

An Application of Queuing Theory to ATM Service Optimization: A Case Study

Abdul-Wahab Nawusu Yakubu* Dr. Ussiph Najim
Department of Computer Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
* E-mail of the corresponding author: mryawn@yahoo.com

Abstract

Unmanaged queues are detrimental to the gainful operation of service systems and results in a lot of other managerial problems. This paper applies queuing theory to determine optimal service level for a case ATM base on a customer-defined criterion of wait time not exceeding eight (8) minutes. In pursuance of this, the prevailing operation characteristics of the case ATM as a queuing system where defined. Direct non-participatory observation and questionnaire were engaged to record time measurements and primary data. Measurements were taken on arrival times and service times of customers who arrived at the terminal within the hours of 8:00 am to 4:00 pm. The Chi-squared Goodness of Fit test was performed on collected data. This established the interarrival times at the case ATM as exponentially distributed. The *M/M/s* queuing model therefore best illustrates the ATM queuing system of the case bank. A queuing theory based decision support system was developed as a result and applied to analyse and suggest improvement in waiting time. Two ATMs at a service rate of 0.60 customers per minute is found to be optimal for the case bank albeit waiting time are found to be relatively higher during the hours of 11:00am to 1:00pm and month endings. The research thus reveals that although queuing theory is applicable in finding optimal service levels, waiting time might still be lengthy because of external factors. Service unavailability was observed to be a contributory factor to queue formation at the case ATM. A routine maintenance regime should be actively implemented in curtailing such problems. For short term queue management however, backup-staffs could be engaged during peak periods to handle any additional demand instead of the alternative of installing the rather capital intensive ATM which might be of less utility for most business hours. Queue management should also be made an active part of the bank's overall strategic queue management processes.

Keywords: Queuing Theory, Waiting Time, Service Rate, Arrival Rate, ATM, Optimal Service Level

1. Introduction

Bankley *et al.* (2006) cited in Musara and Fatoki (2010) opine that technology provides enhanced insight into handling old and new tasks. Technology has changed not only the way we do business but has also changed virtually all sphere of human life. Abor (2005) affirms that technology affects even the direction of an economy and its capacity for continued growth. Human beings consciously or unconsciously interact with products of technology in almost all their daily activities. These products have made the performance of activities which hitherto were carried our stressfully and unproductively much more convenient, faster, easier and more accurate. In the banking industry, information and communication technology is playing a major role in addressing operational challenges such as quicker exchange of data, information processing, record storage and retrieval and many more. The Automated Teller Machine (ATM) is one of the several electronic banking channels used in the banking industry of Ghana. According to Aldajani and Alfares (2009), automated teller machines are among the most important service facilities in the banking industry.

In Ghana, ATM as a banking instrument has enjoyed widespread acceptance and usage. In a survey conducted by Abor (2005), more than half of respondents revealed their preference for ATM as a conduit to conducting transactions. ATMs particularly when installed off-site serve to keep customers away from bank halls. Unfortunately most ATMs in Ghana are on-site operating very much as another department of the bank. The issue of queue control in bank halls via ATM is essentially defeated because ultimately access to these facilities is limited to customers going to the bank. ATMs themselves have as a result become subjects of large service demands which directly translate to queues for services when these demands cannot be quickly satisfied. This situation becomes compounded and more evident during festive periods and month endings, around which time demand for cash is high.

Queuing theory known by various other names such as the traffic theory, congestion theory and the theory of mass service (Copper, 1981) is a mathematically based technique for analysing waiting lines (queues). In queuing theory, queuing models are used to approximate a real queuing situation or system so that the queuing behaviour can be analysed mathematically. It has been used successfully in the studies of queue behaviour problems, optimization problems and the statistical inference of queuing system (Xiao and Zhang, 2009). Queues abound. There is queuing at petrol or gas or diesel filing stations if too many customers await service. Vehicles waiting to cross a major road create congestion. There is also queuing in bank halls when customers wait to be attended to by a cashier or to utilize an ATM facility and many other similar situations. A queue is synonymous to a waiting line because of the waiting involved. According to Ford (1980) "Waiting lines develop when "clients" arriving for "service" are delayed prior to being served". If customers are scheduled to visit service facilities, and the scheduling rule strictly adhered to, queues can be avoided. Nevertheless, in reality this is not the case as most of the time customers arrive at these service facilities in a random and uncontrolled manner. Arriving customers who meet a busy server and/or a waiting line of customers, either departs or waits in the queue for his or her turn during which time the customer holds on to the server. After service is completed, customers are generally assumed to leave the system making it available for other customers. Unmanaged queues are detrimental to the gainful operation of service systems and results in a lot of other managerial problems. For instance an ATM that receives and accommodates huge inflow of customers can be detrimental to its smooth running and response time. A slow response would greatly affect the speed at which service is provided to customers. As a results service providers may loss customers who grow impatient and leave the system. The business order now is the ability to acquire and retain customers. This centrally lies in the ability of firms to satisfy and provide better service experiences. As such managers of service systems need to design apt strategies to tackle challenges brought about as a result of lengthy queues. This paper considers two on-site ATMs at the Agricultural Development Bank (ADB) Tamale Main Brach as a means of gathering data and understanding the real issues on the grounds regarding identifying an optimal ATM service level in the mist of queues and service outage.

2. Methodology

2.1 Research Design

Depending on research questions and orientation of the researcher, a choice is made in setting out the research plan. There is experimental design, longitudinal design, cross-sectional design and case study design. Robson (1993) divides these designs into fixed and flexible research designs. Others have referred to this distinction as quantitative research designs and qualitative research designs respectively.

