Improving Customer Waiting Time at a DMV Center Using Discrete-Event Simulation

Georges M. Arnaout (Ph.D. Student)*; Shannon Bowling (Ph.D.)
Engineering Management and Systems Engineering - Old Dominion University (Norfolk, VA)
garna001@odu.edu

Abstract. Virginia’s Department of Motor Vehicles (DMV) serves a customer base of approximately 5.6 million licensed drivers and ID card holders and 7 million registered vehicle owners. DMV has more daily face-to-face contact with Virginia’s citizens than any other state agency [1]. The DMV faces a major difficulty in keeping up with the excessively large customers’ arrival rate. The consequences are queues building up, stretching out to the entrance doors (and sometimes even outside) and customers complaining. While the DMV state employees are trying to serve at their fastest pace, the remarkably large queues indicate that there is a serious problem that the DMV faces in its services, which must be dealt with rapidly. Simulation is considered as one of the best tools for evaluating and improving complex systems. In this paper, we use it to model one of the DMV centers located in Norfolk, VA. The simulation model is modeled in Arena 10.0 from Rockwell systems. The data used is collected from experts of the DMV Virginia headquarter located in Richmond. The model created was verified and validated. The intent of this study is to identify key problems causing the delays at the DMV centers and suggest possible solutions to minimize the customers’ waiting time. In addition, two tentative hypotheses aiming to improve the model’s design are tested and validated.

1. INTRODUCTION
The usage of simulation has increased noticeably in the recent years due to the advancement of computer technology. The act of simulating behaviors and situations has been adopted in multiple areas like military, social behavior, flight simulators, robotics, etc. In this paper, we use Discrete Event Simulation (DES) since our aim is modeling the DMV system as it progresses over discrete times in a non-continuous fashion. In this model, we attempt to mimic the behavior of the real DMV center system by building our model from variables that are generated from data that is collected from experts of the DMV headquarter in Richmond, VA. The model is examined thoroughly and conclusions and solutions are produced from this study. Additionally, the study identifies two possible scenarios to enhance the system, and determines if they present a statistical significance to the model.

1.1 OBJECTIVES OF THE STUDY
The study is conducted in order to (1) minimize the customer waiting time at the DMV – 850 Widgeon Road, Norfolk, (2) give insights towards minimizing the customer waiting time at all the DMV centers, statewide, or around the country, (3) attempt to improve the existing model (i.e. Should we add another check-in window?), and (4) give suggestions aiming for optimizing the system. The focus of the study will be on reducing the following three delays: 1-The ticketing waiting time: time needed to obtain a service ticket. 2-The service window waiting time: time needed to reach the service window and be serviced. 3-The transaction time: time needed to be serviced.

2.0 THE MODEL
The model built using Arena (version 10) from Rockwell Systems, is a miniature of the existing DMV center located at 850 Widgeon Rd, Norfolk, VA.

2.1 Model Details
Customers arrive in a stochastic way according to an inter-arrival rate produced by exhaustive observations conducted at the Widgeon center. The model has a main queue called the “check-in"
queue. This is the main queue where all customers have to pass through in order to get their tickets and proceed to the nearby seating area, waiting for their ticket number to be announced. When a customer's ticket number is announced, the customer proceeds to one of the service windows in order to be served in a FIFO manner. As there are 14 service windows, according to our extensive observations, only 10 windows are being used at the same time. Thus, for the sake of the study, we consider 10 windows with 10 servers serving on them. There is a separate M/M/1 queue for each of the windows (i.e. 10 separate queues) that we will emphasize graphically in our model. In the real system, we cannot observe these queues separately as the customers are seated all together. There is a range of 10 to 14 servers serving at these windows according to a ‘real’ weekly schedule obtained from the Widgeon DMV center. The waiting times, transaction times, and other delays are generated from the collected data (please refer to the Data Collection section). The model runs for 8 hours on weekdays (serving from 9am until 5pm) and 4 hours on Saturday (8am until 12pm). Each DMV agent is allowed to take one ‘30 minutes break’ during the day on weekdays, and no breaks on Saturdays. The breaks are divided between three groups. First group of employees can take the break between 11am and 2pm. The second group can take the break between 1pm and 2pm. Finally, the third group of employees can take the break between 2pm and 3pm. As a summary, the model works like the following: Customers arrive to the system, wait at the check in (ticketing) queue in order to obtain a ticket. Then, customers go wait in line in order to obtain a ticket (ticketing wait time) before moving to one of the 10 available service windows. The customers wait for another delay until they are serviced by a DMV agent, which is the service wait time. Then, the customers are serviced with a service delay called Transaction Time. Finally, the customers leave the system.

