APPLICATION OF SIMULATION TO THE BANKING INDUSTRY

K. L. Slepicka  
University of Notre Dame  
Notre Dame, Indiana

G. A. Spohrer  
Crowe, Chizek and Company  
South Bend, Indiana

ABSTRACT

Computer simulation is examined as a tool to evaluate how the addition of an on line teller terminal system (OLTTS) will affect customer service in a bank. The St. Joseph Bank and Trust Company, in South Bend, Indiana, will be used as a case study bank to provide real data for this study.

Before running the simulation model, industrial engineering techniques were used to gather and analyze the data which was needed as input to the model. To analyze the present performance of the bank's tellers, time studies were used to determine customer service times. Pre-determined motion-time data from a bank which installed an on line system was used to estimate customer service times once the OLTTS was added. Statistical techniques were used to reduce the data to several normal distributions.

The following section discusses why computer simulation is used and presents an overview of the languages available. The actual model is discussed next, followed by a discussion of other possible uses of computer simulation techniques to solve problems in banking.

1. SIMULATION OVERVIEW

Simulation is a solution technique for defining and analyzing a model of a system. The system to be modeled in this study is a bank, which is made up of a complex system of queues. There are three basic elements in any queuing situation: an arrival process, a service mechanism, and operating rules for the arrivals and servers (Pritsker 1974). Each one of these can contain elements of uncertainty due to the stochastic nature of the physical process. For example, thirty customers may arrive per hour at one branch bank, but the time between arrivals will not always be two minutes. Similarly, the time required for service can vary greatly among customers. Analytical solutions to queuing problems usually require assumptions to simplify the physical situation, and give answers in terms of expected values.

Computer simulation handles many queuing situations which analytical methods cannot, including the jockeying, reneging, and balking of customers. Simulation can also handle customer service distributions other than exponential and, in a multi-server system, different average service times for servers. Different situations can be easily modeled by changing a few input data cards.

Activity times and the time between arrivals in many physical situations are governed by probability distributions. Real data from the system being modeled can be collected and tested to determine if it fits a particular distribution using statistical goodness of fit tests. The computer model then randomly samples from the distribution to determine how long the next activity will take or when the next arrival will occur in the computer model.
This paper will focus on a discrete simulation of a bank. In discrete simulation, time is advanced in discrete steps. Using the notation of Pritsker, the system is defined as the portion of the universe that the model affects. Systems consist of related groups of objects called entities. Each entity has a set of attributes which completely define the entity to the system. Entities engage in activities which occur over time. Activities start and end in events, and events do not take any time. The state of the system is changed by an event. The time between events is a random variable based on the physical characteristics of the system. The system does not change between events, thus the computer advances time in non-uniform steps to the next event. This type of discrete simulation is called next-event simulation.

The simulation analyst could write an entire computer program every time a simulation of a new system is desired, but in any simulation, certain processes are always used. Block simulation languages were developed to help the analyst provide many subroutines and functions commonly used. These routines include random number generation, the time advancing mechanism, and statistical collection capabilities.

GPSS, SIMSCRIPT, and SIMULA are among the most popular discrete block languages. GASPIV is also in this group and is used in this project.

The next section discusses the assumptions made about the branch that was modeled and how the customer service times were arrived at.

2. MODEL DEVELOPMENT

The St. Joseph Bank has thirteen branches in South Bend and vicinity. The Roseland branch was chosen as having a typical mix of customers and business volumes. This branch has five lobby tellers, and two drive-up tellers serving three drive-up lanes. Meetings were held with the personnel at the branch to discuss the operation of the branch and to give the analyst a better understanding of the physical situation to be modeled. Also discussed were the limitations of simulation and what physical situations could be simplified. After these discussions, the following mutually acceptable assumptions were made:

- All tellers in one area follow the same customer service time distribution. Differences in speed of customer service are omitted for all lobby tellers and for all drive-up tellers.
- All tellers can handle all types of customer transactions. There are no special service queues.
- Miller (1975) states that universal tellers lead to better overall productivity gains using an on-line system.
- Idle time for a teller is defined as the time a teller is not occupied with a customer. For example, a teller may be counting cash when a customer arrives at the window. The teller will stop to service the customer, and then resume counting. According to this definition the teller is idle while counting cash.
- The model simulates one hour of the branch operation. During this hour, no change in the number of tellers is possible.
- Flip-flopping of a teller between the lobby and drive-up is not a normal procedure, and is omitted.
- In the lobby, one to twelve tellers are possible.
- Jockeying is allowed in the lobby and occurs when a line is two or more persons shorter than an adjacent line. The person at the end of the longer line will jockey to the end of the shorter line.
- In the drive-up, two configurations are possible. Either there is one teller serving one lane via a sliding drawer, or there are two tellers serving two lanes via one sliding drawer and two vacuum tube units. All three lanes are fed by one approach lane.
- In the lobby, two queueing schemes are possible. Either there are as many lines (queues) as there are tellers, or there is only one queue and the person at the head of the queue goes to the next available teller.
- Balking is possible in both the lobby and the drive-up. Once there is a certain number of customers in the system, the next arriving customer will not enter the system. There is a 50 percent chance that a customer balking at the drive-up will park and enter the branch lobby to be served.
- When faced with several queues, the customer will pick the shortest one.
- All queues follow a first-in-first-out (FIFO) priority.

