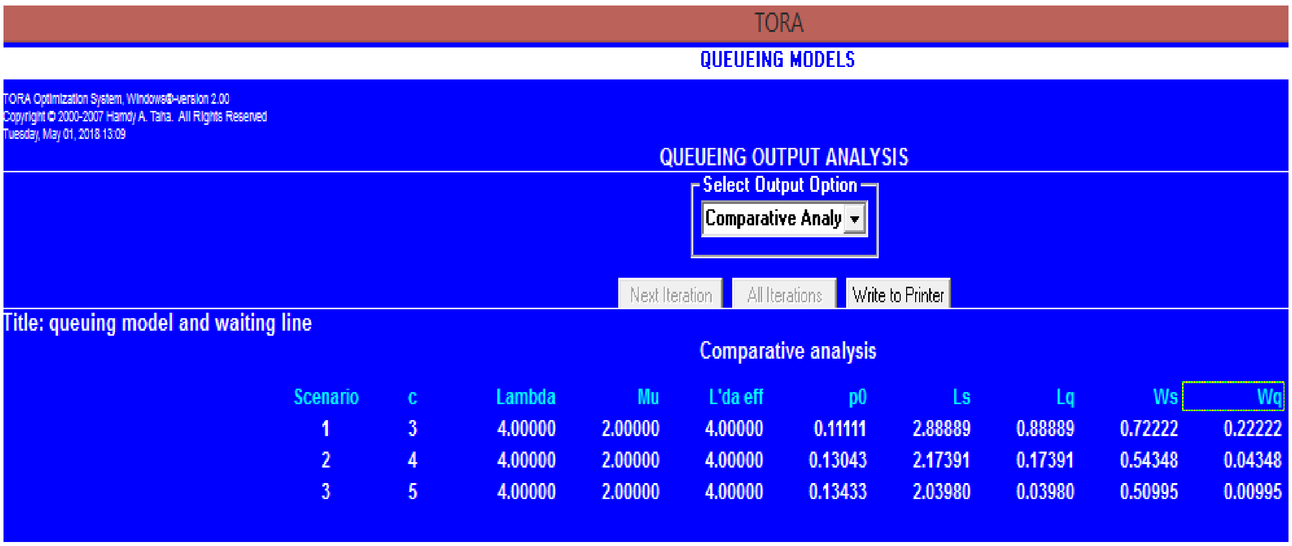
M/M/S queuing simulation for Agrani Bank using TORA SOFTWARE: The simulation was finished utilizing TORA software (Figure-1). It was discovered that inter arrival rate and service rate. Poisson distribution takes after between inter-arrival rate and service rate takes after Exponential distribution as per information gathered from Agrani Bank constrained. As a result, analyzing scenario M/M/S queuing model was considered