The case study design is applied appropriately in this paper. Case study as an empirical inquiry is chosen because it allows focus to be placed on the queue phenomenon within its real-life context. A case study design was chosen also because the topical issue was customer queuing at ATMs; this could not be considered more perceptively without the context, the banking environment and more specifically the settings created by the ATM-customer interaction. It was in these settings that data was gathered and utilized. It would have been impossible for the researcher to have a true picture of customer queuing without considering the context within which it occurred. The case study approach thus offered us the opportunity for relevant aspects of queuing problem to be studied in some depth within a limited time scale. Moreover, it allowed us to cover contextual conditions relevant to the phenomenon under study. The design here is particularly a single case where we considered the two ATMs of the main branch bank of ADB in Tamale. This appropriately provides us the needed environment to glean required data for analysis.

2.2 Data Collection

The first step to structuring collected data for analysis is how that data would be captured in the first place (Kantner *et al.*, 2005). In this paper, two types to data were collected and used. These are primary data and secondary data. Secondary data was obtained through an intensive review of relevant literature on the queuing problem from journal articles, textbooks and many usable electronic sources. Primary data was collected in two periods in a month long field studies via observation and questionnaire. The first three weeks was devoted to time studies which involved recording on an observation form three key required quantities (customer arrival time, service start time and service completion time) as customers reach at and leave the ATM terminal. In the second period of data collection, we utilized a closed-ended structured questionnaire to gather data on customers' point of view about service quality in their interaction with the ATMs and in particular their definition of an acceptable time to wait in queue.

2.3 Sampling Procedure and Population

Convenient sampling as a nonprobability sampling technique was utilized in administering research questions. This technique was employed not only because respondent were necessarily easy to recruit, but also because of their convenient accessibility. Participants were the most willing and available to be engaged at the material moment. By so doing data was collected faster, easier and inexpensively without requiring the practical details of using randomize sampling techniques.

ATM cardholders and users were qualified for selection for observation and questioning. This qualification ensured that the participants understood firstly the nature of the questionnaire and its relevance to bettering the current situation they obviously are part of. The questionnaires were administered to two groups of customers, using the self-administered and face-to- face techniques. Two groups of customers are considered. The first group comprised of customers who could read and write and opted to fill the questionnaire out by themselves. The second and other group comprised of customers who could not read or write and those who simply preferred that questions be read out to them whiles they respond. For the first group, the questionnaires were self-administered and immediately return at site. On average customers in this group responded quickly without requiring any clarification or having queries about it. This may show that the questions administered were not associated with bad design or missing responses. We administered questionnaires to the other group members in a face-to-face fashion. We read out questions and best-fit choices and answers made by individual respondents in this group.

Participants for the time studies were also ATM users who arrived at the ATM terminal between the hours of 8:00 AM and 4:00 PM each of the study period excluding Sundays. This design was done purposively to meet the desired objectives. Since the bank had many repeat customers we ensured that no customer was allowed to participate more than once in the administration of questionnaires. However, several repeat customers might have been captured in the time studies. This we do not consider in our analysis and do not believe would have a significant effect on results.

3. Analysis and Results

3.1 Extraction of Time Measurements

The following procedure is applied to extract needed information from time measurements. Upon arrival at the ATM terminal, customers are indexed for identification by $i = 1, 2, 3, ..., n$; each associated with the following time measures.

- The arrival time of customer of customer i.
- The service start time of customer i.
- The service completion time of customer i.
- The inter-arrival time between customers i and $i+1$.
- The wait time of customer i in the queue
- The service time of customer i.
- The wait time of customer i in system

Averages for the interarrival times and service times within each hour of the eight hour period per day are taken. These are then used to obtain the mean operational characteristics of the case ATM queuing system by further applying the appropriate M/M/s queuing model equations.

3.2 Analysis of Questionnaire Data

Data gathered from questionnaire were coded and then analyzed using the Statistical Package for Social Scientist (SPSS) and Easy Fit. These were utilized to present interpretation from findings and as well perform test of hypothesizes on distribution of service times and interarrival times respectively. Each questionnaire question and the corresponding 106 responses were entered in SPSS and descriptive statistics in the form of frequencies and simple percentages are used to represent questionnaire results. Means and standard deviations are also computed to provide additional elaboration on data.

3.3 The ATM Queuing Model

Based on observations and Chi-Square Goodness of fit test on collected time measurements, we conclude that the case bank's ATM system as a queuing system is best illustrated as an *M/M/s* queuing model. It specifies a multi-server queuing mathematical model where the *M* signifies exponential interarrival times and exponential service times and *s* represent the number of facilities (ATMs) in the system. Customers arrive at the service system at random points in time to seek service also in a random manner. The service system operates in such a manner that for each arrival, if both ATMs are busy then the customer enters a queue; else the arriving customer immediately enters service. As customers depart, one or both ATMs becomes idle, else a customer is selected from the queue to enter service. The *M/M/s* model is one of the most commonly used queuing models to analyse queuing problems involving multiple channel systems servicing a single unlimited waiting line of customers in a FCFS fashion. The model enables computation of performance measures such as average customer wait times (in queue and in system) and queue lengths (in queue and in system) given arrival rates, service rates and number of servers.