It should be emphasized that some of these assumptions are made to accommodate the bank being modeled and are not limitations of computer simulation. For example, with minor modifications, each teller within the model can be made to follow a different distribution for customer service.

This model will be used to determine how customer service will be affected by the addition of an OLTTS. The first step in verifying the model was to be sure all logic relations were valid within the model. Once the model was running properly, using hypothesized customer service time distributions, real data was collected to determine customer service times in the lobby and drive-up of the branch being modeled.

Both the lobby and the drive-up arrivals were assumed to follow a Poisson process, thus the time between arrivals was exponentially distributed in both cases. This is typical of many physical queueing situations.

The service time distributions in the lobby and drive-up were thought to be normal, but perhaps with different means. The drive-up mean was expected to be lower because drive-up tellers perform a more limited range of transaction than tellers in the lobby. For example:
The drive-up vacuum tubes are not large enough to hold commercial transactions, and not powerful enough to send rolls of coin. Therefore, customers having large or long transactions will bring their business into the lobby.

The drive-up tellers are not allowed to handle transactions requiring the customer to sign a document (such as a traveler's check) in front of the teller, thus these customers also will come into the lobby.

Lobby tellers are nearer to their customers than are drive-up tellers, and spend more time exchanging pleasantries with their customers. Thus the same transaction will take longer in the lobby than it will at the drive-up.

In the lobby, service begins when the customer arrives at the teller window and ends when the customer leaves. At the drive-up, the customer arrives at a lane and must send the transaction to the teller via either the vacuum tubes or the sliding drawer. Once the teller receives the transaction, it is processed by the teller and sent back to the customer. The customer does not leave the system until he/she takes the transaction out of the tube or drawer, and drives away. Due to this more complicated service mechanism, the drive-up was first studied to determine how long it takes: a) from the time the customer arrives until the teller receives the transaction, and b) from the time the teller returns the transaction until the customer leaves.

The data showed that the time it took for the transaction to pass through the vacuum units was the same for lanes two and three. However, this time was longer than the time it took for transactions to pass through the sliding drawer in lane one. Variations in observed time did occur because some customers would quickly send their transaction, while others would fumble in a purse or pocket, and others would sit at the vacuum unit and write out their check to be cashed. Similarly, upon receiving their transactions back, some customers would drive away quickly, while others would count their change, and put it in their purse or pocket before driving away. The results of this part of the study showed the mean in lane one to be .2088 minutes, with a standard deviation of .146 minutes. For lanes two and three, the mean was determined to be .4335 minutes, with a standard deviation of .180 minutes. A Kolmogorov-Smirnov (K-S) goodness of fit test was performed, and both distributions were normal.

Two additional distributions were required to analyze the customer service provided by the teller: one which follows present customer service - off line - and a second for customer service when the OLTSs is added. Time studies were made at one branch for both lobby tellers and drive-up tellers. Observations included the time of the transaction as well as the type of the transaction. The time study resulted in data points arranged into several categories, each category being one type of transaction. K-S tests performed on each type of transaction showed that transaction times were normally distributed. The overall distributions were assumed to be normal because the sum of a normal distribution is also a normal distribution.

Because the case study bank is only now considering the purchase of the on line system, data for customer service with an OLTS at the St. Joseph Bank is not available. A bank in Ohio, however, did install an OLTS and performed a study of teller transactions with the system before installing it. The pre-determined motion-time system of Methods-Time Measurement (MTM) was used to derive customer service times. This data, developed by Shogren, allows an overall distribution of teller transaction time to be calculated for an on line system.