Arrival rate, $\lambda = \frac{N}{T} = \frac{300}{75} = 4$

Service rate, $\mu = \frac{N}{S} = \frac{300}{150} = 2$

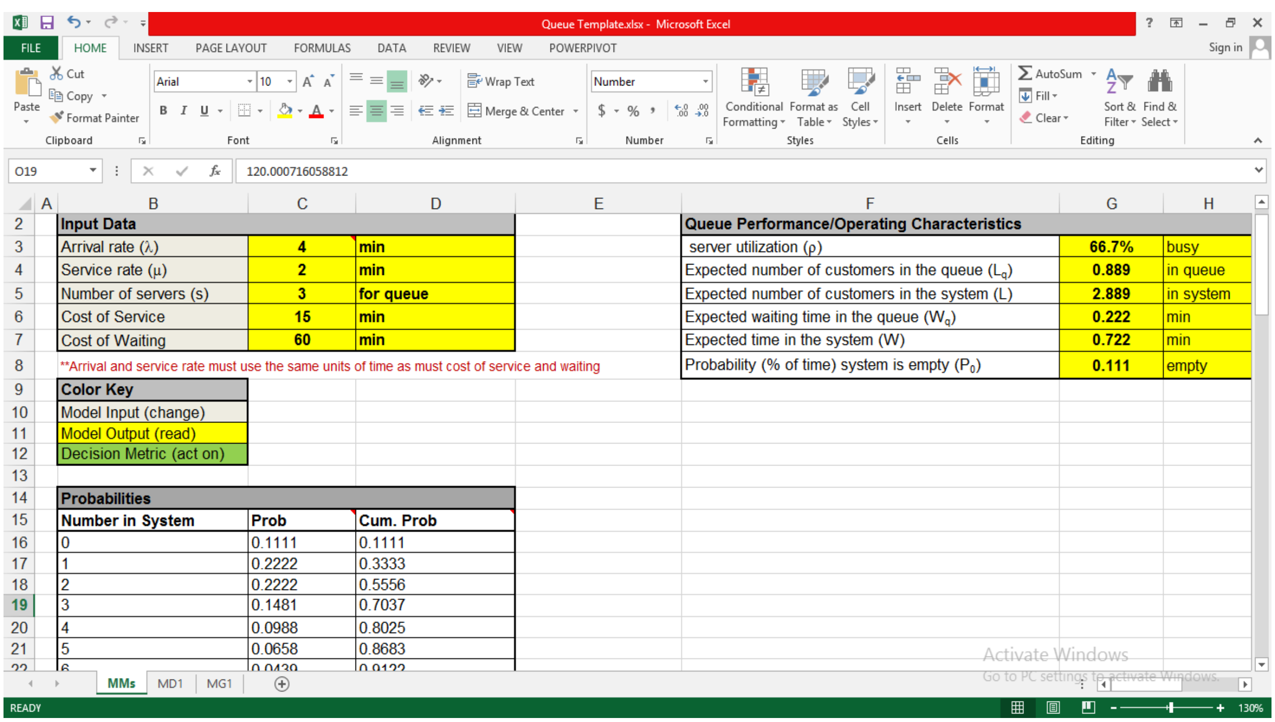
Data validation: The process of ensuring that data which have undergone data cleansing to ensure that they have data quality, which are both correct and useful are introduced as data validation. Figure-4, 5 and 6 represents the data validation for 3 server, 4 server, 5 server sequentially.

Table-4 has been the summarized from Figure-2, 3 and 4 where queue performance has been characterized for multiple server system.



Scenario	c	Lambda	Mu	L'da eff	p0	Ls	Lq	Ws	Wq
1	3	4.00000	2.00000	4.00000	0.11111	2.88889	0.88889	0.72222	0.22222
2	4	4.00000	2.00000	4.00000	0.13043	2.17391	0.17391	0.54348	0.04348
3	5	4.00000	2.00000	4.00000	0.13433	2.03980	0.03980	0.50995	0.00995

Figure-1: Comparative analysis of waiting line of multiple scenario.



Input Data		
Arrival rate (λ)	4	min
Service rate (μ)	2	min
Number of servers (s)	3	for queue
Cost of Service	15	min
Cost of Waiting	60	min

Queue Performance/Operating Characteristics		
server utilization (ρ)	66.7%	busy
Expected number of customers in the queue (L_q)	0.889	in queue
Expected number of customers in the system (L)	2.889	in system
Expected waiting time in the queue (W_q)	0.222	min
Expected time in the system (W)	0.722	min
Probability (% of time) system is empty (P_0)	0.111	empty

Number in System	Prob	Cum. Prob
0	0.1111	0.1111
1	0.2222	0.3333
2	0.2222	0.5556
3	0.1481	0.7037
4	0.0988	0.8025
5	0.0658	0.8683

Figure-2: Data validation for 3 server.