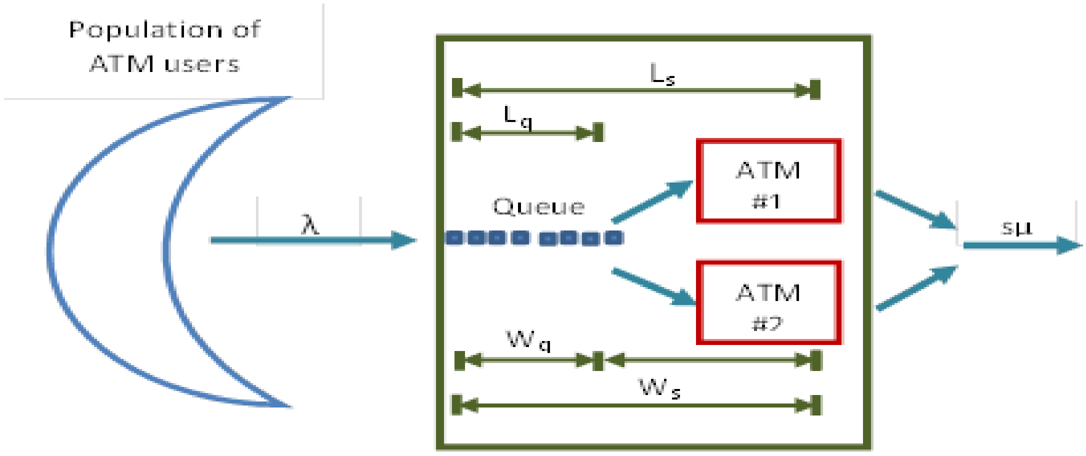

Figure 1. A Conceptual Model of the Case ATM queuing system

3.3.1 Model Assumptions

The model above is implemented with the following assumption implemented.

- Service times are exponentially distributed.
- Identical service facilities (since same kind of transactions are performed on both ATM).
- No customer leaved the queue without being served.
- A queue with unlimited waiting space that feeds into s identical servers.
- Customers are served on FCFS (First Come, First Serve) basis.

3.3.2 Modelling Arrival Rate and Service Rate

Determination of a suitable model for a system requires better understand and knowledge of features of the system such as arrival pattern of customers, service distribution, queue arrangement and number of servers. In this respect, we observed that the pattern of arrival of customers to the case ATMs varied. On most occasions customers entered singly. Again, arrivals to the ATMs did not form any regular pattern since customers did not arrive at constant intervals. Furthermore, the pattern of arrival of customers was unscheduled as customers entered without any ordering and the arrival did not follow any specific pre-defined plans. Moreover, the distribution of arrivals was random and inter-arrival times were not constant. Inter-arrival times most commonly fall into one of the following distribution patterns: a Poisson distribution, a deterministic distribution, or a general distribution Dharmawirya and Adi (2011). Service time distribution on the other hand can be constant, exponential, hyperexponential, hypo-exponential or general. Data from observation lends credence to the fact that interarrival times are exponentially distributed. This is buttressed by Palm (1943) and Khinchin (1960) cited in Creemers and Lambrecht (n.d.) who expressed that the arrival processes of customers to a service facility will tend to a Poisson process. Lariviere and Van Mieghem (2004) affirm that the specification of interarrival times as exponential is accurate for many service systems.

3.3.3 Determining Performance Measures of the ATM Queue Model

This paper is most interested in the customer waiting time. There are two types of customer waiting times; the time a customer spends in the queue and the total time a customer spends in the system. Since we are dealing with human beings, we are specifically concerned with customer waiting time in queue. It is waiting in queue that is dissatisfactory to customers and affects greatly their service experience. Other interesting areas of ATM performance are related to the ATM activity and include the percentage of time they may be busy or idle. The following performance measures of the system would be computed:

- λ: The mean customer arrival rate per minute.
- μ: The mean customer service rate per minute.
- ρ: Traffic intensity; the probability that the ATMs are busy at random time (t) within the interval.
- L: Average number of customers in the system per time period.
- Lq: Average number in the queue per time period.
- W: Average time spent at the terminal, including the waiting time in queue.
- Wq: Average customer waiting time in the queue.

3.4 Questionnaire Results

Table 1. Summary statistics of questionnaire responses

QUESTION DETAILS	RESPONDS (N=106)	PERCENTAGE (%)
Where/How do you prefer to perform bank transaction?		
Banking Hall	9	8.5
ATM	97	91.5
Which of the following operations do you often perform at the ATM?		
Cash Withdrawal	43	40.6
Balance Enquiry	1	0.9
Balance Enquiry Then Withdrawal	60	56.6
Pin Change	1	0.9
Others	1	0.9
On average, how much cash do you withdraw from an ATM transaction?		
Less than GH¢100	6	5.7
GH¢100 – GH¢200	33	31.1
GH¢210 – GH¢300	36	34.0
GH¢310 – GH¢400	23	21.7
More than GH¢400	8	7.5
What is of most concern to you in your use of the ATM for business transaction?		
Queues/Waiting line	65	61.3
Service unavailability	34	32.1
High service charge	6	5.7
Others	1	0.9
How much time in queue do you consider acceptable?		
1 – 5 minutes	56	52.8
6 – 10 minutes	26	24.5
11 – 15 minutes	11	10.4
16 – 20 minutes	4	3.8
21 – 25 minutes	0	0

QUESTION DETAILS	RESPONDS (N=106)	PERCENTAGE (%)
26 – 30 minutes	9	8.5
> 30	0	0.0
In general, how would you rate the quality of ATM service provided by the bank?		
Highly satisfactory		
Satisfactory	3	2.8
Fairly satisfactory	27	25.5
Not satisfactory	60	56.6
	16	15.1

On average, customers felt that approximately eight (8) minutes is an acceptable time to wait in queue. However, as acceptable waiting responses were based on individual perspectives, description of what constitutes an acceptable waiting time tend to be around the interval one to five minutes. Fifty two per cent (52%) felt that a waiting time of one to five minutes was a reasonable time to wait. Beyond the acceptable waiting time threshold there is no guarantee of customers been patient to wait and perform transactions of particular interest. A customer may renege after this time period; accruing to the bank as lost profit from lost business. This loss of business may occur immediately because the customer grows impatient and leaves or in the future, because the customer is sufficiently irritated that he or she does not come again.

