

ADAPTIVE SIGNAL CONTROL FOR NON-LANE BASED VEHICLE MOVEMENTS

Sadguna Nuli¹, Saladi Sri Venkata Subbarao²

¹ Department of Civil Engineering, CVR College of Engineering, Mangalpalli, Vastu Nagar, Hyderabad - 501510, Telangana, India

² School of Engineering Sciences, Mahindra Ecole Centrale, Bahadurpally, Hyderabad – 500043, Telangana, India

Received 15 January 2018; accepted 10 April 2018

Abstract: State-of-the-art adaptive signal control models, commonly used in developed countries, uses upstream or advance detector information for determining control plan. These models work well in lane-based traffic conditions especially when multiple junctions are involved. However, traffic in many cities across the world is heterogeneous, characterised by non-lane based movement and mixed-vehicle type, which causes inaccurate estimation of turning proportions. These difficulties can be addressed to a great extent by placing the detector at the stop-line rather than in advance or upstream location. However, there is no truly adaptive traffic control models exists for such traffic conditions and sensor location. In this study, a signal control model is proposed that is truly adaptive and uses stop-line detector information. The model aims at real-time allocation of green times through actor-critic reinforcement learning; an approach originated from the machine learning community. This approach can learn relationships between signal control actions and their effect on the traffic system while determining optimal control policy. To test the performance of the model a typical four-way four phase intersection with variable flow was simulated using a traffic micro simulator (VISSIM) and interfaced with the proposed model. The performance of the model was compared with the traditional vehicle actuated system. The results using this approach shows significant improvement over traditional control, especially for varying traffic demand.

Key words: intersection, adaptive, traffic signal, reinforcement learning.

1. Introduction

Operating traffic signals in urban areas requires proper timings, so that varying demands can be managed effectively. Traditional algorithms, which are optimized off-line, usually generate a library of signal timing plans each with fixed stage duration and sequence. Plans are retrieved from the library for implementation according to the time of day and the day of week.

Such plans require regular updating, otherwise the performance declines over the time. Most of the operating signal systems today are traffic responsive or vehicle actuated. The responsiveness to traffic is that the allocation of green times is adjusted according to real-time traffic information. The real-time traffic data is usually detected by using inductive loops. Moving to adaptive traffic signal control, a substantial number of highly sophisticated

¹ Corresponding author: sadgunanuli@cvr.ac.in

and complex models are existing. However, the basic limitation common to all or most of these models is that turning proportions estimated using traffic flow models do not account for traffic heterogeneity and non-lane based vehicle movement, which results in considerable error in the estimations of corresponding green times. It may be noted that non-lane based mixed traffic conditions are very significant in most of the developing countries like India, and was well documented (Asaithambi and Shravani, 2017). In this study, a stop-line based adaptive traffic signal control model is developed to address such traffic conditions. The proposed framework addresses above limitations by using utilized green times and discharge information of the previous cycles obtained through detectors located at the stop-line.

1.1. Literature Review

At a signalized intersection, traffic signals typically operate in one of three different control modes: pre-timed, actuated, and adaptive control. In pre-timed control, all of the control parameters are fixed and pre-set off-line. These techniques are useful in generating the parameters for fixed timing plans of conventional pre-timed urban traffic control systems, where the traffic conditions during different time periods of the day (e.g. peak hours, off-peak hours) are estimated based on the historical information. In vehicle actuated control, the performance of the system can be rated between pre-timed and adaptive control systems. Actual green time falls between the pre-set minimum and maximum values. Actuated control strategy can partially solve the criticism attributed to the pre-timed control strategy in a sense that it can respond to the real-time traffic arrivals of the current green phase. However,

this actuated control strategy does not take into consideration of the queue lengths on other conflicting movements, and may result in suboptimal control especially when the traffic arrival rate is significantly different across all the phases (Zhang *et al.*, 2007). Adaptive traffic signal control performs much better than above two control strategies because they normally look ahead in time and space. Adaptive control strategies can be broadly categorized into heuristics, fuzzy logic and optimization models.

Heuristics such as rule based systems were proposed to have real time adaptive traffic signal control systems (Lin, 1988; Owen and Stallard, 1995). Further, to improve the estimations of traffic state Hu *et al.*, (2007), Chou and Teng (2002), and Zhang *et al.*, (2005) have applied fuzzy logic to model the traffic signal control problem in various aspects such as traffic movements, geometry, and phasing scheme. Although both of these models heuristics and fuzzy logic requires minimum computation time, they are not optimal when the traffic flow do not follow historical pattern. To address this, many authors proposed various optimization models such as DYPIC, MOVA, SCOOT, OPAC, and RHODES (Robertson and Bretherton, 1974; Vincent and Peirce, 1988; Robertson and Bretherton, 1991; Gartner, 1983; Mirchandani and Head, 2001). These models consider traffic signal control as a multi-stage decision making problem and solved it by move based search methods (gradient based or genetic algorithms) or constructive search methods (dynamic programming or branch and bound techniques) for obtaining optimal solution. Although, the models search for optimum solution, well known criticism about move based search methods is that solution is local optimum, whereas it is global optimum

in case of constructive search methods. However, under certain circumstances such as large-scale models like signal control, the excessive computation requirement of constructive search methods not suitable for real time applications. Another criticism is that optimization modelling approaches require accurate traffic arrival information with respect various turning movements, for the next few minutes to determine the best control plan. This information is often affected by traffic dispersion as well as non-lane based movement of vehicles, and is very difficult to obtain.

To overcome afore mentioned problems associated with optimization-based methods, especially those methods based on dynamic programming, Bingham (2001), Abdulhai *et al.*, (2003), and Xie *et al.*, (2010) have applied reinforcement learning (RL) technique to model traffic signal control. There are two major advantages of using the reinforcement learning to solve multi stage decision problems over using dynamic programming. First, the RL does not require state-transition probabilities and traffic arrival predictions as inputs. It can learn the state-transition probabilities interactively from the system operations. Secondly, after the model is trained, it has the same low computational requirement as that of rule-based methods have. Thus, it is more suitable for real-time applications.

Therefore, several problems with the existing applications can be identified as follows: (1) No adaptive traffic control model exist to account for non-lane based vehicle movements. (2) Most of the models were basically designed with respect to corridor or network level rather than isolated intersection. Therefore, the solution for

isolated intersection is sub-optimal. (3) Few models require more number of detectors in the form of stop-line as well as advanced/upstream locations, which is difficult from practical implementation point of view.

Therefore, the objective of this study is to design an adaptive traffic signal control model for an isolated intersection using stop-line detector information, so that the model can address the non-lane based vehicle movements. Further, the model does not require any accurate predictions as an input which otherwise effects performance of the control policy significantly.

2. Methodology

The basic idea of the proposed model is to bring adaptive feature to the stop-line based vehicle actuated controller through reinforcement learning as shown in Fig. 1. The signal control begins with initialization of various controller and model parameters such as minimum green g^{\min} , maximum green g^{\max} , unit extension e_o , and threshold gap h_{th} etc. If the time is more than the control period the controller stops else it works as a traditional vehicle actuated controller during the first cycle. After the first cycle, the model estimates maximum green for each phase of the subsequent cycle based on the utilized green times and discharges in the previous cycles. Thereafter, the controller terminates active phase according to the threshold gap criteria as in vehicle actuated control. The model learns about these control actions (assigning maximum green) through reinforcement learning. Reinforcement learning (RL) is an area of machine learning concerned with how an agent ought to take actions in an environment to maximize some notion of cumulative reward.

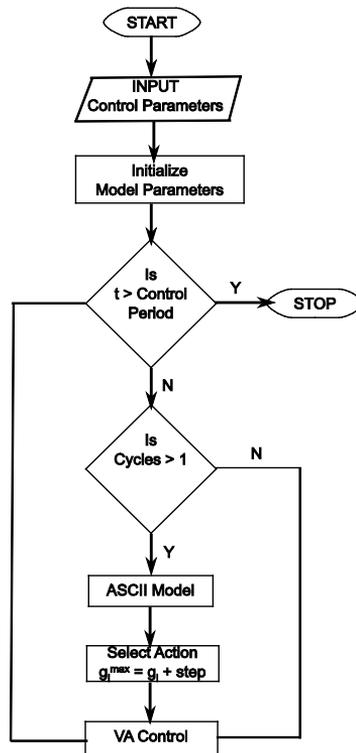


Fig. 1.
ASCII Controller

To deal with large scale problem such as multi-phase signal control, proposed model Adaptive Signal Control for Isolated Intersection (ASCII) integrates concepts of neural network and fuzzy logic into actor-critic reinforcement learning. Here fuzzy logic helps in representing large state space very efficiently whereas neural network generalizes the training experience to the states that were not visited during learning process. The design elements of the proposed model in terms of the typical reinforcement learning (i.e. state, action, and reward) are discussed below.

As the queue length cannot be measured using stop-line detector but utilized greens

can be measured accurately, the state of the system is represented by a vector of N components that are the utilized greens associated with each phase of the previous cycle. The action of the control agent is the maximum green for each phase of the upcoming cycle. The action space is represented by neighbourhood of the utilized greens in the previous cycle. The reward is defined as the increase in the total discharge, i.e., the difference between total discharges between two successive cycles. If the reward has a positive value, this means that the total discharge was increased by this value after executing the action. However, a negative reward value indicates that the action results in decreased discharge. The objective of the

model is to determine an optimal maximum green time so that the intersection discharge is maximized over the control period.

2.1. Modeling

Like neural networks, proposed ASCII model has four layers as shown in Fig. 2. The first 3 layers together represent state of the system and the action space is represented by fourth

layer. The learning experience is stored in connection weights λ_j^k and w_j^k between 3rd and 4th layers. The model can be described as follows: The first layer is the input layer. It receives traffic state variable values and sends them to different fuzzy membership functions in the second layer. Each node in the first layer represents an input variable. Each node in the second layer is a fuzzy set with a fuzzy membership function associated with it.

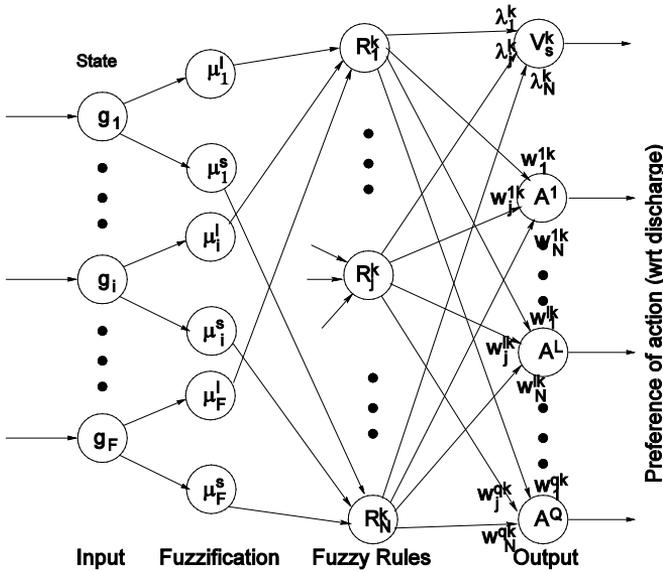


Fig. 2. Data Flow Across ASCII Model

The inputs to the second layer are the state variable values, and the outputs of the second layer are fuzzy membership function values. The inputs and fuzzy sets of the second layer constitute many linguistic terms such as *Green is Short* and *Green is Long*. Thus, the output of the second layer can be

considered as degrees of membership values associated with the short and long fuzzy variables (Zhang et al., 2007). The third layer corresponds to antecedent part of fuzzy rules in the fuzzy logic controller. For example, a sample fuzzy rule in the case of four phase signal control can be represented as:

$$if \{g_1 \text{ is long AND } g_2 \text{ is long AND } g_3 \text{ is long AND } g_4 \text{ is short}\} \text{ then } \{increment g^{\max}\} \tag{1}$$

Therefore, the output of the third layer can be considered as firing strengths. The fourth layer is a collection of nodes representing consequence part of fuzzy rules. The first node stands for the critic, and its output value indicates how good the current state value is i.e., in the present case it represents expected discharge by the end of control policy. The remaining nodes correspond to the available actions that can be taken, and their output values are the preferences to choose each action given the current state inputs.

The critic and action values of the policy are computed as follows: The input and fuzzification parts are same as that in typical fuzzy logic controller. Given the latest utilized greens as inputs, the *short* and *long* fuzzy variables compute membership values and these values are feed into the third layer of the model. Assuming the j^{th} fuzzy rule has the following F antecedents, Eq. (2):

$$g_i \in \mu_{ij}^m \quad \text{where } i = 1, \dots, F \quad (2)$$

Then the firing strength of the j^{th} fuzzy rule is Eq. (3):

$$R_j^k = \prod_{i=1}^F \mu_{ij}^m \quad (3)$$

The critic value V_s^k is computed as Eq. (4):

$$V_s^k = \sum_{j=1}^N R_j^k \lambda_j^k \quad (4)$$

Where, λ_j^k is weight connecting the j^{th} fuzzy rule and the critic output in the k^{th} cycle; similarly, the preference of choosing each action A^q is computed as Eq. (5):

$$A^q = \sum_{j=1}^N R_j^k w_j^{qk} \quad (5)$$

Where, w_j^{qk} is weight connecting the j^{th} fuzzy rule and the q^{th} action output in the k^{th} cycle.

The training of the model is done using well-known gradient descent technique. However, temporal difference (TD) error is used while applying this technique. The advantage of temporal-difference method is that it can learn directly from raw experience without either the model of the system's dynamics or targeted output. The TD methods update estimates based in part on other learned estimates, without having to wait for the final outcome.

Therefore, a TD error δ in estimating the critic value is computed as Eq. (6):

$$\delta = (V^{k+1} - V^k) + \gamma V_s^k - V_s^k \quad (6)$$

Where, V^{k+1} and V^k are cycle discharges in the $(k+1)^{\text{th}}$ and k^{th} cycles; γ is discount rate accounted for future discharge; V_s^k is the critic value estimated for the new state of green times in the $(k+1)^{\text{th}}$ cycle, V_s^k is the critic value estimated for the state of green times in the k^{th} cycle; and the new connection weights are updated as Eqs. (7-8):

$$\lambda_j^{k+1} = \lambda_j^k + \alpha \delta R_j^k \quad (7)$$

$$w_j^{q(k+1)} = w_j^{qk} + \alpha \delta R_j^k \quad (8)$$

To facilitate the agent to search overall state-action space during training, a well-known ϵ -greedy method is used for selecting a particular action. The ϵ -greedy learner selects greedy action most of the time except for a small amount (ϵ) of time, it selects a random action uniformly. The value of ϵ is chosen to

decrease gradually with iterations. This will result in more exploration at the beginning of the learning process which enables the agent to search the overall state-action space and gradually emphasizes exploitation as the model converges to the optimal policy (El-Tantawy and Abdulhai, 2010).

3. Model Testing

The proposed real-time adaptive signal control model was evaluated by measuring intersection delay, total throughput, and queue length. The performance of the widely used stop-line based vehicle actuated control was used as a bench mark and is compared to the proposed model. The

models were implemented using a scalable, high performance microscopic simulator VISSIM. It has been widely used in the testing of various algorithms and evaluation of various Intelligent Transportation Systems (ITS) strategies because of its powerful Application Programming Interfaces (API). The proposed model was developed as a VISSIM plug-in through API programming.

To demonstrate the scalability of the model, a typical four phase isolated intersection as shown in Fig. 3 was chosen. Since all the three movements from each approach are combined into a single phase, only one detector is placed at each stop line else more number of detectors are required depending upon the phasing scheme.

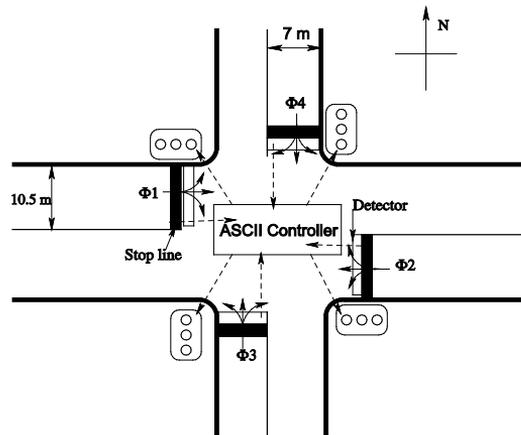


Fig. 3.

Illustrates Intersection Geometry, Phasing, Detectors and Controller Communication Details

Further, to test the robustness of the model diverging flow patterns (high volume on west and east approaches and low volume on south and north approaches) with typical day time variations as shown in Fig. 4 was given. To account for the mixed vehicle type and no-lane based vehicle movement, the VISSIM model of the test intersection was developed as per the guide lines given by Mathew and

Radhakrishnan (2010). Since performance of the vehicle actuated controller largely depends upon the optimization of controller parameters, the parameters such as minimum green (g^{\min}), maximum green (g^{\max}), unit extension (e_0), and threshold gap (h_{th}) are optimized as 10 s, 45 s, 3 s, and 2 s respectively by trial and error for the given traffic, geometric, and control conditions (FHWA, 2016).

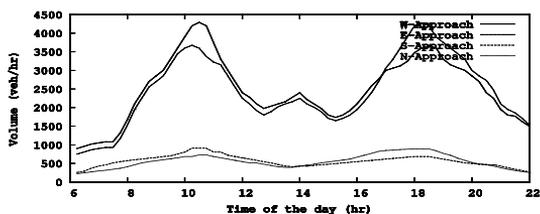


Fig. 4.
Flow Pattern on all the Four Approaches of the Intersection

During the simulation traffic volumes and allotted green times were measured for most of the day, specifically from morning 6 hours to evening 22 hours.

The volumes and green times obtained from vehicle actuated (VA) control model were presented in Fig. 5a and 5b respectively.

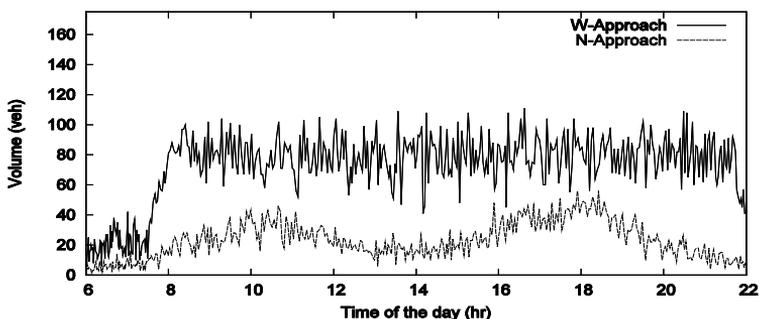


Fig. 5a.
Measured Volumes under Vehicle Actuated (VA) Model

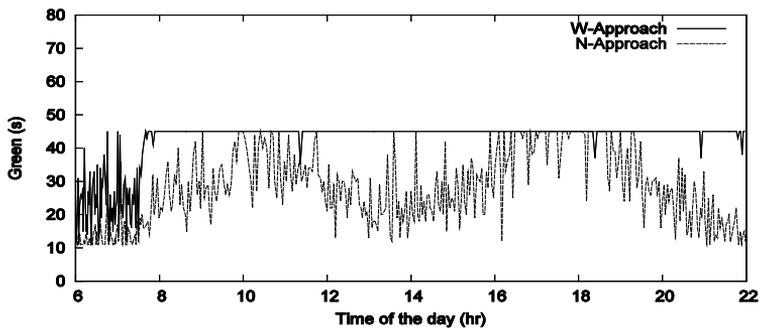


Fig. 5b.
Green Times obtained from Vehicle Actuated (VA) Model

Similarly, volumes and green times obtained from ASCII model were presented in Fig. 6a and 6b respectively. From the Fig. 5a and 5b it is clearly understood that though there is increasing demand on the West approach VA control was saturated to a pre-set maximum green of 45 s. However, proposed ASCII model sensitive to vary its green time according to the flow pattern on its approaches (Fig. 6a and 6b).

The use of reinforcement learning helped the model to look at the whole intersection in space and time ahead, and act accordingly by varying maximum green time in every cycle. Hence, it can be inferred that the proposed model is adaptive to real time traffic pattern. In addition, it can be noticed that the green times are well utilized without any wastage because of the use of stop-line detector.

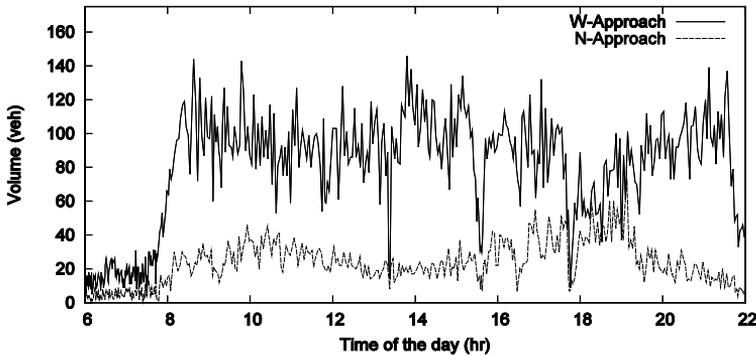


Fig. 6a.

Measured Volumes under ASCII Model

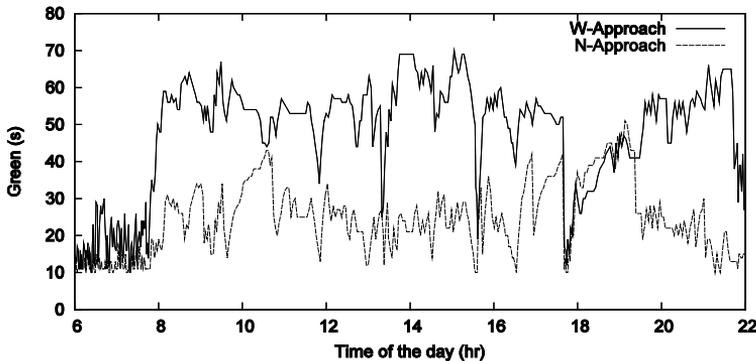


Fig. 6b.

Green Times obtained by ASCII Model

The overall performance comparison was summarised in Table 1. The delay decreased about 10 % and 24 % on West and East (high volume) approaches respectively when

compared with VA control. However, the delay increased about 91 % and 142 % on South and North (low volume) approaches respectively as expected. It may be noted

that such high delays are common in mixed traffic conditions due to complex interaction between vehicles (Bhuyan and Rao, 2011). Although delay increased on low volume approaches, overall intersection delay decreased about 3 %. Similarly, queue decreased about 10 % and 19 % on West and East approaches respectively, and increased about 96 % and 158 % on South and North approaches. Further, overall intersection queue decreased about 6 %.

In accordance with delay and queue, the intersection discharge increased on

high volume approaches (5 and 9 %), and marginally decreased on low volume approaches (0 and 2 %) due to lack of demand. Finally, overall intersection discharge increased (5 %).

Therefore, it can be summarized that the proposed control model is able to improve the intersection discharging capacity without penalizing overall intersection delay and queue values. Hence, it can be concluded that ASCII model is sensitive to look at the whole intersection rather than active phase (tunnel vision) alone as in VA controller.

Table 1

Comparison of Delay, Queue, and Discharge obtained from VA and ASCII Models

Approach	Model	Delay (s/veh)	% Change wrt VA	Queue (m)	% Change wrt VA	Discharge (Veh)	% Change wrt VA
West	VA	150	-10	152	-10	29005	5
	ASCII	135		137		30544	
East	VA	152	-24	153	-19	28305	9
	ASCII	116		124		30844	
South	VA	57	91	28	96	9088	0
	ASCII	109		55		9087	
North	VA	55	142	26	158	8930	-2
	ASCII	133		67		8778	
Intersection	VA	128	-3	122	-6	75328	5
	ASCII	124		115		79253	

4. Conclusion

This paper presented the frame work and evaluation of “Adaptive Signal Control of an Isolated Intersection (ASCII)”. The proposed model could bring adaptive feature to the vehicle actuated controller by estimating maximum green timings using stop-line detector information and neuro-fuzzy actor critic reinforcement learning. The problem of traffic heterogeneity and non-lane based vehicle movement is addressed by moving detector to the stop-line, measuring utilized green times and discharges. Therefore, it can be observed that the information given

to the model is more reliable even under such traffic conditions because it is not dependent on any forecasting as in other state-of-the-art models. Further, the model is more convenient with respect to field implementation as it requires only measured stop-line detectors information.

As the test results indicated that ASCII is a promising adaptive traffic control model in dealing with large scale problem of multi-phase signal control, and its applicability to real time operation, the model can be extended to control a corridor involving multiple intersections.

Acknowledgement

The authors would like to thank Department of Civil Engineering, IIT Bombay, India for their support in carrying out the experiments.

References

- Abdulhai, B.; Pringle, R.; Karakoulas, G.J. 2003. Reinforcement learning for true adaptive traffic signal control, *Journal of Transportation Engineering* 129(3): 278-285.
- Asaithambi, G.; Shravani, G. 2017. Overtaking behaviour of vehicles on undivided roads in non-lane based mixed traffic conditions, *Journal of Traffic and Transportation Engineering (English Edition)* 4(3): 252-261.
- Bhuyan, P.K.; Rao, K.V. 2011. Application of GPS and clustering techniques in defining LOS criteria of signalized intersections for Indian cities, *Highway Research Journal* 4(1): 69-75.
- Bingham, E. 2001. Reinforcement learning in neurofuzzy traffic signal control, *European Journal of Operational Research* 131(2): 232-241.
- Chou, C.H.; Teng, J.C. 2002. A fuzzy logic controller for traffic junction signals, *Information Sciences* 143(1): 73-97.
- El-Tantawy, S.; Abdulhai, B. 2010. An agent-based learning towards decentralized and coordinated traffic signal control. In *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 665-670.
- FHWA. 2016. Traffic signal timing manual, Federal highway administration. Available from internet: <<http://www.ops.fhwa.dot.gov>>.
- Gartner, N.H. 1983. Opac: a demand-responsive strategy for traffic signal control, *Transportation Research Record: Journal of the Transportation Research Board* 906: 75-81.
- Hu, Y.; Thomas, P.; Stonier, R.J. 2007. Traffic signal control using fuzzy logic and evolutionary algorithms. In *Proceedings of the IEEE Congress on Evolutionary Computation*, 1785-1792.
- Lin, F.B. 1988. Comparative analysis of two logics for adaptive control of isolated intersections, *Transportation Research Record: Journal of Transportation Research Board* 1194: 6-14.
- Mathew, T.V.; Radhakrishnan, P. 2010. Calibration of microsimulation models for nonlane-based heterogeneous traffic at signalized intersections, *Journal of Urban Planning and Development* 136(1): 59-66.
- Mirchandani, P.; Head, L. 2001. A real-time traffic signal control system: architecture, algorithms, and analysis, *Transportation Research Part C: Emerging Technologies* 9(6): 415-432.
- Owen, L.E.; Stallard, C.M. 1995. Rule-based approach to real-time distributed adaptive signal control, *Transportation Research Record: Journal of Transportation Research Board* 1683: 95-101.
- Robertson, D.I.; Bretherton, R.D. 1991. Optimizing networks of traffic signals in real time-the SCOOT method, *IEEE Transactions on vehicular technology* 40(1): 11-15.
- Robertson, D.I.; Bretherton, R.D. 1974. Optimum control of an intersection for any known sequence of vehicle arrivals. In *Proceedings of the 2nd IFAC/IFIP/IFORS Symposium on Traffic Control and Transportation Systems*, 3-17.
- Vincent, R.; Peirce, J. 1988. Mova: Traffic responsive, self-optimising signal control for isolated intersections, *Technical Report No.70*. Transport Research Laboratory. 21 p.
- Xie, Y.; Zhang, Y.; Li, L. 2010. Neuro-fuzzy reinforcement learning for adaptive intersection traffic signal control. In *Proceedings of the Transportation Research Board 89th Annual Meeting* (No. 10-0369).

Zhang, L.; Li, H.; Prevedouros, P.D. 2005. Signal control for oversaturated intersections using fuzzy logic, In *Proceedings of the Transportation Research Board 84th Annual Meeting Compendium of Papers*, [CD-rom].

Zhang, Y.; Xie, Y.; Ye, Z. 2007. Development and Evaluation of a Multi-Agent Based Neuro-Fuzzy Arterial Traffic Signal Control System. Technical Report No. 473700-00092, Texas A&M University, College Station, Texas.