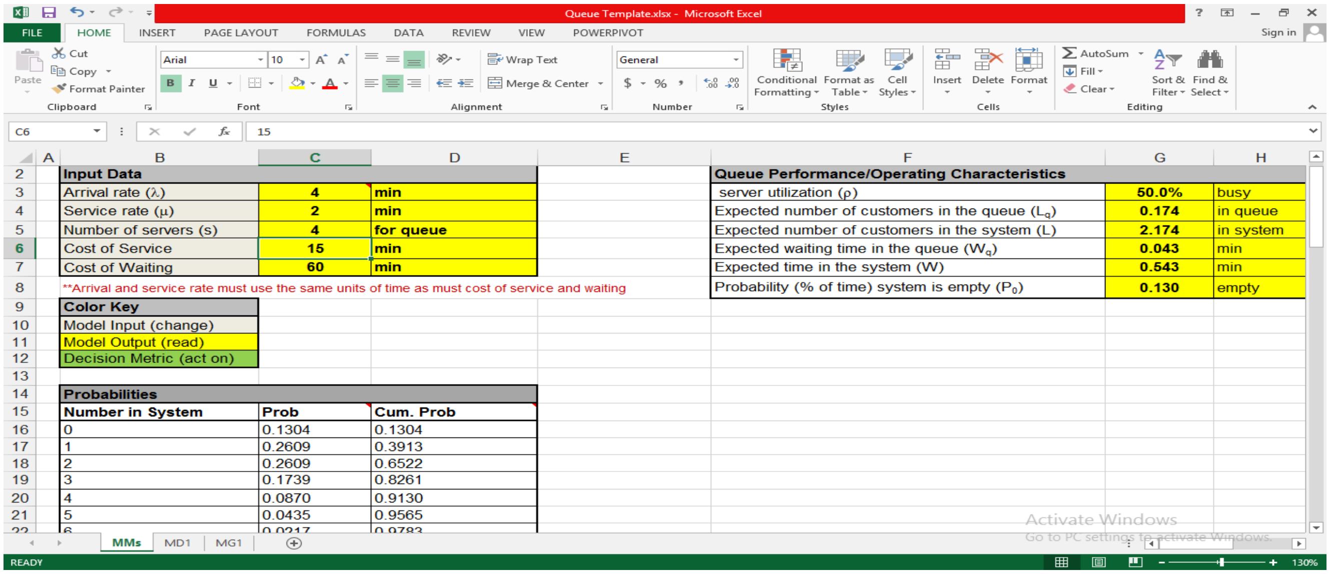


Figure-3: Data validation for 4 server.

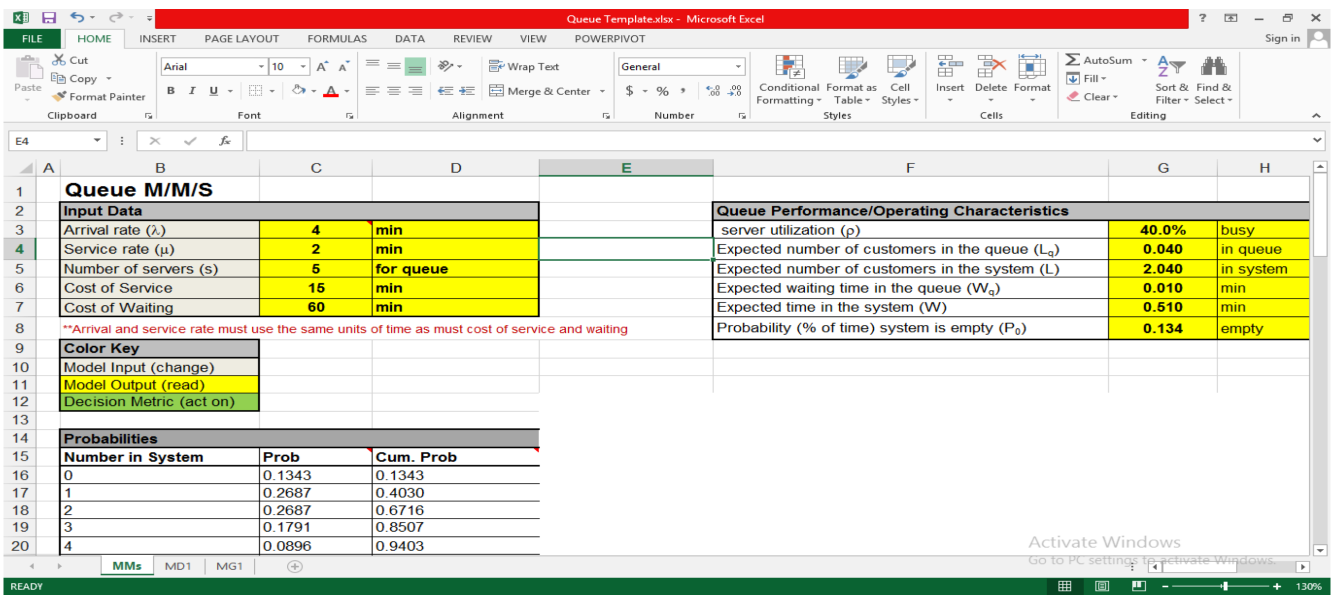


Figure-4: Data validation for 5 server.

Table-4: Queue performance characteristics.