For the different types of transactions in both the lobby and the drive-up, there is an observed average time from the time study - without an OLTS - and the MTM time for the same transaction types with an OLTS. The overall customer service average is computed by taking the average time for each type of transaction that occurs and multiplying it by the frequency of that transaction type. The value obtained for the average of customer service in the lobby with the OLTS is greater than the value obtained for the drive-up. This difference accounts for the broader range of transactions which occur in the lobby of the bank. The average in the lobby - with the OLTS - was also adjusted to account for the conversation between tellers and customers, which occurs in the lobby to a greater extent than at the drive-up.

At the lobby or the drive-up, the observed average of the time to cash a check should be the same, as the procedure is the same in either case.

<table>
<thead>
<tr>
<th></th>
<th>Drive-up observed average</th>
<th>Lobby observed average</th>
<th>MTM data value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.865 min</td>
<td>1.321 min</td>
<td>.707 min</td>
</tr>
</tbody>
</table>

Notice the MTM value with an OLTS is approximately 20 percent less than the observed average in the drive-up. Yet because of additional teller-customer interaction in the lobby, the comparison between the lobby average and the MTM value would not be valid. Therefore, one-half of a minute was added to the average of the MTM values when comparing them with the lobby averages, to account for the conversation. The final results, shown in Table 1, are four normal distributions, for lobby and drive-up, with and without an on line teller terminal system.

TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>Time Study (Without OLTS)</th>
<th>MTM (With OLTS)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LOBBY</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.107</td>
<td>.6944</td>
</tr>
<tr>
<td>Average</td>
<td>.9900</td>
<td>.9910</td>
</tr>
<tr>
<td><strong>DRIVE-UP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>.5759</td>
<td>.2654</td>
</tr>
</tbody>
</table>
Now that these distributions for customer service are known, all that is required for the simulation is the arrival process. The simulations will be run at three arrival levels: sixty, ninety, and one hundred twenty arrivals per hour. These arrival levels correspond to the branch being not very busy, busy, and quite busy for the simulated hour of operation. The simulation starts with the bank and queues empty. The program runs for sixty simulated minutes to allow the system to fill with customers and to reach steady state. Then new statistics are collected as the program runs for another sixty simulated minutes. The program provides information on the percentage of time each teller was busy, the number of customers in the system, the number of busy tellers, the number of customers in each queue, the number of customers who balked, and customer service variables. These consist of the time customers spend in the system and the time customers spend waiting to be served. The program also provides histograms of the values of these two customer service variables, so the percentage of each variable which is under a certain time can be easily seen.

In general, banks use two measures of how well they are serving their customers. The first is customer waiting time, and the second is the percentage of customers who wait less than a specified period of time. The St. Joseph Bank rates itself on customer service according to how long customers have to wait. If customers wait less than three minutes, the bank is doing well. A wait of seven minutes is considered unacceptable.

3. RESULTS

In the following section, inputs to the computer simulation model are varied, and their effects on customer waiting are examined. For each set of simulation runs, two conditions are compared at each of the three arrival rates - sixty, ninety, and one hundred twenty per hour. The pertinent statistics from these computer runs are summarized in Table 2.

The first set of simulations compares customer service with and without an OLTS, using the same number of queues as tellers in the lobby. Table 2, runs 1 through 6, show the results. The only factors which change between runs are the customer service times. The results show that the on line system improves customer service for both the average customer waiting time and the percent of customers who wait less than three minutes to be served. Notice that the addition of the OLTS improves the percentage dramatically at an arrival rate of one hundred twenty customers per hour, but the percentage does not change at an arrival rate of sixty customers per hour. When the branch is less busy, the improvement in customer service is less noticeable.

The second set of simulations was made for the same situation as the first set, except that in this series of runs, the lobby has one queue which feeds all five lobby tellers. The results, shown in Table 2, runs 7 through 9, show that the addition of the on line system improves customer service.

