The following publication E. R. Magsino and I. W. H. Ho, "An intelligent highway tollgate queue selector for improving server utilization and vehicle waiting time," 2016 IEEE Region 10 Symposium (TENSYMP), Bali, Indonesia, 2016, pp. 271-276 is available at https://doi.org/10.1109/TENCONSpring.2016.7519417.

An Intelligent Highway Tollgate Queue Selector for Improving Server Utilization and Vehicle Waiting Time

Elmer R. Magsino and Ivan W. H. Ho

Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University
elmer.magsino@connect.polyu.hk and ivanwh.ho@polyu.edu.hk

Abstract — **On a highway setup, a vehicle will most probably use a tollgate server that has the shortest queue thinking that it is the fastest exit. In this paper, an intelligent highway tollgate queue selector using fuzzy logic is proposed and simulated in Matlab SimEvents. Its aim is to automatically select the most appropriate tollgate server for a vehicle to ensure the shortest waiting time while trying to balance the server's utilization. Two policies are considered in this study, namely: (1) Shortest Queue (SQ) and (2) Fuzzy Logic-Controlled Queue (FLCQ). Results indicate that the FLCQ policies reduce the average waiting time and queue length by approximately 50% of those obtained from the SQ policy while guaranteeing an equal utilization among available servers. These findings are valid for light and heavy, homogeneous and nonhomogeneous vehicle arrivals. To further improve the decision making of the fuzzy logic controller, traffic flow information collected by remote road-side units can be exploited to allow the control of various system parameters (e.g., service time) in advance.**

Keywords—**Tollgate system, server utilization, vehicle waiting time, queue selector, fuzzy logic**

I. INTRODUCTION

Intelligent Transportation System (ITS) is an integration of various disciplines dealing with services and applications to transportation, its infrastructure and traffic management. Its objectives are dissemination of pertinent traveler's information, reduction of traffic congestion and travel time and improve traffic safety and efficiency, thus improving quality of life [1, 2].

One key transportation infrastructure is the highway that normally has more vehicle lanes than the conventional roads. This is to accommodate a higher vehicle volume. It is also characterized by having no traffic light signals, presence of entry/exit ramps and tollgates. On a freeway, there are three volume-density-based control regulations to prevent or eliminate congestion, namely: (1) on-ramp (most efficient), (2) speed control and (3) merging control [3]. In this paper, the tollgates at the end of the freeway are also classified as one of these approaches, and we focus on minimizing the queueing time of vehicles at the tollgates. If the approaches above fail to maintain a stable traffic flow, a large volume of vehicles bunch up at these exit tollgates. This is another source of inconvenience to vehicle owners and travelers.

Earlier works have focused on the optimal allocation and management of highway tollgates. In [4], the optimal number of operating tollgates was derived based on the economic costs of the tollgates. The tollgate system was based on an M/D/1 model whether in inbound or outbound directions. According to their findings, with a total limit of 12 operating tollgates both for inbound and outbound, for an inbound arrival rate of 8.33 cars/min, four tollgates are needed while for an outbound rate of 18.33 cars/min, eight tollgates are required. If we consider a peak-hour traffic of 30 cars/min (1800 cars/hr) per road lane as in [5], then the total operating tollgate exceeds the set limit of 12 just to accommodate a single (inbound/outbound) lane.

The traffic delays in tollgates at Port Authority were studied in [6] and the proposed solution was the development of an efficient scheduling of its manpower to reduce traffic delays and minimize its operational costs. The study was done in a span of 14 months and was able to achieve scheduling efficiencies of 95% or better. Other research studies analyzing tollgate scheduling scenarios are seen in [7, 8, 9]. In [9], fuzzy logic was used for the dynamic allocation of tollgate plaza capacity based on the desired level of service with minimum expenses. Fuzzy logic systems are considered as one of the most important technologies that can play a significant role in intelligent transportation system [10, 11, 12].

In this research study, for a given combination of the number of highway lanes and number of tollgates, the fuzzy logic controller selects for an incoming vehicle the server queue with the least amount of queueing time while at the same time ensuring that the servers have fair utilizations. The simulated system is a real-time decision making controller based on a homogeneous assumption of service times, homogenous and nonhomogeneous vehicle arrival rates and fixed queue capacities. This addresses the growing vehicle volume while the highway infrastructure are constant.