No of servers	3	4	5	Servers availability
Utilization of Average server (p)	66.7%	50.0%	40.0%	busy
No. of average customers in the queue (L_q)	0.889	0.174	0.040	in the queue
No. of average customers in the system (Ls)	2.889	2.174	2.040	in system
Average time of waiting in the queue (W_q)	0.222	0.043	0.010	minimum
Average time in the system (Ws)	0.722	0.543	0.510	minimum
Probability (% of time) system is empty (P_0)	0.111	0.130	0.134	empty

Queuing Cost Calculation: For evaluating and determining the optimum number of servers in the system, two trade off costs must be considered in making these decisions: i. Cost of service, ii. Cost of waiting.

The management is helped by this economic analysis of costs to make a trade-off between the decreased waiting time costs of customers derived from providing that service and the increased costs of providing better service.

According to Agrani Bank Authority there is waiting cost and service cost per min. It varies from time to time. An assumption of cost has been made in this case, the cost of service is 15tk/min and cost of waiting is 60tk/min.

Cost of service, Cs= 15tk/min
Cost of waiting, Cw=60tk/min
Expected Service Cost E(SC)=SCs
Expected Waiting Cost E(WC)=(λWs)Cw
Expected Total Cost E(TC)= SCs + (λWs)Cw

$$=3 \times 15 + (4 \times .72222) \times 60$$

=218.33tk

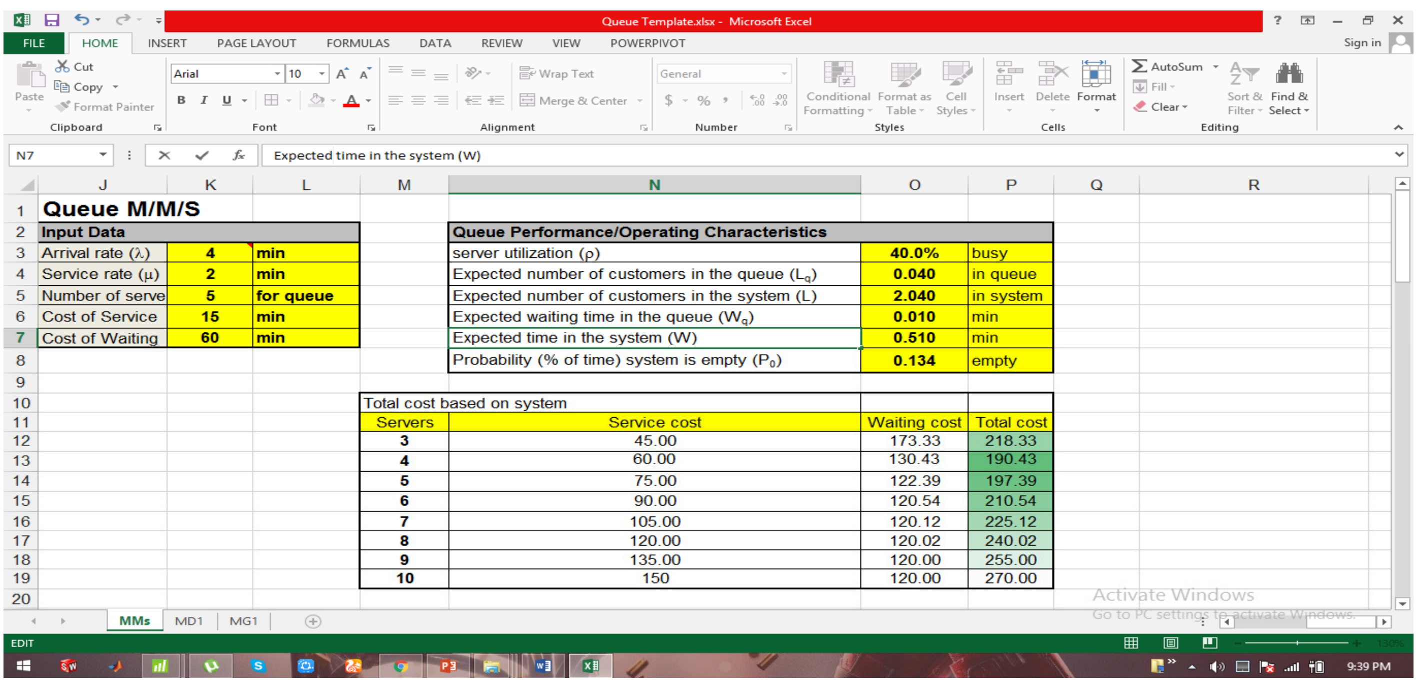
Where: S=No. of servers, Cs=Each server's service cost, λ =Rate of Arrival, Ws=Average time of an arrival spend in the system, Cw= Cost of opportunity waiting by customers.