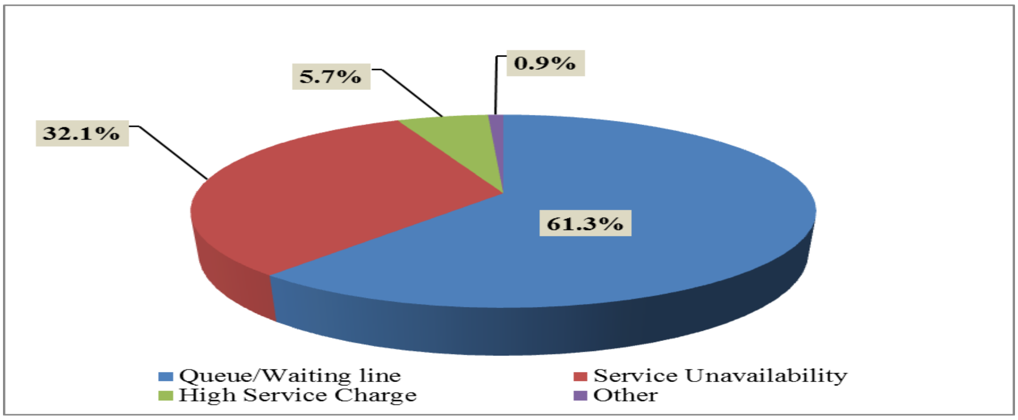

Figure 2. Customers concerns in the use of the ATM for transaction performance

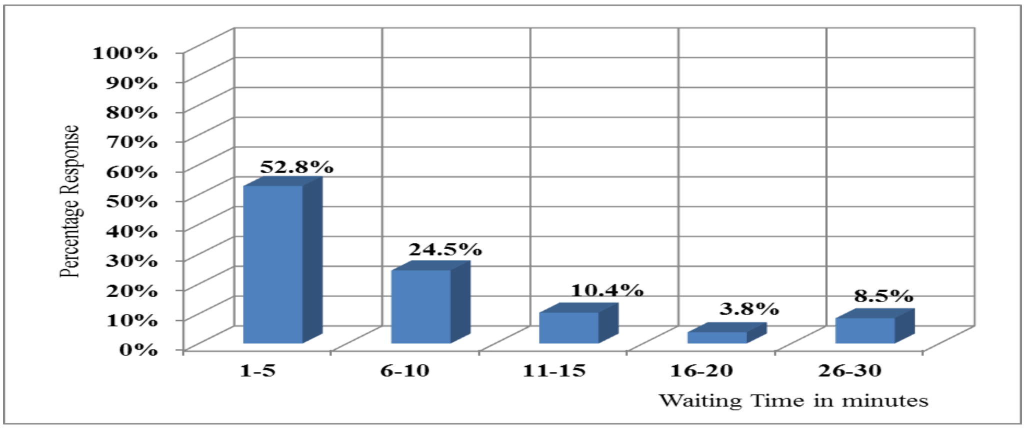

Figure 3. Customers definition of acceptable waiting time

3.5 Optimization of the ATM Service Level

The task of queuing system optimization is subject to clearly stated criteria defined usually by the management in charge. Mostly this decision is made with particular attention paid to the cost of operating the system with the objective of meeting the stated criterion or criteria at the possible lowest cost. Sensitivity test proved that number of ATMs, arrival rate and service rate are factors that influence the length of time a customer waits in queue and the performance of the system in general. In this paper we look at optimization in the following areas with the view to reducing waiting time.

1. Determine optimum number of ATMs to maintain at reduced wait time.
2. Determine the optimum service rate at which to operate with existing capacity and
3. Determine a favourable queue arrangement or order.

3.5.1 Cost Model for the ATM Queuing System

We define the total cost for the ATM queuing system with *s* ATMs as the sum of the operating cost of running the ATMs and the wait cost as customers wait in queue. That is;

$$TC = TC_s + TC_w \qquad (1.0)$$

Where;

TC = total cost for operating the ATM queuing system with s ATMs.

TC_s = total operating cost of s ATMs at the case bank.

TC_w = total cost of customer wait in queue.

The total operating cost of the s ATMs (TC_s) is equal to the cost of running an ATM (C_s) per unit time multiplied by the number of *ATMs* (s).

Let TC_{ATM1} = cost of running ATM_1; whether it is busy or not. Then;

$$TC_{ATM1} = C_s$$

Similarly let TC_{ATM2} = cost of running ATM_2; whether it is busy or not. Then as above;

$$TC_{ATM2} = C_s$$

From the preceding;

$$TC_s = TC_{ATM1} + TC_{ATM2}$$

$$\Rightarrow TC_s = C_s + C_s$$

But since the ATMs are similar as they provide same services then;

$$TC_{ATM1} = TC_{ATM2}$$

Therefore

$$TC_S = s \times C_s \qquad (1.1)$$

Cost of operating an ATM (C_s) may include cost of labour to fix mechanical faults, cost of labour to periodically stock the machines with cash and the cost of installation of additional ATMs. In the same way, total cost of customer wait in queue at the case ATMs (TC_w) is equal to the wait cost (C_w) of customers in queue per unit time multiplied by the average number of customers that are present in queue (L_q) per unit time. Wait cost may refer to the cost of losing customers (not keeping them in queue) who are unhappy and are less likely to return due to long wait in queue. The average daily cash withdrawal at the ATM if known could be used as the waiting cost value. The total cost of customer wait in queue at the case ATMs can therefore be expressed as;

$$TC_w = L_q \times C_w \qquad (1.2)$$

From equation 1.1 and 1.2, equation 1.0 becomes;

$$TC = (s \times C_s) \times (L_q \times C_w) \qquad (1.3)$$

Equation 1.3 would be used for all cost calculation.

3.5.2 Optimal number of ATMs

We define the optimal number of ATMs as the number of ATMs that ought to be setup to achieve basically two objectives of reduced customer wait time and reduce cost of providing services. The challenge therefore is determining whether or not cost would be less by maintaining, reducing or increasing current capacity. The case service system is a multi-server system consisting of two automated teller machines with a single line of customers. We investigate below for the optimal number of ATMs using sample busy period and less busy period data. The cost values of running the system at current capacity are firstly computed. The operating cost is then determined by varying the number of ATMs (s). Example cost elements are C_s = GH¢1000 and C_w = GH¢500. The flowchart that handles this sub-task of our developed decision support system is depicted in figure. 4. The flowchart yields the optimal number of ATMs based on the current wait and the desired wait in queue.