This paper is organized as follows: Section II discusses the development of the highway model for simulation, the two policies to be compared, i.e. Shortest Queue (SQ) and Fuzzy Logic-Controlled Queue (FLCQ) policies. Section III provides the simulation results and discussion of the two policies based on a set of performance metrics. Finally, this paper is concluded in Section IV. Future directives are also given in this section.

II. DEVELOPMENT OF AN INTELLIGENT HIGHWAY TOLLGATE QUEUE SELECTOR

In this section, we discuss the model development of the highway system in Matlab/Simulink and the shortest queue and fuzzy logic controlled policies.

A. Matlab/Simulink Highway System Development

Fig. 1 shows a section of a highway having N vehicle lanes and i exit tollgates with corresponding l_i queue capacities that is considered in this work. An access point (AP) is placed on the highway system where the vehicle communicates to obtain which server to queue according to the proposed intelligent highway queue selector. Table I summarizes the assumptions used in developing the Matlab SimEvents model.

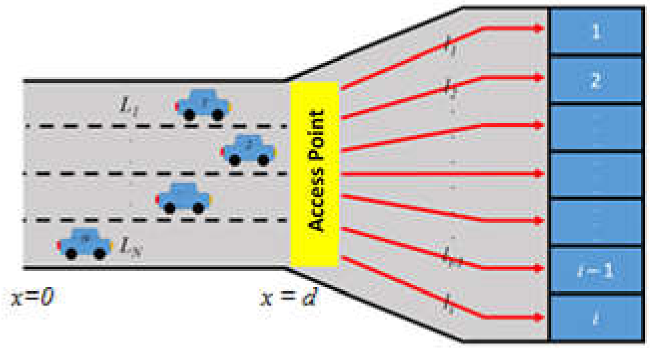


Figure 1. Highway Tollgate System

Fig. 2 depicts the Matlab/Simulink model incorporating modeling blocks from SimEvents Library of a Fuzzy-controlled Tollgate system for the highway segment shown in Fig. 1. The three blocks highlighted by the dashed box represent the highway tollgate system following the SQ policy.

The three main blocks are: (1) VehicleGeneration, (2) TollgateQueues and (3) TollGateServers. The VehicleGeneration block generates vehicles following a (non)homogeneous Poisson distribution. The stochastic traffic model used came from [5] and we assume that there is no car joining or leaving the highway segment before the tollgate. The highway scenario in this study is modeled as a single stage queuing model with multiple queues and multiple parallel servers, (M/M/c/K). Though vehicles originate in more than one highway lane before $x = d$, service will be dependent on a first come, first served basis (FIFO).

TABLE I. ASSUMPTIONS FOR THE HIGHWAY AND TOLLWAY MATLAB SIMEVENTS MODEL

1. Each lane $L_{z = 1, \ldots, N}$ is characterized by an independent, identical Poisson distribution with its own (non)homogeneous arrival rate of vehicles.
2. All vehicles have the same length.
3. Tollgate queue lengths ($l_{x=1 \ldots i}$) are fixed and based on the number of vehicles it can accommodate.
4. Tollgate service times ($T_{Sy=1 \ldots i}$) are characterized by a deterministic (exponential) distribution.

The TollgateQueues block determines which tollgate a vehicle will go through to exit the highway. The queue capacities are defined before any simulation run depending on the tollgate configuration. Fig. 1 represents a symmetric configuration. The TollGateServers block contains all the available servers. Service time (T_S) can be set to be a constant or exponential distribution.

B. Shortest Queue and Fuzzy Logic-Controlled Policies

There are two policies governing how a vehicle queues in a server that are presented: (1) Shortest Queue (SQ) [7] and (2) the proposed Fuzzy logic-controlled Queue (FLCQ). The SQ Policy is used at a certain instance by any approaching vehicle to select which server currently has the shortest queue. This selection does not consider which server has the quickest service time or the highest utilization. On the other hand, the FLCQ Policy determines which lane the approaching vehicle should queue. The FLCQ takes into consideration the server's current lane density and its service time to make the decision.