The following Figure-5 representing the changes in total cost with the increasing number of servers.

Results and discussion

Analyzing the data of Table-4 with a view to minimizing the total cost of the company the following table has been drawn to measure the performance of the M/M/S queuing model for the concerned company.

From Table-5 it can be seen that total cost is minimum when the no. of server is 4. So, if the bank keeps 4 server then it will be optimized and thus customer satisfaction level will be high as with the increase no of server they don't have to wait longer time period in the queue. Thus increasing the utilization rate and minimization of average waiting time has also done.



Queue M/M/S

Input Data		
Arrival rate (λ)	4	min
Service rate (μ)	2	min
Number of serve	5	for queue
Cost of Service	15	min
Cost of Waiting	60	min

Queue Performance/Operating Characteristics		
server utilization (ρ)	40.0%	busy
Expected number of customers in the queue (L_q)	0.040	in queue
Expected number of customers in the system (L)	2.040	in system
Expected waiting time in the queue (W_q)	0.010	min
Expected time in the system (W)	0.510	min
Probability (% of time) system is empty (P_0)	0.134	empty

Total cost based on system

Servers	Service cost	Waiting cost	Total cost
3	45.00	173.33	218.33
4	60.00	130.43	190.43
5	75.00	122.39	197.39
6	90.00	120.54	210.54
7	105.00	120.12	225.12
8	120.00	120.02	240.02
9	135.00	120.00	255.00
10	150	120.00	270.00

Figure-5: Cost calculation.

Table-5: Summery table of performance measure of multi-server queuing model at the bank.

Scenario	No of servers (S)	Arrival rate (λ)	Service rate (μ)	Utilization (%)	Po	Ls	Lq	Ws (hours)	Wq (hours)	Total cost(hrs)
1	3	4	2	66.7	0.11111	2.88889	0.88889	0.72222	0.22222	218.33
2	4	4	2	50	0.13042	2.17391	0.17391	0.54348	0.04348	190.43
3	5	4	2	40	0.13433	2.03980	0.03980	0.50995	0.00995	197.43

Table-6: Comparison between existing and proposed queuing model results.

	Existing result	Proposed result
Arrival rate, λ	4 minute	4 minute
Service rate, μ	2 minute	2 minute
Utilization of Average server (p)	66.7 %	50.0 %
No. of average customers in the queue (L_q)	0.889 queue	0.174 queue
No. of average customers in the system (Ls)	2.889 system	2.174 system
Average waiting time in the queue (W_q)	0.222 minute	0.043 minute
Average time in the system (Ws)	0.722 minute	0.543 minute
Probability (% of time) system is empty (P_0)	0.111 empty	0.130 empty

From Table-7 it can be seen that waiting cost likewise has been decreased contrasted with past existing waiting cost at Agrani Bank Limited.

Graphical analysis: Figure-6 representing the expected service cost vs. no of server's graph. From this figure see that if increases the number of server the service cost will be increased.

Figure-7 representing the expected waiting cost vs. no of server's graph. It can be seen that, if increases the number of server the service cost will be decreased.

Figure-8 representing the expected total cost vs. no of server's graph. From this graph it can be seen that with the increased number of server the total cost will be increased for 3 server but decreased for 4 server.

Table-7: Comparison of queuing cost.

Type of cost	Existing queuing cost	Proposed queuing cost	Reduction of queuing cost
Expected service cost	45	60	27.96
Expected waiting cost	173.33	130.32	
Expected total cost	218.3	190.32	