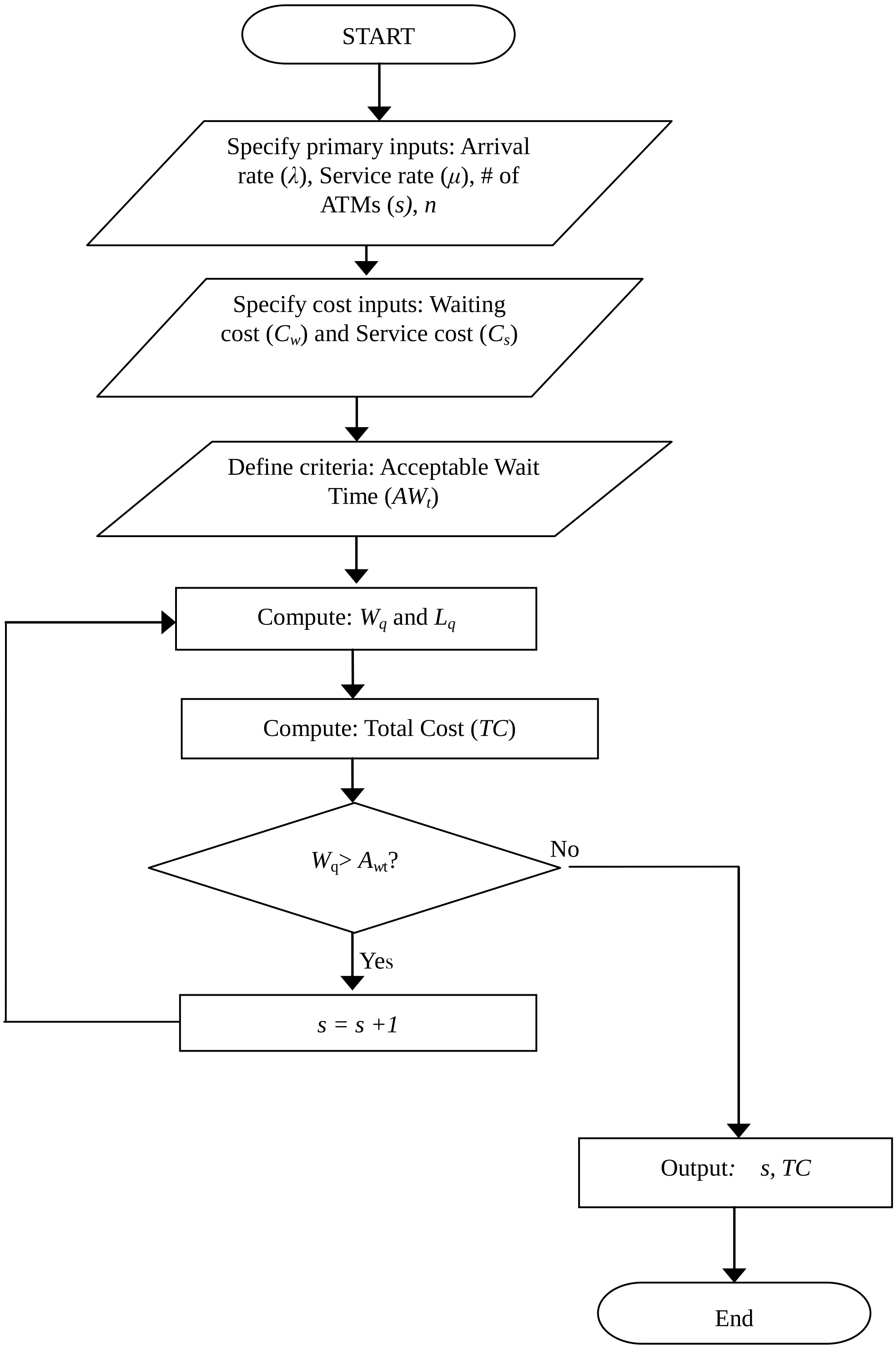

Figure 4. Java program flow for determining optimal number of ATMs based on wait time

Table 2. Optimal number of ATMs for sample busy period (arrival rate = 0.99 service rate = 0.51)

No. of ATMs	Total System Cost	Queue Length (L_q)	Waiting Time (W_q)	ATM Utilization(ρ)
1[1]	-	-	-	-
2	GH₵ 17775.68	31.551	31.870 Min.	97.06 %
3	GH₵ 3382.43	0.765	0.773 Min.	64.71 %
4	GH₵ 4075.33	0.151	0.152 Min.	48.53 %
5	GH₵ 5016.99	0.034	0.034 Min.	38.82 %
6	GH₵ 6003.77	0.008	0.008 Min.	32.35 %
7	GH₵ 7000.79	0.002	0.002 Min.	27.73 %
8	GH₵ 8000.15	0.000	0.000 Min.	24.26 %
9	GH₵ 9000.03	0.000	0.000 Min.	21.57 %
10	GH₵ 10000.00	0.000	0.000 Min.	19.41 %

From table 2: we observe that with existing capacity of two ATM machines, the total cost of running the system is GH₵ 17775.68. Customers at this level wait on average of approximately 32 minute to get the chance of performing transactions. However, if an addition ATM is installed the total operating cost of the system decreases to GH₵ 3382.43. This is GH₵ 14393.25 less compared to when two ATMs. The waiting time in queue reduces also with additional ATMs. The minimum total system cost is therefore achieved by running three ATM machines. It is evident therefore that there is considerable cost saving in running three ATMs at current arrival rate and service rate instead of two or more than three. Therefore we arrive at a conclusion that the cost for operating three ATMs for this scenario is relatively cheaper while at the same time desired reduced wait time in queue is met. More importantly at capacity of three ATMs, customers wait time is only at a tolerable 1 minute. Utilization moreover is at an appreciable level of approximately 65% and diminishes unfavorably with additional ATMs. Therefore to meet demands of customers for this scenario three ATMs need to be setup which is the optimal.