2.2 Model Constraints

Our study has several limitations due to the time factor and the nature of the study:

- The types of services that customers request are ignored. Due to the limitation of the data collected, the types of services are overlooked and all the service delays are recorded as one service type delay. For example, the time that a customer spends for obtaining new car license plates is combined with the time of another customer trying to obtain a driving license (which is remarkably longer).
- Customers that leave the DMV center for any reasons (e.g. missing documents) are still counted in the time statistics but not modeled in our system.
- The customer’s inter-arrival rate and the waiting time at the main ticketing queue are obtained by interviewing experts as well as extensive observation at the DMV Widgeon center [3].
- Holidays as well as the busiest days of the month (i.e. first day and last day of the month) are not counted in our model. However, Fridays and especially Saturdays are considered busier than the other days.
- The DMV employees’ weekly schedule is considered static although it changes weekly.
- We assume that customers arrive one at a time to the DMV center.
- All the units used in this study are in minutes.

2.3 Model Design

The model built using Arena 10.0 is represented in figure1. The key model variables are:

- 10 service windows
- 10 Service Queues - (Waiting Time)
- Customers arriving in a stochastic way to the center
- Service Time (or Trans Time)
- Main queue – customer check in (time needed to acquire a ticket)
- Number of DMV agents defined by a schedule ranging between 10-14 during the day (including breaks)

Figure 1: DMV Model Design in Arena 10.0

A: Entity creating the users arriving to the DMV in a stochastic way.
B: Recorder that catches the time of arrival of the customers for further calculation
C: Module that seize the agent on the main check in queue in order to serve the customer
D: Module that delays the service time of the agent according to a distribution explained in the Data Collection section
E: Module that releases the agent after finishing from servicing the customer at the check in queue
F: Module that delays the customer service time before reaching the service window according to the next available window queue (refer to the Data Collection section).

<table>
<thead>
<tr>
<th>Table 1: DMV Weekly Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
<tr>
<td>F</td>
</tr>
</tbody>
</table>

G: Module that decides which is the next available service window according to a small program to calculate the service window containing the least number of customers waiting in its queue.

H: Ten service windows that serve the customers.

I: Recorder that calculates the flow time of the customers in the system using the previous recorder (B).

J: Module that allows the customers to exit and leave the system.

3. DATA COLLECTION

The data is mainly collected from extensive observation of the center, interviewing experts from the headquarters, and from the DMV weekly data sheets provided by the DMV experts [2]. The weekly data collected from the DMV experts provided with around 100 data points that were used to generate the distributions of the related delays. The customers *inter-arrival rate* is generated from a schedule that resulted from extensive observation and interviewing the system's experts. The arrival schedule is implemented from Monday through Saturday. On weekdays, the customers arrive between 9am and 5pm. On Saturdays, the customers arrive between 8am and 12:00pm. The *ticketing waiting time* is generated from a Triangular distribution that resulted from extensive observation of the center and the behavior of customers. The expression used is $EXPO(0.62)$. The *service waiting time* delay is generated from a Weibull distribution that resulted from plotting the data points on a histogram (refer to Figure 2), provided by the DMV weekly data. The expression produced is: $4 + WEIB(17, 2)$. We considered this distribution a good fit for our collected data because it has a very low Square Error (which is 0.005248) and the p-value is remarkably larger than 0.05.

![Figure 2: Histogram of Service Waiting Time Delay](image1)

The *transaction time (service time)* is generated from a log normal distribution that resulted from plotting the data points provided by the DMV weekly data in the histogram shown in Figure 3. The expression produced is: $5.2 + LOGN(1.81, 0.803)$. Figure 3 plots the histogram of the data collected and includes the distribution summary. We consider this distribution a good fit for our collected data because it has a very low Square Error (which is 0.0628) and the p-value: 0.101 is remarkably larger than 0.05.

![Figure 3: Histogram of the Transaction Time Delay](image2)
4. OUTPUT ANALYSIS

After running the model for 10 replications where each replication represents a week composed of 6 business days, we came up with the following results:

- The weekly average of customers coming to the DMV center of Widgeon Road is 2054 customers.
- The total average waiting time for each customer (or customer flow time) is 41.65 minutes.

The queuing delays are represented in details in Table 2 below:

<table>
<thead>
<tr>
<th>Waiting Time</th>
<th>Average</th>
<th>1st</th>
<th>Median</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Check-in Queue</td>
<td>2.77725</td>
<td>1.83</td>
<td>1.9113</td>
<td>7.9614</td>
</tr>
<tr>
<td>Service Window 1 Queue</td>
<td>2.30725</td>
<td>0.72</td>
<td>1.0726</td>
<td>4.3867</td>
</tr>
<tr>
<td>Service Window 2 Queue</td>
<td>2.30851</td>
<td>0.72</td>
<td>1.0726</td>
<td>4.3867</td>
</tr>
<tr>
<td>Service Window 3 Queue</td>
<td>2.90847</td>
<td>1.09</td>
<td>1.3725</td>
<td>8.2350</td>
</tr>
<tr>
<td>Service Window 4 Queue</td>
<td>18.7113</td>
<td>4.18</td>
<td>7.0683</td>
<td>27.7253</td>
</tr>
<tr>
<td>Service Window 5 Queue</td>
<td>18.4785</td>
<td>5.95</td>
<td>8.0317</td>
<td>34.3370</td>
</tr>
<tr>
<td>Service Window 6 Queue</td>
<td>20.9472</td>
<td>5.42</td>
<td>9.1003</td>
<td>38.6350</td>
</tr>
<tr>
<td>Service Window 7 Queue</td>
<td>23.0520</td>
<td>5.09</td>
<td>9.1411</td>
<td>40.6500</td>
</tr>
<tr>
<td>Service Window 8 Queue</td>
<td>18.0203</td>
<td>6.60</td>
<td>9.4556</td>
<td>44.2974</td>
</tr>
<tr>
<td>Service Window 9 Queue</td>
<td>20.3454</td>
<td>6.90</td>
<td>9.7793</td>
<td>46.2830</td>
</tr>
<tr>
<td>Service Window 10 Queue</td>
<td>27.2823</td>
<td>6.87</td>
<td>10.3778</td>
<td>46.2237</td>
</tr>
<tr>
<td>Service Window 11 Queue</td>
<td>28.4406</td>
<td>6.99</td>
<td>10.9392</td>
<td>47.8204</td>
</tr>
</tbody>
</table>