The SQ Policy pseudocode is shown in Table II below. A vehicle approaching a tollgate server makes its decision based on the current server lane density.

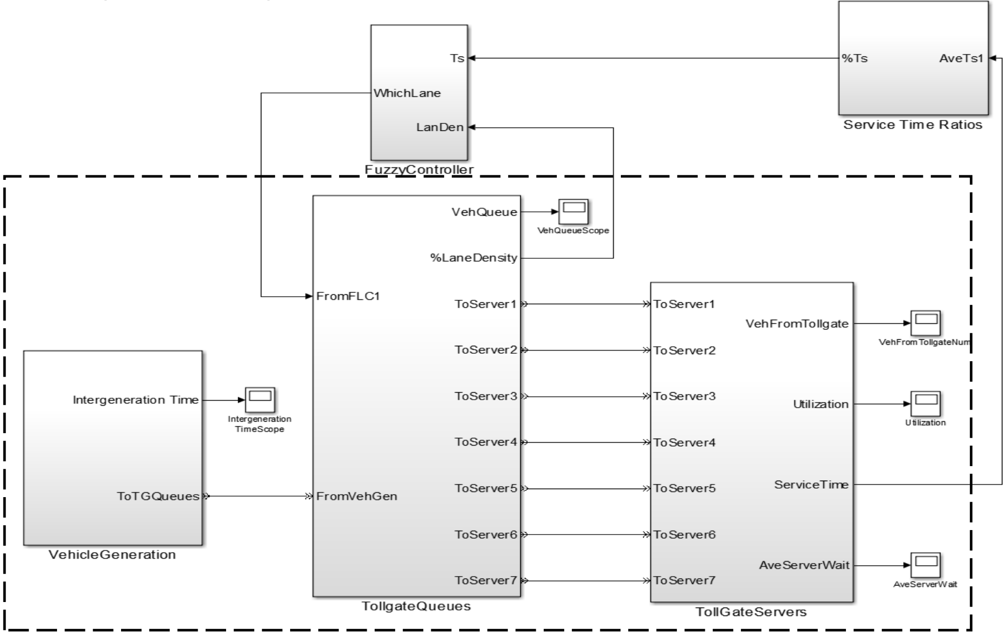


Figure 2. Simulink Model of a Fuzzy-controlled Tollgate System

TABLE II. SHORTEST QUEUE PSEUDOCODE

```
Get Lane Densities, $QD_i$
Determine ALL Possible Queues ($PQ$) with minimum $QD$
if $PQ > 1$
    Select a queue based on a uniform random number
end if
Output Queue number
```

Queue Density (QD_i) is the ratio between the number of vehicles in a queue and the queue capacity. Mathematically, it is expressed as:

$$QD_i = \frac{n_i}{l_i} \qquad (1)$$

where:

n_i = number of cars in queue i
l_i = queue i car capacity

Normally, if there are more servers having the same number of vehicles in queue, it will choose any of the middle servers because of the traveled distance involved. Therefore, during a light traffic flow, tollgate servers having queues with the least queue capacities will be highly utilized.

For the FLCQ policy, the crisp inputs chosen to be fuzzified are the lane density and service time ratio (TS_{Ratio}). TS_{Ratio} is defined as:

$$TS_{Ratio} = \frac{T_{Si}}{T_{S\max}} \qquad (2)$$

where:

T_{Si} = service time of server i
$T_{S\max}$ = maximum service time from all servers

The service time ratio allows the fuzzy logic controller to have an idea on the relative service time of a certain tollgate with respect to the other tollgates. The defuzzified output is the probability of the tollgate being chosen. Fuzzy control is then developed using the rule base that is generally implemented by:

$$QP = fuzzy(TS_{Ratio}, QD)$$

where the server time ratio and queue density are the preconditions while the consequent is the queue probability (QP). Creating the fuzzy rules are based on the idea that when the server time ratio is small (thus being the fastest), and the lane density is low (thus being the shortest), the probability of choosing the server queue is high. The fuzzy rule base is shown in Table III.

