

Application of Queueing Theory to ATM Service

Najeeb Al-Matar
Saudi Arabia

Abstract

Unmanaged queues may sometimes become detrimental to the operations of the banking activities whereby it can even become a major problem to the management of the banking services. The use of Automated Teller Machines (ATMs) in banks prevents losing customers because of long waits in the queue. For this case, application of Queueing theory can be of benefit to the running of the banking activities thus reducing the waiting time and improving on utilization time. In this case study, the researcher uses a direct non-participatory observational method to collect data over a period of one month within the banking hours between 9:30 AM to 4:30 PM in one of the local banks in town which are analyzed and various workings illustrated. Later, the researcher gives recommendations on how the ultimate goal of improving service time and reducing waiting time can be achieved.

Additionally, arrival rate, waiting time in the queue, service rate, utility rate and approximated average number of customers in a line are illustrated to enable the surveyor to solve the problem in Queueing model analysis. Further, M/M/1 Queueing model has been used to explain Queueing system at the ATM facility. In this case, the Queueing theory is employed to solve the problem of long queues to achieve the required satisfactory service. Lastly, the researcher concludes the importance of Queueing model analysis to a bank ATM especially when maximum service delivery and customer satisfaction is an enterprise's primary mission to earn an economic profit and remain economically relevant in the banking industry.

Introduction

Queueing is a systematic arrangement of people, jobs or customers to obtain a required service and could be dealt with one at a time. According to Crowther & Davila, (2011), technology offers critical solutions to both old and new problems facing humankind. For instance, not only that technology has changed how business is conducted, but it has also virtually enhanced all spheres and aspects of human dignity. Whether willingly or unwillingly, people come into contact with the products of technology in every activity they perform daily. Equally, technology has made those tasks that were once done in a stressful and unproductive to be completed faster, easier, and more accurate. In the banking industry, information and technology have enhanced data retrieval, storage, processing and much more. For instance, the ATMs have made it easier in reducing congestion in the banking halls. Further, Crowther & Davila (2011) say that the world economy has been made so liquid, and the demand for money has increased necessitating more money transactions for money. In this regard, there are tremendous reforms in financial sector especially banks which tend to reduce cost and maximize profit to satisfy customers' utility maximally.

With all these efforts, it has been realized that a queue remains as one of the criteria employed during the delivery of services in banks. This is profoundly captured in banks ATMs or within the banking halls. Apart from the banking industry, queues are also experienced in traffic lights, clinics, supermarkets, bus stops and many others. Krolark, Felts, and Marble (1970) say that queues are commonly formed when customers exceed the number of available servers. However, some customers may not wait for the service when the total number of clients to be served the number of serving facilities available. Therefore, a queue can be defined as a line best explained by the point of customers from which they can form the service point.

In most cases, waiting time in a line is a challenge to the financial sector, and managers should seek to improve their mode of operation to satisfy the demand of customers wholly. Moreover, waiting time increases cost directly as it requires more space hence additional cost to the management. Regardless of the introduction of ATMs in banks, still the problem of reducing waiting time upholds. Specifically, ATMs are adopted to improve the services and allow efficiency at the shortest time

possible to its users. Therefore, the business sector like banks tries their level best to improve its services by minimizing service time and give a customer better experience. This study will assist managers to solve the underlying problem such that they can achieve quality, maximum and timely service delivery to their clients.

Queueing system

According to Law (1974), a Queueing system is described by the service mechanism, arrival or input pattern of customers, and customers' behavior. Primarily, Queueing systems comprise of one or more servers that offer services to the arriving clients. Further, Queueing systems may contain a finite or infinite population in their queues. Due to a limited number of servers and different velocity of money circulation, customers arrive at the service facility in a random fashion (International Annual Conference of the German Operations Research Society, & Lubbock, 2016). Therefore, a queue can be classified as having a limited or unlimited capacity as the case of a banking queue with infinite queue capacity length.

Queueing Theory

The Queueing theory can be borrowed back to Copenhagen Telephone Company in 1908 which requested Erlang to work on the keep up times in a telephone switch (Aniekan & Udoanya, 2014). In this study, Agner realized that the number of phone calls followed a Poisson distribution and were exponentially distributed. Thus, Queueing theory is a systematic study of mathematical models in Queueing systems which are used to evaluate the service processes especially the one with random variable properties in demand and service times (Lavenberg, 1974). Queueing systems can be handled in such a manner where there can be a well-defined model such as First-Come-First-Served (FCFS) discipline (Harel & Mouche, 2013). However, the criteria chosen will significantly affect the outcome of the line itself. In most cases, a system may have one or two servers to provide services to customers. The arrival of clients at service center takes a random fashion depending the time of demand for money and the velocity of money circulation. Thus, a bank ATM is an appropriate example to consider for this study.