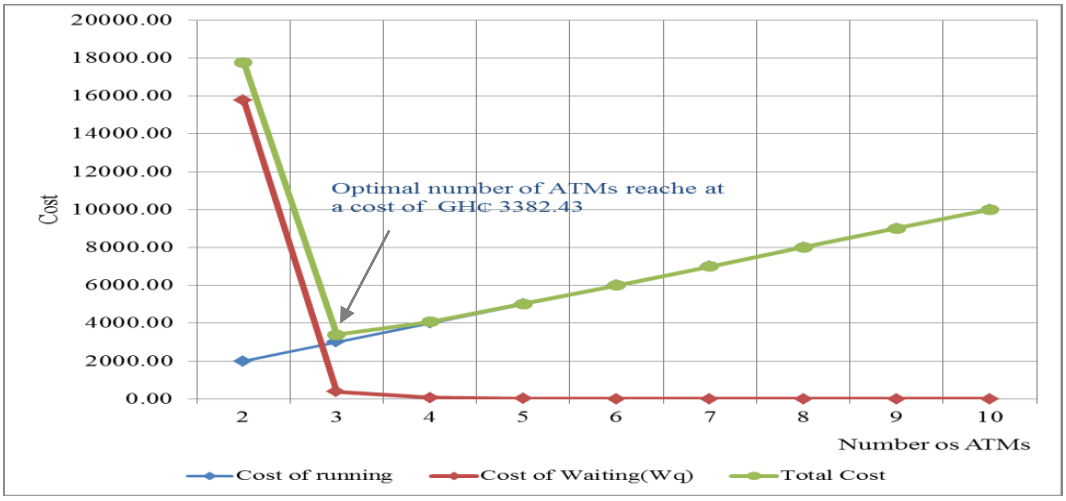

Figure 5. Optimal number of ATMs for sample less busy period

3.5.3 Optimal Service Rate

Figure 5, depicts the program flow for yielding optimal service rate that will minimize current wait time in queue to a desired wait time showing cost at each service rate. Table 3 shows results using sample busy period data from the case bank at various service capacities. At the present capacity of s = 2 ATMs, the optimal service rate (μ^*) should be equal to 0.55 to meet the desired reduction of customer wait time in queue. This yields a minimized system cost of GH₵ 5836.84 compared to GH₵ 17775.68 for service rate at the present 0.51. Management may also instead of operating at the optimal service rate of 0.55, maintain the current rate of service of 0.51 but increase the number of ATMs to three (3). If this decision is taken, a further decrease in system cost is achieved from GH₵ 5836.84 to GH₵ 3382.43. Management however would have to be contended

[1]* Queue grows beyond bound if 1 ATM is presented at the current arrival rate.

with a reduction in system utilization; from 90.00% to 64.71%. The optimal service rate for $s > 3$ is 0.51. Overall system cost increases with additional ATMs beyond 3. In this light system utilization severely deteriorates. Therefore customer service rate for this sample period should be 0.55 customers per minute to keep customers wait time at the centre less than 8 minutes and as well ensure a substantial level of utility of the facilities.