TABLE III. FLCQ POLICY FUZZY RULE BASE

TS_{Ratio} / QD	F	OK	S
SQ	***HP***	***HP***	***HP***
OK	***HP***	***OK***	***OK***
LQ	***HP***	***OK***	***LP***

The *LD* inputs are categorized as: SQ (short queue), OK (nearing half-filled queue) and LQ (long queue). On the other hand, the TS_{Ratio} inputs are characterized by: F (fast), OK (just right) and S (slow). There are two types of membership functions (MFs) used to represent the inputs, namely, (1) triangular MFs and (2) trapezoidal MFs. These are shown in Figs. 3 and 4 below respectively.

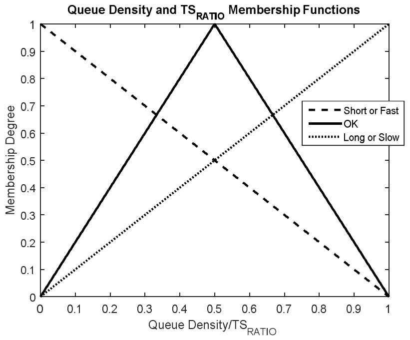


Figure 3. Queue Density and Service Time Ratios using Triangular Input Membership Functions

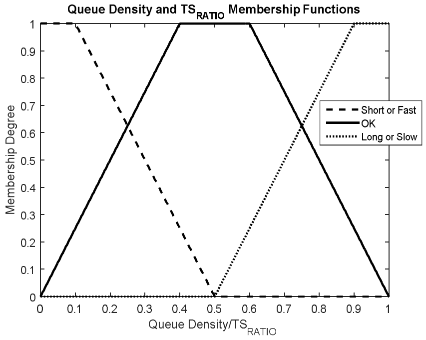


Figure 4. Queue Density and Service Time Ratios using Trapezoidal Input Membership Functions

The possible outputs for these two approaches in input membership functions are: HP (high probability), OK (more or less 50% of being selected) and LP (low probability). These are depicted in Fig. 5. The linear membership functions were chosen because of the advantages it holds, i.e. simple implementation and fast computation.

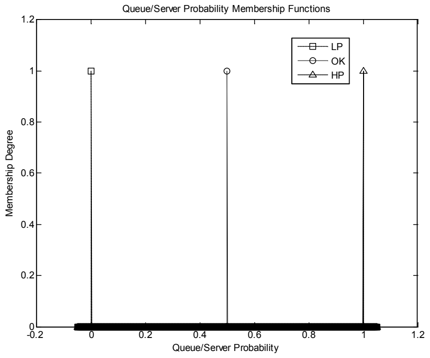


Figure 5. Lane Probability Output Membership Functions

To clearly see the input-output relationships of the proposed fuzzy logic controllers, the surface views of the fuzzy rule base in Table III along with the two approaches in Figs 3–4 are shown in Figs. 6–7 respectively. The surface views show all possible combinations of inputs and its corresponding probability output.

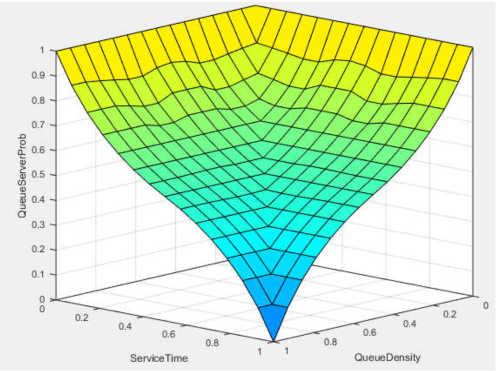


Figure 6. Fuzzy surface view (Triangular Input MFs)

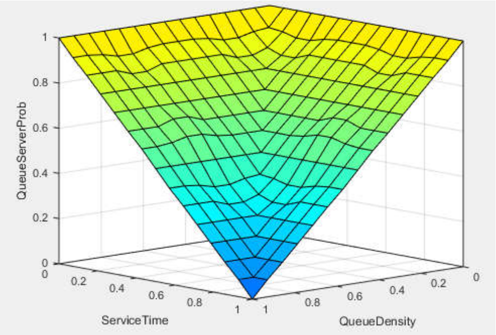


Figure 7. Fuzzy surface view (Trapezoidal Input MFs)