<table>
<thead>
<tr>
<th>Run No.</th>
<th>Area of Operation</th>
<th>Number of Tellers</th>
<th>Number of Queues</th>
<th>Planned Arrival Rate (Per Hr)</th>
<th>Actual Arrivals (Per Hr)</th>
<th>% Customers Who Wait Less Than 3 Minutes With OLTS</th>
<th>Average Waiting Time (Minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lobby</td>
<td>5</td>
<td>5</td>
<td>120</td>
<td>155</td>
<td>91.0</td>
<td>28.1</td>
</tr>
<tr>
<td>2</td>
<td>Lobby</td>
<td>5</td>
<td>5</td>
<td>90</td>
<td>104</td>
<td>100.0</td>
<td>99.0</td>
</tr>
<tr>
<td>3</td>
<td>Lobby</td>
<td>5</td>
<td>5</td>
<td>60</td>
<td>61</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>4</td>
<td>Drive-Up</td>
<td>2</td>
<td>3</td>
<td>120</td>
<td>115</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>5</td>
<td>Drive-Up</td>
<td>2</td>
<td>3</td>
<td>90</td>
<td>89</td>
<td>69.6</td>
<td>51.6</td>
</tr>
<tr>
<td>6</td>
<td>Drive-Up</td>
<td>2</td>
<td>3</td>
<td>60</td>
<td>56</td>
<td>98.3</td>
<td>93.4</td>
</tr>
<tr>
<td>7</td>
<td>Lobby</td>
<td>5</td>
<td>1</td>
<td>120</td>
<td>155</td>
<td>89.8</td>
<td>32.5</td>
</tr>
<tr>
<td>8</td>
<td>Lobby</td>
<td>5</td>
<td>1</td>
<td>90</td>
<td>104</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>9</td>
<td>Lobby</td>
<td>5</td>
<td>1</td>
<td>60</td>
<td>61</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>10</td>
<td>Lobby</td>
<td>4</td>
<td>4</td>
<td>120</td>
<td>155</td>
<td>19.6</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>Lobby</td>
<td>4</td>
<td>4</td>
<td>90</td>
<td>104</td>
<td>99.0</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>Lobby</td>
<td>4</td>
<td>4</td>
<td>60</td>
<td>61</td>
<td>100.0</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>Lobby</td>
<td>5</td>
<td>5</td>
<td>90</td>
<td>104</td>
<td>99.0</td>
<td>98.1</td>
</tr>
<tr>
<td>14</td>
<td>Drive-Up</td>
<td>2</td>
<td>3</td>
<td>50</td>
<td>51</td>
<td>100.0</td>
<td>96.0</td>
</tr>
<tr>
<td>15</td>
<td>Lobby</td>
<td>3</td>
<td>3</td>
<td>20</td>
<td>30</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>16</td>
<td>Drive-Up</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>14</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Also, this queueing mechanism improves customer service over that of each teller in the lobby having a queue.

The final set of simulations considers whether the addition of the OLTS can effect a reduction in the staff of a bank by changing the number of tellers used in the lobby of the bank with and without OLTS.

Four lobby tellers with an OLTS will be compared to five tellers without an OLTS. In both cases, there is one queue per teller in the lobby and the drive-up is not affected. As shown in Table 2, runs 10 through 12, the results are very close, but the five tellers without the OLTS give slightly better customer service. As in previous comparisons, the effect of adding an OLTS is not as pronounced at lower arrival rates.

The St. Joseph Bank had a detailed work management study done in late 1977 which suggested a system for staffing the bank. As a part of the study, arrivals were recorded at every branch. This arrival data will be used as input for the simulation of two branches. The first branch has five lobby tellers, and two drive-up tellers serving three drive-up lanes, and the second branch has three lobby tellers, and one drive-up teller serving one lane. The peak arrival rates at each branch were used.

For each branch operating at peak arrival rates given by the 1977 study, results of a comparison of customer service with and without an OLTS are shown in Table 2, runs 13 through 16. These results parallel the previous results. At both branches, customer service can be expected to improve with the addition of an OLTS.

The overall result of these simulations is that an OLTS does lead to better customer service. A reduced staff using the on-line system will handle the same number of customers as a larger number of tellers without an OLTS. Conversely, the same number of tellers will be able to handle more customers, which might offset additional hiring to handle the increased volume in the future.

The simulation data show that an OLTS will improve customer service. The bank could keep the same staff and maintain the same level of service in the future as the number of customers increased. If on the other hand, the number of customers served by the bank remains at its present level, the bank could reduce its teller staff by 20 percent.

4. CONCLUSIONS

The previous sections have shown that the installation of an OLTS will lower teller transaction time, raising the tellers' productivity. It has also been shown, using computer simulation, that this increased productivity would lead to better customer service. Computer simulation, however, is not limited to modeling the customers in a branch bank. The same concepts used in this model can easily be applied to analyze problems in other areas of banking.

A computer simulation program can be developed to model any stochastic process by simply modifying or changing the attributes of entities within the system being modeled. For example, the flow of documents to and through a bank's proof department can be simulated. Other areas possible to model might include customer flows to evaluate alternative layouts within a branch bank, cash flows between branches and the bank's central office, or even the size and layout of different types of parking lots.

In conclusion, computer simulation has been used in most industries to model many different phenomena, solving problems which have no easy analytical solution. This same technique can be applied successfully in banking and other service industries. Once the model is developed, it can easily be used to examine the effect of different parameters on the variables in the model.

REFERENCES


Personal correspondence with R. S. Shogren, Cleveland, Ohio.