Mathematical Queueing models consist of; arrival time (λ), service time (S), some clients in the system (Q) and waiting time (W) and they assist in analyzing the problem of reducing waiting time while improving service time.

According to Little's theorem Little (1961), the following variables between service rates can be calculated by finding the expectation or mean in the following equations:

That is, at steady state, number of clients in a queue can be expressed as;

$$L = \lambda T \quad \text{.....I}$$

Where, λ is average arrival rate, and T is average service time in a system for a user.

From the steady state equation, three fundamental variables are derived i.e.

L , which is directly proportional to T and λ ,

λ , which increases when L is increased or T decreases, and

T increases when L increases or λ decreases. This tells that λ has a mutual relationship with L and T .

Thus, this calls for an ATM model ($M/M/1$ model) whereby arrival and service time follow a Poisson processes i.e. they are exponentially distributed. In this model, the following variables are required;

λ , as the arrival rate and μ as the service rate.

Implying that the utility function is given as β

$$= \lambda / \mu \quad \text{.....II}$$

The probability of having no customers in the ATM service is given by

$$P = 1 - \beta \quad \text{.....III}$$

However, the likelihood of n customers can be evaluated as;

$$P_n = P_0 \times \beta^n = (1 - \beta) \times \beta^n.$$

Therefore, L (average number of clients in the queue is expressed as;

$$L = \beta / (1 - \beta) = \lambda / \mu - \lambda \quad \text{.....I}$$

Also, the average number of clients in the queue are represented as;

$$L_q = L \times \beta = \beta^2 / (1 - \beta) = \lambda / \mu - \lambda \quad \text{.....V}$$

Equally, the waiting time, W_q

$$= Lq/\lambda = \beta/\mu-\lambda \dots\dots\dots VI$$

Therefore, based on these calculations, the researcher will solve the problem of long queues for the ATM facilities.

Computation Burden

With the question of how to reduce the waiting time in ATM services, it is important to consider the formulation and its solution to waiting time (Denizel, 2003). In practical and theoretical conditions, it can be hard to solve the problem of growth computational time as a function of problem size. For this reason, the issue of waiting time is an application of Queueing model to find ways on how the waiting time can be reduced and finally improve service delivery to the clients. In some cases, the Monte Carlo simulation techniques can also be used to give an approximate but a workable solution. This methodology has been applied to various problems involving stochastic processes and mathematical problems which require the application of Queueing model and where an application of mathematical techniques and experimentation is not real. Therefore the issue of waiting time in the banking industry can be evaluated by applying the Queueing model.

Methodology

Research Design

Setting up a research design depends on the researcher's orientation and research objectives (Piedmont & Village, 2010). There are different types of research designs which are longitudinal, case study and experimental designs. Further, these designs are divided into either fixed or flexible research designs. In this study, a case study has been used as the most appropriate research design since the phenomenon of ATM banking would not be appropriate without the interaction from with the bank clients. However, the researcher has applied a non- direct empirical method of data collection by observing and recording the number of clients visiting an ATM facility in one of the banks in town. The researcher uses primary data recorded in a period of one month in one of the banks in the city. Thus, the aim of this study is to evaluate the performance of an ATM service point. The number of customers arriving at the service point, service time, and average time a customer takes is analyzed in this study. The illustration obtained from the workings will assist in evaluating the performance of service mechanism. Also, the service time is recorded by seeking assistance from the previous analyses. The empirical study was carried for one month, and the following data was recorded:

Table 1: monthly daily customer data by observation of banking time.

Day week	Sunday	Monday	Tuesday	Wednesday	Thursday	Weekend	Weekend
						Friday	Saturday
1	201	165	140	125	117	90	88
2	175	150	130	110	74	65	62
3	145	125	119	98	120	83	70
4	120	102	95	95	90	85	65
Total	641	542	484	428	401	323	285

The results above indicates that the number of customers on Sunday is approximate twice the number of customers on weekends especially on Saturday. Generally, in weekdays the turn up is low as compared to weekends. Probably, this can be attributed to the fact that most customers do not work during the weekends.

Workings;

From the researcher's empirical analysis and consultation with the bank attendants, a customer spends approximately 2.5 minutes at the ATM (W), whereas the queue length is four people (Lq), and lastly the waiting time is 3 minutes. At the same time, on average, the number of clients visiting the