The implication method of the fuzzy logic controller used the *min* operator, while the aggregation of each rule output was done by the *max* operator. In the implication method, the result (a single number) derived from "*and*-ing" the preconditions is used to reshape the consequent by truncating the output fuzzy set. This is done for all set of rules. On the other hand, the aggregation method combines all truncated output fuzzy sets (inputs) from the implication method into a single fuzzy set by getting the maximum value when comparing all inputs. Finally, defuzzification is implemented by finding the centroid of the aggregated output. For example, if the $QD = 0.25$ and $TS_{Ratio} = 0.5$, then $QP = 0.667$. The process discussed below is shown in Fig. 8 for triangular member input functions. The values of both QD and TS_{Ratio} are drawn vertically into its corresponding MFs for all the nine rules. The intersection of this vertical line with the MFs creates the shaded regions (see columns 1 and 2). We compare these two regions point by point and get the minimum of the two. This newly formed region will now be used to reshape the output MFs (see column 3). The aggregation method combines all truncated output MFs by comparing them point by point and getting the maximum value among all compared points (last row and last column text box). Finally, the centroid is determined to arrive at $QP = 0.667$.

All server-queue pairs have their own probability of being chosen. A Matlab function then selects the one with the highest probability. If two or more server-queue pairs are possible, then it chooses randomly following a uniform distribution.

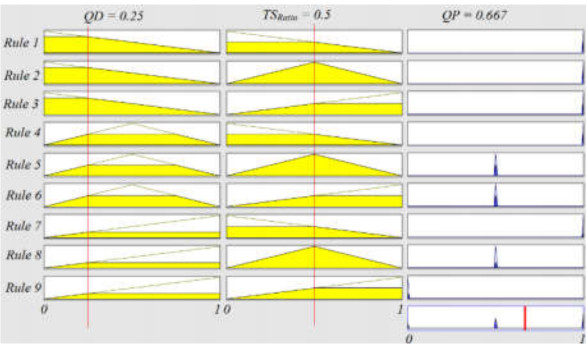


Figure 8. Fuzzy Logic Process

III. SIMULATION RESULTS AND DISCUSSION

The tollgate symmetric configuration shown in Fig. 1 is used to simulate and compare the SQ and FLCQ policies. There are four ($N = 4$) highway lanes and seven ($i = 7$) tollgate servers/queues. The performance metrics used are: (1) Average Queue Length (QL_{Ave}), (2) Average Queue Waiting Time (QW_{Ave}) and (3) Server Utilization ($Util$). The average queue length is defined as the time average of the number of vehicles. The server utilization is defined as ratio of time spent servicing a vehicle over the total simulation time. In queueing theory, it is defined by the expression below.

$$U_i = \frac{\lambda_i}{\mu_i} = \frac{\frac{n_i}{T_{Total}}}{\frac{1}{T_{S_i}}} = \frac{n_i T_{S_i}}{T_{Total}} \tag{3}$$

where:

λ_i = arrival rate at a tollgate server i
μ_i = tollgate i service rate
n_i = number of cars in queue i
T_{Si} = Tollgate server i service time
T_{Total} = Total simulation run time

The highway lane arrival rate of each lane follows either a homogeneous or nonhomogeneous Poisson distribution. The queue capacities (in number of vehicles) are set to $l_1 = l_7 = 50$, $l_2 = l_6 = 40$, $l_3 = l_5 = 30$ and $l_l = 20$ while the service times are modeled exponentially. Table IV summarizes the different arrival rates used.

TABLE IV. SIMULATION PARAMETERS

Simulation Trial	**Arrival Rate Per Highway Lane (cars/min)**			
	Lane 1	***Lane 2***	***Lane 3***	***Lane 4***
1	12	10	8.57	7.5
2	30	20	15	12
3	30,20	15,12	10,8.57	7.5,6.67

The third simulation trial represents the nonhomogeneous arrival rates at the highway lanes. For Trials 1 and 2, the simulation time was set to 7200 sec (2 hrs) while for Trial 3, simulation time was set to 14400 sec (4 hrs). For example, the arrival rate of $[x,y]$ in Table IV Trial 3 signifies that during the first hour of simulation, there are *x* cars/min, for the second hour, there are *y* cars/min and then becomes periodic.