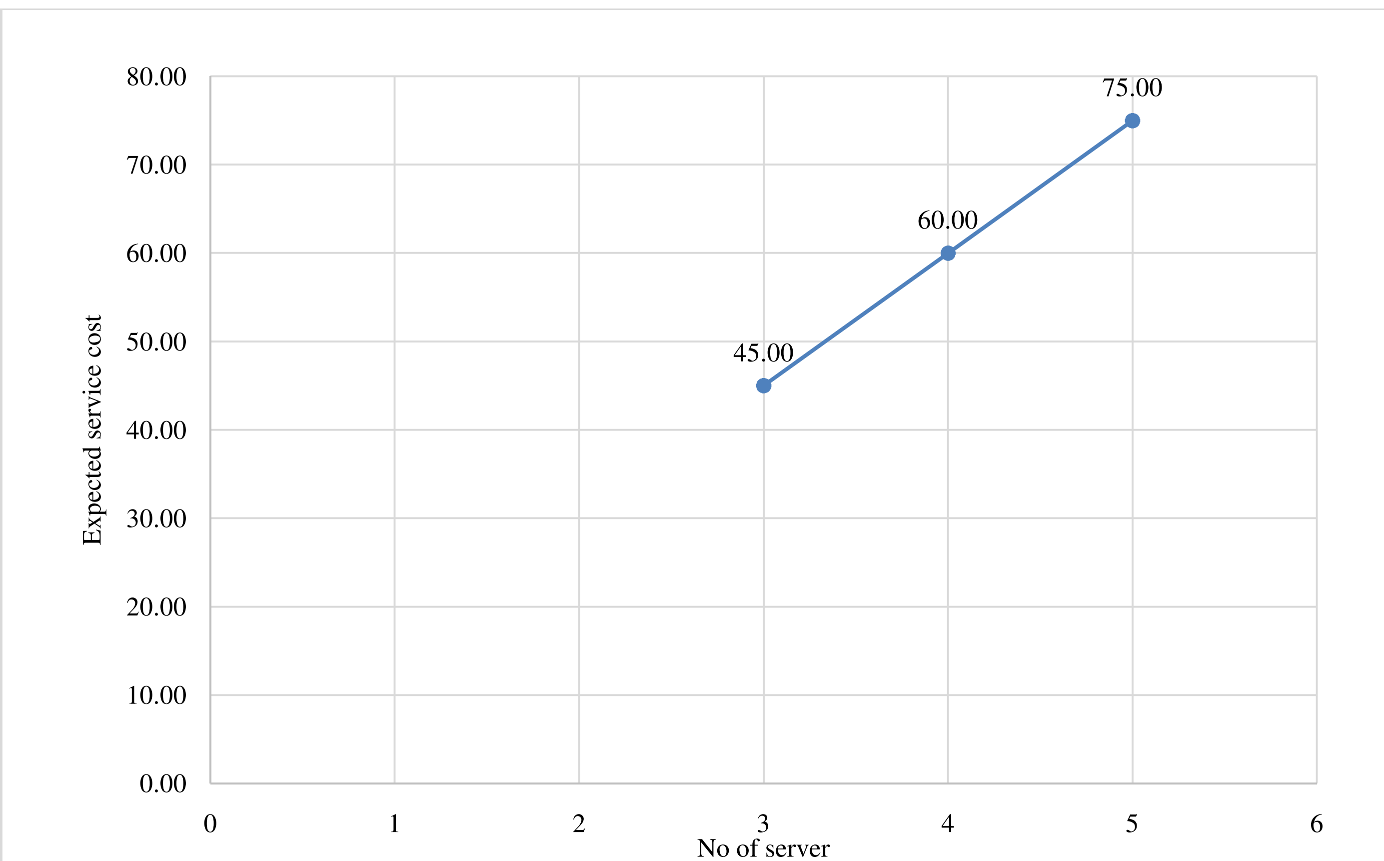


Figure-6: Expected service cost vs. No of server.