ATM facility is approximately 65 customers per hour. Thus, the average arrival rate λ , for the first two days after a weekend, can be evaluated as;

$$\lambda = 65/60 = 1.08 \text{ customers per minute} \approx 1 \text{ CPM}$$

Equally, from equation I,

$$\lambda = 5/25 = 2 \text{ customers}$$

Theoretically, waiting time

$$Wq = Lq / \lambda = 4 / 1.08 \approx 4.3 \text{ minutes}$$

$$L = 1.08 \times 2.5 = 2.7 \text{ customers} \approx 3 \text{ customers.}$$

$$Lq = L \times \beta = 1 \times 2.5 = 2.5 \text{ customers per minute.}$$

Equally,

$$\mu = \lambda (1 + L) / L$$

$$= 1.08 (1 + 2.7) / 2.7$$

$$= 1.48$$

Hence,

$$\beta (\text{utilization function}) = \lambda / \mu = 1.08 / 1.48 = 0.72$$

$$Wq = 0.72 \div (1.48 - 1.08) = 1.8 \text{ minutes.}$$

The average number of people in ATM can be expressed as;

$$L = \lambda T = 1.08 \text{ cpm} \times 2.5 = 2.7 \text{ customers} \approx 3 \text{ customers.}$$

The probability of zero customers

$$P = 1 - \beta = 1 - 0.72 = 0.28$$

This assists in finding the probability of n customers in the system. This can be expressed as;

$$P_n = P_0 \beta^n = (1 - \beta) \beta^n = (0.28) [(0.72)]^n \dots \dots \dots n^{\text{th}} \text{ customer}$$

According to Dhar & Tanzina Rahman (2013), applying the concept of balking, it can be assumed that an impatient customer will exit the queue once the system reaches three customers. Besides, if it is assumed that the maximum capacity of the system is ten clients or jobs in the queue, then the probability of having more than four customers in the system is given as;

$$P_{5-11} = \sum_{n=5}^{11} P^n = 0.28(0.72)^n = 15.42\%$$

Thus, a bank analyst can analyze the probability of a customer going away without getting the service and addressing the problem to offer quality and efficient services to the clients.

Analysis and usefulness of this study

From the calculations above, the utilization value is high especially on weekends. This study assists management of a bank ATM in improving the quality of service in rendering their service. The estimated number of customers joining and leaving the service center, banks can use this analysis as a reference to know the number of ATMs needed to solve the problem of long queues in a banking facility and thus guarantee their customers quality service.

Review of related studies

From prior studies, there have been various studies that have been done on simulation and Queueing systems that have been done other than on ATMs. These are discussed herein:

Guizani, & Wiley Inter-Science (2010), in their survey, they have considered that the approximation of the probabilities using the Queueing models has received considerable attention with the need for simulation and Queueing literature. In this study, the researcher has pointed out that most of the literature covered concentrates heuristically on the derivation changes of measure. As a result of the application of the Monte Carlo simulation techniques, the performance turns out positive in many but fails in some models. Equally, in other studies related to this, few simple models have been used whereby the researchers have not used the Queueing models to solve the challenge of reducing waiting time in different sectors that they based their studies. At the same time, adaptive methods that try to solve similar Queueing problems iteratively have not been applied (Crane and Iglehart, 1974).

In another study by White, Schmidt & Bennett (1975), the study recommends the use of simulation techniques as a viable tool for solving problems that are not analytically tractable that is simulation techniques are applied in analyzing numerically complex models. With this case, different iterative and adaptive methods are used which is different from the present method. Further, Thomopoulos (2013) indicates that moderate Queueing and simulation techniques have been applied with durations of 500 and 1000 tours used and the average simulation computer time to tour is 0.03 seconds for big computers. In this study, the model was subjected to saturated conditions with non-dependable replications for estimating a confidence interval for A. This is because, for regenerative simulation, this comes as the most preferred input level.

Further, according to Lavenberg (1974), the insertion or savings techniques that build a solution in a way that for each step of the methodology, a current configuration is done which is probably infeasible. To achieve the maximum results, an alternative procedure is done which applies some criterion such as the total cost or the one that uses the least expensive method function in the current configuration into the existing routes. More of these techniques have been addressed by Solomon (1987) and Clarke and Wright (1964).

According to Denzel (2003), the application of Lagrangian relaxation procedures of a customer in a banking hall and in front of an ATM facility. Besides, in trying to address the problem of reducing the waiting time. Denzel (2003) uses the interactive optimization as the collective purpose approach. More of these Queueing procedures are discussed by Krolak et al. (1970).

Besides, heuristic approaches such as Simulation Model (SA) have been previously studied. For instance, for the general Queueing problems. Winston & Goldberg (2004) had initially applied the location based heuristic approach. Profoundly known as the capacitated concentrator location problem, the method focuses on expressing the Queueing task as a position problem (Denzel, 2003). This study formed the basis for formulating the Queueing problems because it was solved and the results transformed to solving various problems including the Queueing problem.

Conclusion

It is necessary, and there is a need to eye for balance between service and economic consideration such that there could be no too many servers. This paper has captured the application of mathematical (Queueing model) to an ATM facility. This method is also applicable in evaluating daily data from ATM services. That is, with the high utilization time, the probability of having impatient customers leaving the queue without receiving service can be minimized. Therefore, there is no need for banking managers to increase ATMs to their facilities to serve a relative number of customers. Rather, they can do so by employing extra personnel which is relatively cheap compared to acquiring additional ATM service. By so doing, the owner can enjoy enormous economic profits while at the same time regulating queues which maximize service delivery to their clients. In turn, this will contribute to an improvement of banking services in the financial sector. However, the limitation of this study is that the conclusion has been reached without taking into consideration of the cost model of a system.

Recommendations

There are numerous features of waiting time for ATMs that have not been covered in this study and at the same time not yet covered in prior studies. Therefore, the topics can form interesting questions for subsequent studies. For instance, the time a customer stands and leaves a queue should be taken into consideration and from study subjects. Also, if a client is not conversant with how an ATM works, it will take more time while other customers wait in the queue and this has not been addressed in previous studies. Thus, it can equally form a study topic. At the same time, a case where there is the lack of cash stock in the facility can be taken into consideration and finally analysis of waiting times during peak hours and holidays can form excellent topics of study. The researcher advocates that additional of an extra ATM facility by management to improve service rate and reduce the waiting time is not the optimal solution. Rather, the bank managers can consider training the guards who man ATM facilities on how the service works such that they help those clients who don't know how to

handle an ATM. In this way, the cost of hiring new specialized personnel and the accounting cost of acquiring a new ATM can be reduced. Equally, banks should minimize their policies regarding maintenance so as to avoid some complexities which may arise due to the failure of server machines. This eventually will meet the demand of customers hence banking services will be improved.

References

- Davila, G. A. M., & Crowther, D. (2011). *Technology, Human dignity and Managerial Responsibility: Diversity, rights, and sustainability*. Farnham: Gower.
- Denizel, M. (April 01, 2003). Minimization of the Number of Tool Magazine Setups on Automated Machines: A Lagrangean Decomposition Approach. *Operations Research*, 51, 2, 309-320.
- Guizani, M., & Wiley Inter-Science (Online service). (2010). *Network modeling and simulation: A practical perspective*. Chichester, West Sussex. U.K: Wiley.
- Harel, M.-A., & Mouche, E. (December 01, 2013). 1-D steady state runoff production in light of Queueing theory: Heterogeneity, connectivity, and scale. *Water Resources Research*, 49, 12, 7973-7991.
- International Annual Conference of the German Operations Research Society, & In Lubbock, M. (2016). *Operations research proceedings 2014: Selected Papers of the Annual International Conference of the German Operations Research Society (GOR), RWTH Aachen University, Germany. September 2-5, 2014*.
- Krolark, P., Felts, W., and Marble, J., (1970). "A man-machine approach toward solving the generalized truck dispatching problem," *Transportation Science*, Vol.6, pp 149-170.
- Law A. M. (1974), "Efficient Estimators for Simulated Queueing Systems," ORC 74-7, Operations Research Center, University of California, Berkeley, California.
- Little, J. D. C. (June 01, 1961). A Proof for the Queueing Formula: $L = \lambda W$. *Operations Research*, 9, 3, 383-387.
- Crane M. A., and Iglehart D. L., (1974). "Simulating Stable Stochastic Systems, I: General Multiserver Queues," *J. ACM* 21, pp 103.
- Piedmont, R. L., & Village, A. (2010). *Research in the social scientific study of religion: Volume 21*. Leiden: Brill.
- Dhar S. K., & Tanzina Rahman. (2013). <http://iosrjournals.org/iosr-jm/papers/Vol7-issue1/A0710105.pdf>
- Lavenberg S. S., and Shedler G. S., (1975), "Derivation of Confidence Intervals for Work-Rate Estimators in a Closed Queueing Network," *SIAM J. Comput.* 4, pp 108.
- Lavenberg S. S., (1974), "Efficient Estimation via Simulation of Work-Rates in Closed Queueing Networks," *Proceedings of Computational Statistics*, Physica-Verlag, Vienna, Austria, pp 353-362.
- Thomopoulos, N. T. (2013). *Essentials of Monte Carlo Simulation: Statistical methods for building simulation models*. New York: Springer.
- Udoanya Raymond Manuel, & Aniekan Offiong. (April 01, 2014). Application of Queueing Theory to Automated Teller Machine (ATM) Facilities Using Monte Carlo Simulation. *International Journal of Engineering Science and Technology*, 6, 4, 162-169.
- Vieira, F. H. T., & Bozinis, G. E. (January 01, 2010). *Providing Quality of Service to Computer Networks through Traffic Modeling*.
- White, J. A., Schmidt, J. W., & Bennett, G. K. (1975). *Analysis of Queueing systems*. New York: Academic Press.
- Winston, W. L., & Goldberg, J. B. (2004). *Operations Research: Applications and Algorithms*. Belmont, CA: Thomson/Brooks/Cole.