We observe that there is an average of 3.77 minutes wait time at the check-in queue (ticketing) which we categorized as fair. In addition, looking at the queuing occurring at the service windows, the average transaction time was around 21.65 minutes. Here, it is necessary to mention that the service window delays are not the delays only related to the service time at the window, but also includes the waiting time before the customers get served by an agent. The average service time is around 8 min which in our opinion does not need further improvement.

That leaves the deficiency of the system to only one variable which is the excessive arrival rate of the customers, which in turn, affects all the other delays causing the excessive queues.

5. MODEL VERIFICATION AND VALIDATION

Our V&V was conducted in parallel with the system's experts at the DMV headquarters. By comparing our model's results with their weekly data, we've found that our generated distributions, our model (with its variables), and our results were valid. For the inter-arrival rate of the customers, according to the DMV weekly sheets (of the real system), an average of 2105 customers visited the DMV at Widgeon weekly. According to our model, the average was 2054 customers which is considered very close, and therefore valid. As for the other delays (e.g. Transaction delay), the distributions were verified via emails with a senior analyst at the DMV headquarters in Richmond [reference needed]. The final results of the study and the possible solutions were submitted to the headquarters upon their request and are being studied by their analysts.

6. ALTERNATIVE SCENARIOS

Observing the customers' waiting time at the ticketing window in our model's animation, in Table 2, and in the real system, inspired us to come up with two different alternatives different from having one main ticketing queue.

6.1 Alternative 1

We considered having an additional ticketing window resulting in two parallel check-in queues that are served by two agents. After implementing this addition to our model and running it for 10 replications (just like the original model), we realized that the customer waiting time was slightly reduced from 41.65 minutes to 40.28 minutes. In order to find out if this alternative was worth implementing, we conducted a Paired-t test on the two approaches (this one and the original model). We concluded that the change was not statistically significant (the two means overlapped at 0.05 level).

6.2 Alternative 2

The second alternative was to increase the number of agents by having 14 agents working at all times (including breaks). This alternative has two sub-alternatives. One, by increasing the number of agents to the original model without adding another ticketing queue (i.e. having two ticketing windows), and the other one by adding another ticketing window. After implementing the changes (in both sub-alternatives) and running the model for 10 replications, we concluded that this scenario is also not statistically significant.

7. CONCLUSION

At the beginning of the study, we were considering the proposal of a Self-Check in kiosk to speed up the ticketing phase, as a parallel approach to the main check in service window. Our reason for proposing such an approach was our belief (prior to running the model) that having two separate (but parallel) check-in queues will speed things up and minimize the customers' waiting time. After
implementing and running the model, and after experimenting with the two alternatives proposed (refer to section 6), we concluded that having a self-check in kiosk will not have a significant positive difference on the existing model, and therefore decided to drop that suggestion. Thus in our opinion, this limits the delay to two main system design gaps. Either the service time (transaction time) is relatively high, or the arrival rate is just too excessive. For the first gap, the service time can be reduced by increasing the number of agents but also, increasing the number of service windows proportionally. This would reduce the service time remarkably and affect the overall waiting time of the visiting customers. Here, it must be noted that the pace of the service is relatively fine. The serving pace does not need to be enhanced since according to our records, the average transaction time for each customer is 21.65 minutes (including the time waiting to be serviced), which is relatively fair. Therefore, the queuing is not occurring from the transaction time (i.e. service time). As for the second gap, the arrival rate can only be reduced by offering more online services (but also keeping the option of physically visiting the DMV center for these services). This will reduce the arrival rate of customers since, with the digital age and the ease of access to go online, customers would most probably prefer conducting the transactions online (e.g. from their work office) rather than spending time to visiting the DMV center. Several DMV centers started giving appointments to their customers in order to balance and control their inter-arrival rate. This approach is also feasible and could be implemented at the Widgeon center.

REFERENCES
2. Powell, Minni, personal communication, June 2009
3. White, Robert, personal communication, July 15, 2009