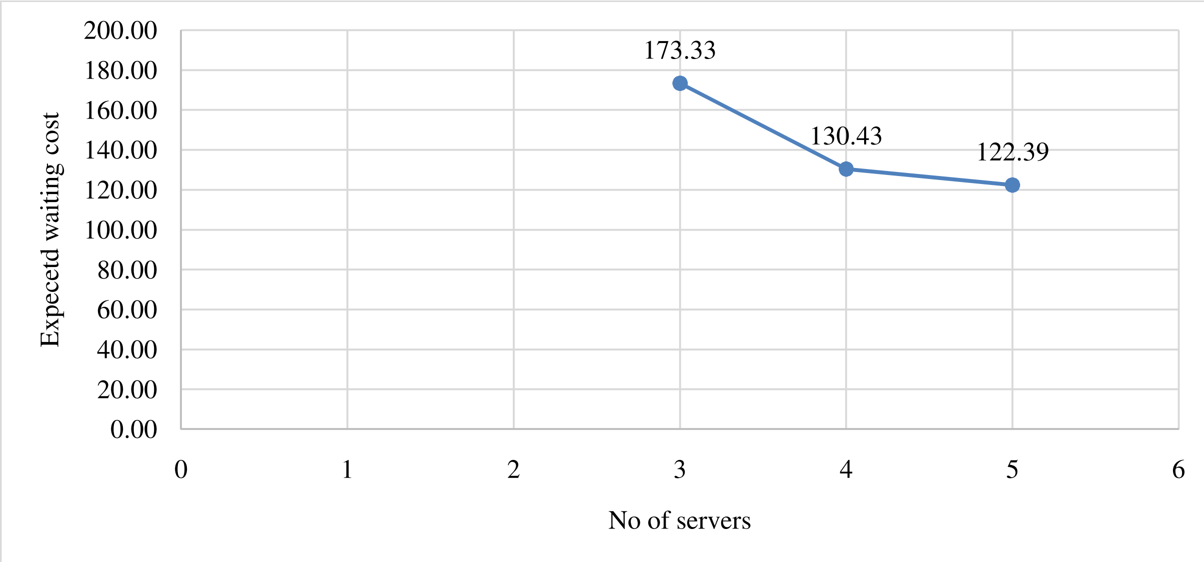


Figure-7: Expected waiting cost vs. No of server.

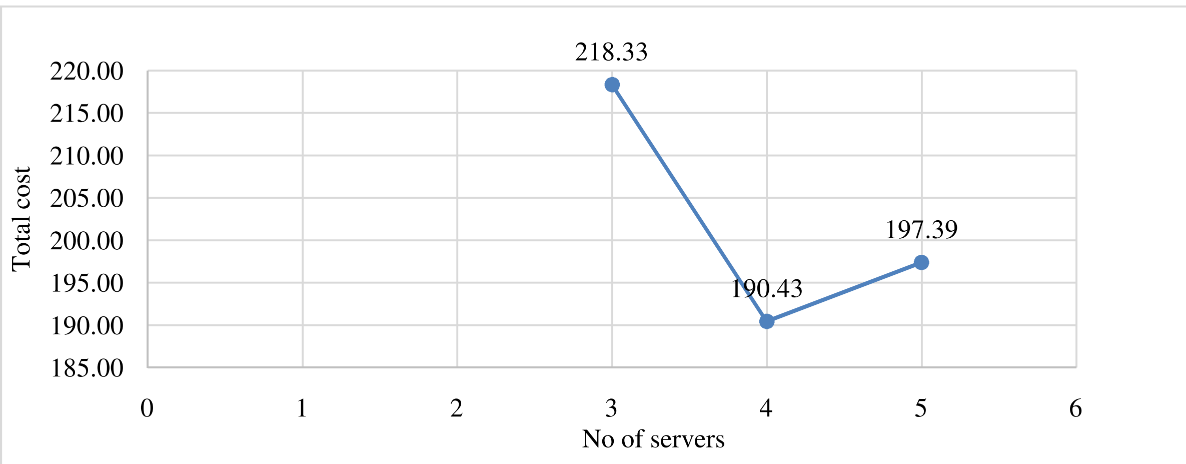


Figure-8: Expected total cost vs. No of server.

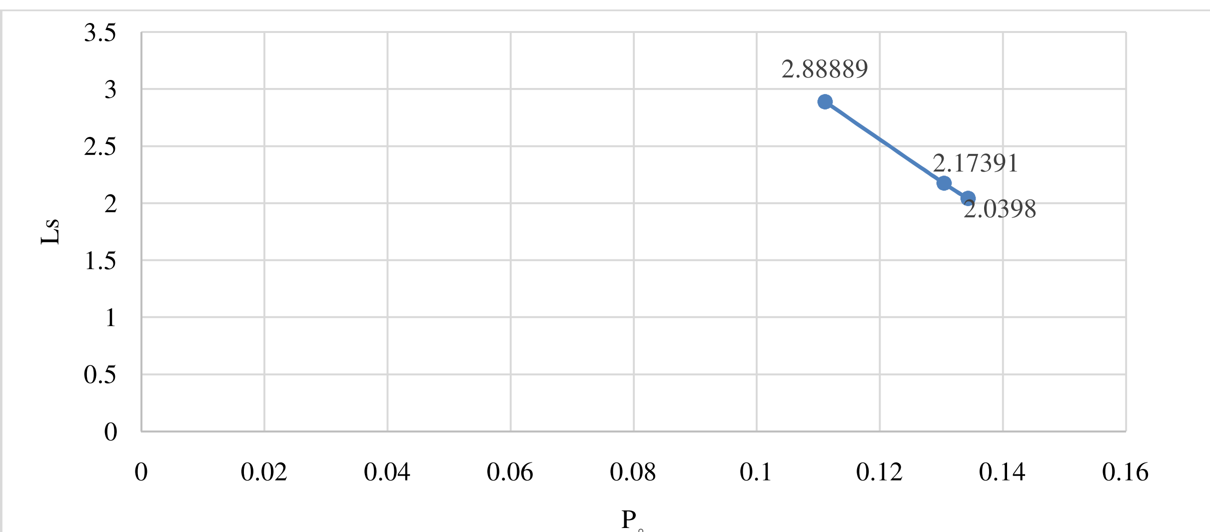


Figure-9: No. of average customers in the system (Ls) vs. Probability (% of time) that the system is empty ($P_{\circ}$).

Conclusion

Through the analysis the result can be determined that the service level increases has been optimized simultaneously with the reduction of customer waiting time, along with the improvement of the efficacy of service level, the customer satisfaction has also increased. Moreover, it is found that the total system cost is minimized at some specific numbers of service windows. This optimal model of the queuing is viable can be demonstrated through the evaluation. The characteristics of queuing at the Agrani Bank Limited were evaluated using a M/M/S queuing model. The length of queue, customers wait time and the overall utilization of servers could be improved which is displayed by the outcomes of the analysis. Service cost increases as a bank tries to raise its level of service with the increase of the service capacity level of servers at the bank from three to four. The cost of time spent waiting on the line drops with the increase of service.

References

1. Toshiba S., Sanjay S. and Anil K.K. (2013). Application of queuing theory for the improvement of bank service. *International Journal of Advanced Computational Engineering and Networking*, 1(4), 15-18.

2. Jhala N. and Bhathawala P. (2016). Application Of Queuing Theory In Banking Sector. *IOSR Journal of Mathematics (IOSR-JM)*, 12(2), 73-75.

3. Bagchi S. (2015). Analyzing Distributed Remote Process Execution Using Queuing Model. *Journal of Internet Technology*, 16(1), 163-170.

4. Mahmood K., Chilwan A., Østerbø O. and Jarschel M. (2015). Modelling of OpenFlow-based software-defined networks: the multiple node case. *IET Networks*, 4(5), 278-284.

5. Al-Mogren A., Iftikhar M., Imran M., Xiong N. and Guizani S. (2015). Performance Analysis of Hybrid Polling Schemes with Multiple Classes of Self-Similar and Long-Range Dependent Traffic Input. *Journal of Internet Technology*, 16(4), 615-628.

6. Rasmussen A., Yu H., Ruepp S., Berger M.S. and Dittmann L. (2014). Efficient round-robin multicast scheduling for input-queued switches. *IET Networks*, 3(4), 275-283.

7. Lee W. and Lambert C.U. (2000). Impact of waiting time on evaluation of service quality and customer satisfaction in foodservice operations. *Foodservice Research International*, 12(4), 241-254.

8. Lee K.W. and Lambert C.U. (2007). Using simulation to manage waiting time in a cafeteria. *Information Technology in Hospitality*, 15(4), 127-141.

9. Banks J., Carson J.S., Nelson B.L. and Nicol D.M. (2001). Discrete-Event System Simulation. *Prentice Hall International series*, 3rd edition, 24-37.

10. Chou C.Y. and Liu H.R. (1999). Simulation study on the queuing system in a fastfoodrestaurant. *Journal of Restaurant & Foodservice Marketing*, 3(2), 23-36.

11. Curin S.A., Vosko J.S., Chan E.W. and Tsimhoni O. (2005). Reducing service time at a busy fast food restaurant on campus. In Proceedings of the 37th conference on Winter simulation. Winter Simulation Conference, 2628-2635.

12. Kuhl M.E., Steiger N.M., Armstrong F.B. and Joines J.A. (2003). Proceedings of the 2005 Winter Simulation Conference. *Ann Arbor, MI: The University of Michigan*, 2629-2635.

13. Adan I.J.B.F., Boxmal O.J. and Resing J.A.C. (2000). Queuing models with multiple waiting lines. *Department of Mathematics and Computer Science*, Eindhoven University of Technology, 1-105.

14. Berhan E. (2015). Bank Service Performance Improvements using Multi-Sever. *Journal of business and management,* 17(6), 65-69.