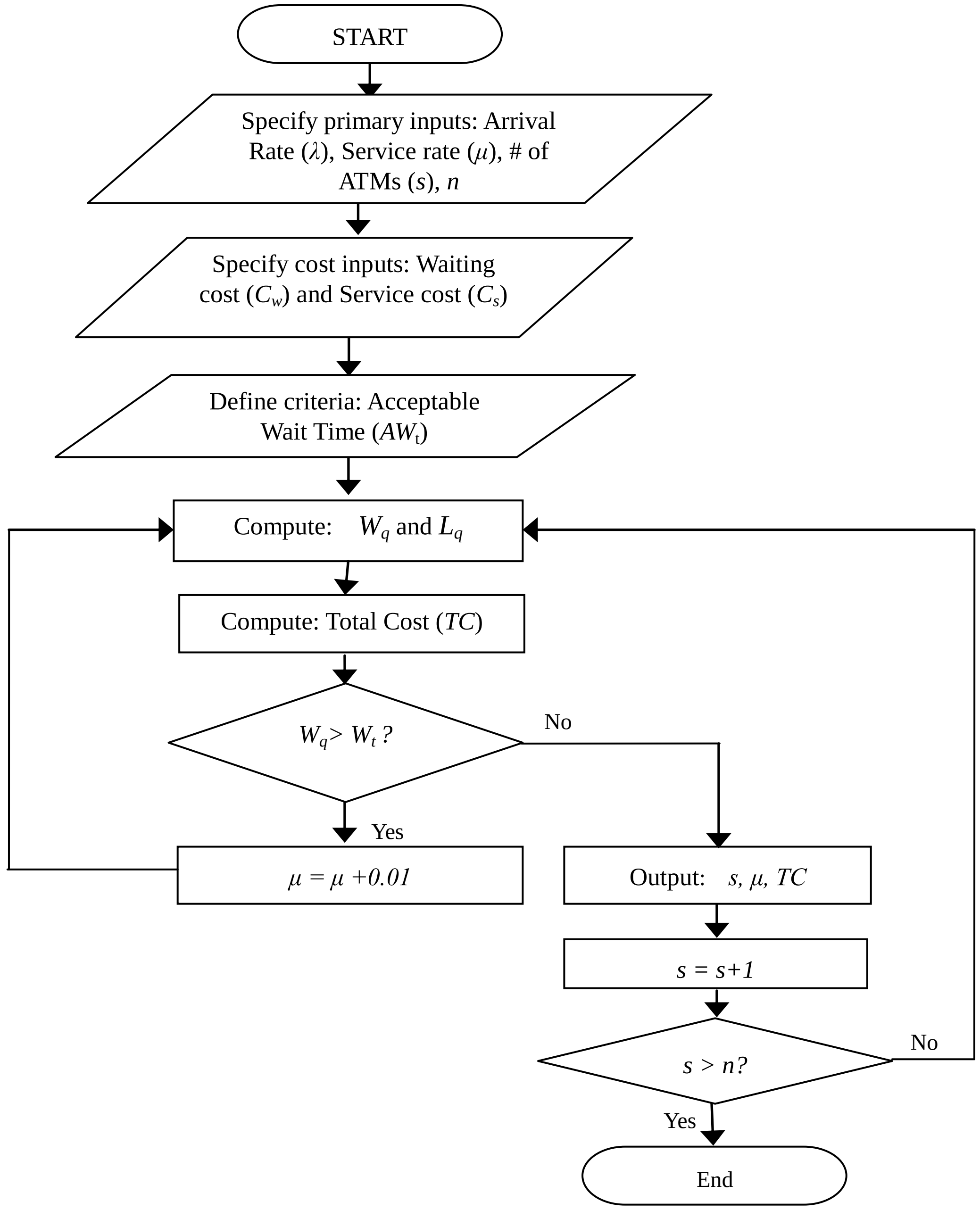

Figure 6. Java program flow for determining optimal service rate for the ATM queuing system

Table 3. Optimal table of service rates for sample busy period (arrival rate = 0.99, service rate = 0.51)

ATMs	Service Rate (**μ**)	Optimal Service Rate (**μ***)	Wait time (W_q)	System Cost	%Utilization (**ρ**)
2	0.51	0.55	7.75 Min.	GH₵ 5836.84	90.00 %
3	0.51	0.51	0.77 Min.	GH₵ 3382.43	64.71 %
4	0.51	0.51	0.15 Min.	GH₵ 4075.33	48.53 %
5	0.51	0.51	0.03 Min.	GH₵ 5016.99	38.82 %
6	0.51	0.51	0.01 Min.	GH₵ 6003.77	32.35 %
7	0.51	0.51	0.00 Min.	GH₵ 7000.79	27.73 %
8	0.51	0.51	0.00 Min.	GH₵ 8000.15	24.26 %
9	0.51	0.51	0.00 Min.	GH₵ 9000.03	21.57 %
10	0.51	0.51	0.00 Min.	GH₵ 10000.0	3.41 %

3. 6 Choice of Queuing Order at the ATM

The order of dispatching of customers from queue to the ATM is an essential indicator that impacts the performance of the system. Since we are dealing with human beings the natural transition order from queue to service point is on a first come first serve (FCFS) basis. The first come first serve queuing discipline is mostly implemented to ensure fairness in service delivery. In a multi-server system such as the one dealt with in this paper, this is not be exactly true. Customers may choose to form a single possibly long queue and transit one after the other for service in the order in which they arrived or they may also choose to create multiple possibly short lines in front of each ATM and receive serve in order of arrival within each queue. The FCFS rule may as well not work because of the high possibility of some customers finishing service far ahead of those who arrived first at the centre. In such instances unevenness may result in service provision which could lead to unhappy customers. We observed from the time studies that both arrangements were practiced at the case bank. The question therefore is which queuing arrangement is better in term of ensuring reduced wait time in queue. Is it a single queue feeding into either ATMs or two separate queues each feeding specifically into each ATM? Following below is an assessment of which arrangement ultimately yields better performance in terms of achieving reduced wait in queue. We consider an arrival rate of 0.99 per minute and service rate of 0.51 per minute of Table 3 above.

Table 5. Comparison of queue system performance for single queue and multiple queues arrangements

Queue Indicator	Queuing Arrangement	
	Single Queue	Multiple Queue
Queue Length (L_q)	31.55	32.03
Waiting Time (W_q) per minutes	31.87	64.71
ATM Utilization (ρ)	97.06%	97.06%
Probability of Waiting	0.96	0.97

Comparison of queue indicators from table 5 for both queue arrangement shows that multiple queues – two lines - generated longer waiting time at the case ATM and consequently results in poor system performance. Therefore, under prevailing conditions at the case bank and on a first come first serve basis, a system that adopts a single line would perform better than one in which multiple lines are formed and fed strictly into each available ATM.

4. Discussion

Results from analysis reveals that approximately ninety two per cent (92%) of respondents prefer to perform cash withdrawal at the case ATM terminals. Abor (2010) produces similar trend with more than half of sampled respondents preferring ATM services to other bank delivery channels. Various and varying reasons are given for this preference. Ease of use, convenience and availability during weekends are more prominent reasons for the preference of ATM as a point of transaction performance. Survey data demonstrates however that, customers in their interactions with the ATM (s) are most concern about meeting a lengthy waiting line of customers. This concern comes about in large part because of the massive preference for the service facility since there is an easy translation of this preference to lengthy waiting time especially if demands for services cannot be quickly satisfied.

On average, customers felt that approximately eight (8) minutes was an acceptable time to wait. However, as acceptable waiting responses were based on individual perspectives, description of what constitutes an acceptable waiting time tend to be around the interval one to five minutes. Over fifty two per cent (52%) of customers specified at least five minutes in their definition of an acceptable waiting time. Beyond this waiting time threshold, there is no guarantee of customers been patient to wait and perform transactions of particular interest. A customer may renege after the 8 minutes; accruing to the bank as lost profit from lost business. This loss of business may occur immediately (because the customer grows impatient and leaves) or in the future (because the customer is sufficiently irritated that he or she does not come again).