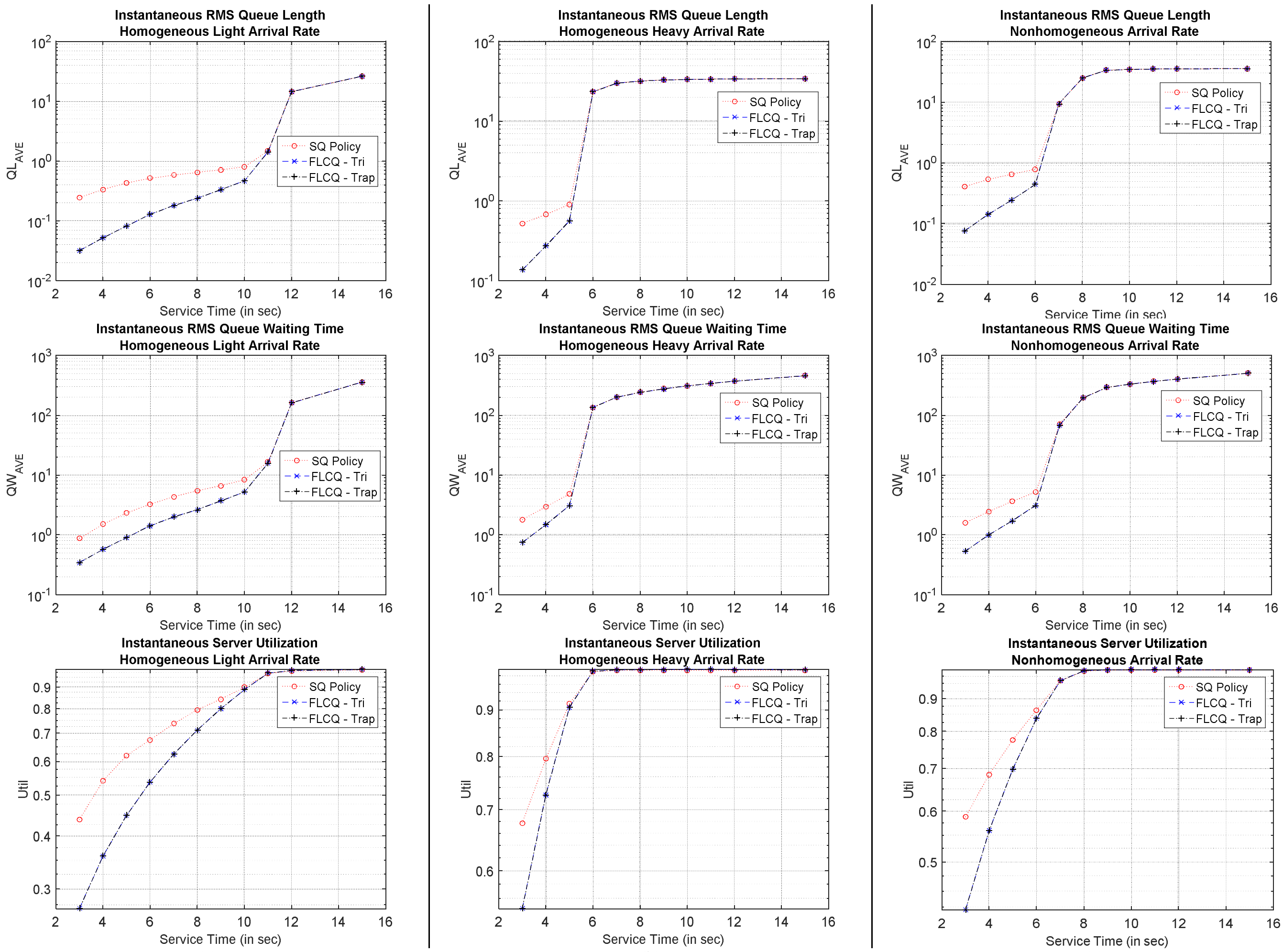
Fig. 9 shows the combined RMS results of all simulation trials of all queues/servers. The average queue length, waiting time and server utilizations of all queue/servers were observed under varying but homogeneous servers' service times with values from 3 – 15 seconds. Note that the y-axis is in semilog scale to highlight the differences between the policies at various traffic conditions.

Regardless of traffic condition, the service times creates a knee in the performance metrics that separates the interval of fast and slow service times. During fast service times, the FLCQ policies provide shorter average queue length and waiting time for a vehicle approaching a tollgate compared to the SQ policy. It introduced a maximum approximate improvement of 50%. There is also a decrease in the server utilization that practically translates to the servers not tiring out quickly. During the interval of slow service times, i.e. after the knee, the performance metrics are generally the same for all policies.

The maximum RMS queue capacity of 38.54 vehicles is reached and the average waiting time is at its maximum and is equivalent to the instantaneous RMS queue length multiplied by the service time. Finally, the server utilization can be seen to be fully utilized. Eq. (3) is used to obtain the numerical value of the instantaneous server utilization.

Between the two FLCQ policies, the triangular MFs offer a better response than the trapezoidal MFs just before the response's knee. From the surface views of the two FLCQ policies, we note that trapezoidal MFs provide a constant change on the queue/server's probability to be chosen. This means that all queues/servers have an equiprobable chance of selection.

On the other hand, examining the triangular MFs surface view reveals that the shorter lane densities and faster service times are given a higher range of probabilities over the half-filled and "OK" service time ratios. Also, if the traffic condition is heavy, a very small chance of being selected is given to a certain queue/server. In this sense, it can be taken that triangular MFs lead to a more reactive policy than the trapezoidal MFs if following the 9-rule base defined in Table III.



(a) (b) (c)

Figure 9. Simulation Results for (a) Light, (b) Heavy and (c) Nonhomogeneous Arrival Rates

Fig. 10 shows the response of the two FLCQ policies over the instantaneous RMS queue length performance metric for each queue/server. It can also be seen in Fig. 10 that the queue length has been equalized. It follows here that the average queue waiting time is also the same as well as the server's utilization. This is not the case in the Shortest Queue policy.

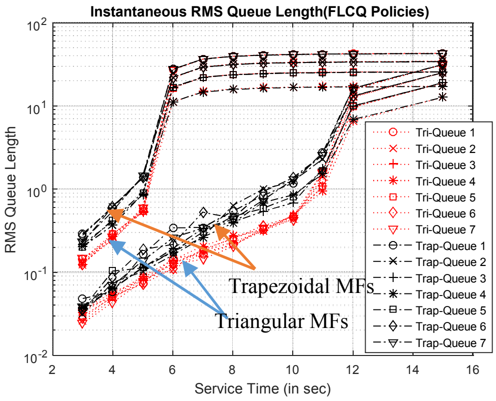


Figure 10. Comparison between the Triangular and Trapezoidal MFs for the FLCQ policy

To further enhance the advantage of the FLCQ over the SQ policy especially during longer service times, the server was provided an early warning capability once the lane density is approaching its maximum. In this simulation run, it was assumed that after the FLCQ signal has been given, the servers were able to reduce half of its service time in a linear function. Fig. 11 shows there is a decrease in the RMS server utilization after the addition of this capability. In practice, road side units (RSUs) can be installed along the highway to monitor the traffic density. The FLCQ constantly communicates with these RSUs to update the current traffic density information and decides if there is a need to inform the server to hasten its service time.

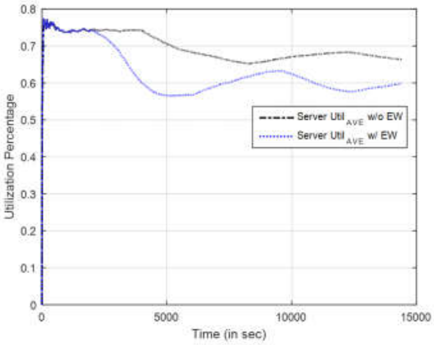


Figure 11. Comparison of FLCQs Average Server Utilization with and without early warning (EW) capability

IV. CONCLUSION AND RECOMMENDATIONS

In this paper, we have successfully shown that the incorporation of a fuzzy logic controller in a highway tollgate system can effectively decrease a vehicle's average waiting time and average queue length, and improve the server's utilization. This is true for both homogeneous and nonhomogeneous vehicular arrivals in the highway lanes especially with tollgates of fast service time. As an added feature to the FLCQ selector, an early warning signal was introduced to allow the server to reduce its service time when the traffic is building up due to relatively long service times.

This work can be further extended by including a practical modeling of the service time reduction and the scalability of servers, i.e. deactivation of existing servers or addition of needed servers based on existing infrastructure. The nonhomogeneous service time of manual servers can also be taken into consideration. Finally, allowing communications between vehicles and highway infrastructure of patterns/features detected by camera or sensors [13] will further develop a more intelligent and highly adaptive highway tollgate system in the future.

ACKNOWLEDGMENT

This work is partially supported by the Early Career Scheme (Project No. 25200714) established under the University Grant Committee of the Hong Kong Special Administrative Region, China, the National Natural Science Foundation of China (Project No. 61401384), and The Hong Kong Polytechnic University (Projects 4-ZZCZ, 4-ZZDF, 4-ZZEF, G-YBK6, and G-YN17).

REFERENCES

[1] L. Tsu-Tian, "Research on Intelligent Transportation Systems in Taiwan," in *27th Chinese Control Conference*, Kunming, Yunnan, China, 2008.

[2] X. Yan, H. Zhang and C. Wu, "Research and Development of Intelligent Transportation Systems," in *11th International Symposium on Distributed Computing and Applications to Business, Engineering & Science*, Guilin, China, 2012.

[3] Y. Wang, W. Bin and Z. Hui, "The National Freeway Control System - Further Development with Fuzzy Logic Theory," in *International Conference on Industrial Technology*, Guangzhou, China, 1994.

[4] L. He, Q. Gao, S. Li and Z. Peng, "Optimal Allocation of Operating Toll Booths at Highway Toll Station," in *Sixth International Conference on Business Intelligence and Financial Engineering*, Hangzhou, China, 2013.

[5] I. W.-H. Ho, K. K. Leung and J. W. Polak, "Stochastic Model and Connectivity Dynamics," *IEEE/ACM Transactions on Networking for VANETs in Signalized Road Systems,* vol. 19, no. 1, pp. 195-208, 2011.

[6] L. C. Edie, "Traffic Delays at Toll Booths," *Journal of the Operations Research Society of America,* vol. 2, no. 2, pp. 107-138, 1954.

[7] E. Correa, C. Metzner and N. Nino, "TollSim: Simulation and evaluation of toll stations," *International Transaction in Operational Research,* vol. 11, pp. 121-138, 2004.

[8] J. Danko and V. Gulewicz, "Operational planning for electronic toll collection: a unique approach to computer modeling/analysis," in *Winter Simulation Conference*, Phoeniz, Arizona, USA, 1991.

[9] E. Ivanjko, M. Matulin and S. Mrvelj, "Dynamic capacity allocation of tollbooth plaza based on fuzzy logic," in *23th International Central European Conference on Information and Intelligent Systems*, Varazdin, Croatia, 2012.

[10] K. N. Karna, "Artificial Intelligence in Intelligent Transportation Systems," *Journal of Intelligent Transportation Systems,* vol. 2, no. 3, pp. iii-viii, 1995.

[11] M. Davarynejad and J. Vrancken, "A Survey of Fuzzy Set Theory in Intelligent Transportation: State of the art and future trends," in *IEEE International Conference on Systems, Man, and Cybernetics*, San Antonio Texas, USA, 2009.

[12] N. K. Swain, "A Survey of Application of Fuzzy Logic in Intelligent Transportation Systems (ITS) and Rural ITS," in *IEEE SoutheastCon*, Memphis, Tennesse, USA, 2006.

[13] J. Chen, B. Carrion Schafer and I. W.-H. Ho, "Implementation of a FPGA-Based Feature Detection and Networking System for Real-time Traffic Monitoring," arXiv:1603.06669, 2016.