The optimal service rate and the optimal number of ATM for the case bank is 0.60 and 2 respectively. These strikingly are same as the current service rate and current capacity provided for service at the case bank. As a result, the evidence available suggest no need in improving the current rate and capacity, noting that order aspects of the ATM service availability must be ensured so that capacity does not fall below the current to warrant congestion. A major weakness observed about the case ATM is the common phenomena of service unavailability arising because of multiple reasons such as unavailability of cash, delay in loading cash and malfunctioning in mechanical parts. Service unavailability was so extensive during the period of observation that it rendered the system unusable during some periods of demand. This placed huge demands on the ATMs when service suddenly became available. Vasumathi and Dhanavanthan (2010) observed that customer preferred ATMs that are relatively stable in functioning than those that frequently were faced with "Out of Service Problem". The case ATMs need to be kept relatively stable in functioning than is currently to enhance customer satisfaction with service provision.

Additionally, it is realized that waiting time are shorter and system performance improved when customers form a single queue than when multiple separate queues are formed. This is in consonance with Xiao and Zhang (2009) conclusions in their study. A change from the current mixed arrangement (multi queues in some cases and single in other cases) to a strictly single queue arrangement will make a substantial improvement in the system's performance. Management could relay such information to customers as part of the instruction guide that accompanies ATM card issuance. It is advised further that future changes should be evaluated by employing queuing analysis to predict the efficacy of the proposed modifications, prior to actual implementation.

5. Conclusion

The experience of waiting long hours at ATMs to say the least is annoying to customers and unprofessional on the part of providers who superintend over this. Long waiting time in queue accounts for the loss of customers and attendant losses such as loss of goodwill and reduction in customer satisfaction. This paper has provided a queuing theory based approach to optimize ATM service level in the midst of fluctuating demands. The case bank operates two ATMs which are considered in queuing terms as the servers. ATM users are considered customers (a term used generically to refer to entities that request for service at a particular service centre). After running a series of goodness of fit test on arrival times and service times it was considered appropriate to represent the interaction between customers and the ATMs in a typical queuing system arrangement as an M/M/s queuing model. The prevailing operation characteristics are thus computed as output of the M/M/s model via an application of a decision support system. Service times and arrival times served as basic inputs to the system. Customer service times and arrival times were measured by conducting an observational study at the case bank. Questionnaires were also run to bring to perspective views of customers on bordering issue of customers waiting in queues at ATM terminals. The population considered was ATM customers who arrive at the terminal between the hours of 8:00 am and 4:00 pm. This was done over a three weeks period except Sundays. The customer queuing characteristics were observed and the general system performance computed. Results further show that long waiting time occurs from about 11:00 am to 1:00 pm within which period demand is high. Queues are shorter in the early mornings and late evenings.

The problem of queues and long waiting time is multifaceted and related to; the number of service windows, the arrival rate and service rate and other external factors. It has been determined that two ATMs that are currently offered to customers are sufficient to handle customer demands at an optimal service rate of 0.60. However, waiting times are still high during certain hours of the day and the month. This is because much of the presented system capacity is unused during such periods primarily brought about by failures in system operation. When this capacity suddenly becomes available for use, possibly huge demands from customers who were denied service prior to a particular active period creates congestion at the center since customer arrivals are not scheduled. Therefore, the problem of long waiting at the ATM terminals is not caused only because of insufficient ATMs but also because of service unavailability, which is due majorly to component faults. This has been determined as a major weakness of the case ATM operations. The nature of queue formation at the ATM as well varies. Study results show that single queue formation at the case ATM turn to yield better system performance than when multiple queues are formed. The correlation between waiting time in bank halls and at ATM terminals makes for an exciting research.

References

Abor, J. (2005), “Technological innovations and banking in Ghana: An evaluation of customers’ perceptions”. Legon: University of Ghana. Available: http://www.financialcertified.com/Technological%20Innovations%20and%20Banking%20in%20Ghana%20-%20AAFM.doc (May 11, 2012)

Aldajani, M. A., and Alfares, H. K. (2009), Location of banking automatic teller machines based on convolution. Systems Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.

Baxtex, P. and Jack, S. (2008), Quantitative case study: study design and implementation for novice researchers. McMaster University, West Hamilton, Ontario, Canada.

Copper, R. B. (1981), Introduction of queuing theory. 2nd ed. Elsevier North Holland, Inc., New York.

Creemers, S. Lambrecht, M. and, (n. d.), Queuing models for appointment-driven systems.

Dharmawirya, M. and Adi, E. (2011), “Case Study for Restaurant Queuing Model”. International Conference on Management and Artificial Intelligence, IPEDR, vol.6.

Ford, J. L. (1980), Two applications of queuing theory. Department of Human Service Studies, Cornell University, New York.

Kantner, L., Sova, D. and Anschuetz, L. (2005), Organizing qualitative data from lab and field: challenges and method.

Lariviere, M. and Van Mieghem, J. (2004), “Strategically seeking service: how competition can generate Poisson arrivals”. Manufacturing & Service Operations Management, vol. 6, pp. 23-40.

Musara, M. and Fatoki, O. (2010), “Has technological innovations resulted in increased efficiency and cost savings for banks’ customers?” African Journal of Business Management, vol. 4(9), pp. 1813-1821.

Robson, C. (1993), Real-world research: A resource for social scientists and practitioner – researchers. Malden: Blackwell Publishing.

Stake, R. E. (1995), The art of case study research. Thousand Oaks, CA: Sage.

Vasumathi, A. and Dhanavanthan, P. (2010), Application of simulation technique in queuing model for ATM facility.

Xiao, H. and Zhang, G. (2009), The queuing theory application in bank service optimization. Institute of Information & System Engineering, Henan University of Finance & Economics, Zhengzhou.

Yin, R. K. (2003), Case study research: Design and methods. 3rd ed. Thousand Oaks, CA: Sage.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Recent conferences: http://www.iiste.org/conference/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar