A Multi-agent Auction-based Approach for Modeling of Signalized Intersections

Mehdi Mashayekhi[1], George List[1]

[1]North Carolina State University, Raleigh, North Carolina, USA
{mmashay2, gflist}@ncsu.edu

Abstract. This paper shows how the traffic interactions of an intersection can be regulated by means of auction theory, multi-agent systems and machine learning techniques. In the proposed work, the intersection space is viewed as a spatially complex set of scarce commodities whose temporal non-overlapping usage is adjudicated through an interactive bidding process that involves multiple agents. A machine learning technique is used to optimize the bidding strategies of the agents. The multi-tiered decentralized agent-based framework proposed here increases modularity by decomposing the intersection into smaller sub-parts, adds flexibility and effectiveness by designing it to account for variable stage sequence. Moreover, unlike most other control strategies which are either based on delay or queue length, the proposed method is based on both of these traffic-related metrics. The proposed method is used to regulate the traffic for a network of six intersections and its results are compared with two other control strategies including pre-timed and fully actuated.

Keywords: Multi Agent Systems, Machine Learning, Auctions, Adaptation, Traffic Signal

1 Introduction

The efficient and effective movement of people and goods is critical if society is to achieve economic prosperity, energy efficiency, environmental sustainability, global competitiveness and other objectives. Much of the transportation activity that occurs in urban areas is by highway, and much of that is via surface arterials.

The single biggest impediment to efficient operation of surface arterials is signal timing. While traffic signals are necessary to ensure the safe movement of all vehicles, their use reduces efficiency. Trips take longer than they would if the vehicle-to-vehicle conflicts could be eliminated without signal control. Thus, optimizing the signal control and making it traffic adaptive is critically important.

Off-line optimization has been standard practice for a long time. Today there are well-established algorithms such as TRANSYT [1], which generate optimal coordinated plans for fixed-time operation. The main weakness of such methods is that their plans are computed for a static situation, based on historical data. But that situation never actually exists in the network. Other systems, like SCOOT, Split Cycle and

Offset Optimization Technique [2], are similar to TRANSYT but are traffic-responsive (they use real-time data from detectors to update the signal control settings). SCATS, the Sydney Coordinated Adaptive Traffic System [3], also makes real-time decisions about the control strategy based on detector inputs. The main difference between SCOOT and SCATS is that the latter is a hierarchical system while the former is not.

The strength of systems like SCOOT and SCATS is their ability to adapt to the traffic flow conditions. However, these systems have limitations. They work with a fixed cycle length. Their coordination patterns have to repeat on a single-cycle basis, and their responsiveness is slow (tempered) to ensure stable operation. Their off-line or even on-line optimization of the signal timings has difficulty adjusting quickly to changing traffic patterns. And yet, traffic flows are highly dynamic. Thus, optimal signal timing plans are difficult to determine in advance.

PRODYN [4], OPAC [5], and UTOPIA [6] are also examples of adaptive systems that are not centralized, but their relatively complex computation and communication schemes make their deployment costly [7].

Given the growth rates foreseen for urban traffic in the future, more flexible and robust approaches are necessary. Hence, making signals smarter is the objective of this paper. The intent is to make signal control more sensitive to the evolving traffic streams and more intelligent about coordination. This will produce savings in energy consumption, pollutant emissions, and delays.

During the last few years, multi-agent systems (MAS) have become a promising application domain within artificial intelligence (AI) for optimizing traffic signal control. MAS techniques can be applied to situations where the conditions evolve dynamically. They can capture the important details at the level of individual entities and produce useful control results. They can be used in a variety of ways to emulate system behavior. They can be active, heterogeneous participants in an environment representing the system of interest and engage in information processing and decision making. Their behavior can be visualized, monitored, and validated at individual agent level, leading to new possibilities for analyzing, debugging, and developing signal control strategies.

This paper shows how the traffic interactions of an intersection can be regulated by means of auction theory, MAS, and machine learning (ML) techniques. In the proposed work, the intersection space is viewed as a spatially complex set of scarce commodities whose temporal non-overlapping usage is to be adjudicated through an interactive bidding process that involves multiple agents. A ML technique is used to optimize the bidding strategies of the agents. The proposed method offers the following features and characteristics: 1) decentralized design and operation, which is usually less expensive comparing to centralized approaches; 2) variable staging sequence, i.e., it is not hampered or constrained by a prescribed stage sequence as is common with all actuated, semi-actuated coordinated, and pre-timed control strategies which have fixed staging sequence; 3) control logic based simultaneously on queue length and delay (unlike most other control strategies which are either based on delay or queue length); 4) self-learning, i.e., decreases human intervention in the operation after implementation; 5) model-free, i.e., does not need a model of traffic pattern that is challenging to acquire; 6) robust, i.e. with no single point of failure.

The paper is organized as follows: Section 2 reviews some related research. Thereafter, Section 3 presents the details of the proposed model. Section 4 presents the experimental results that have been carried out in a network of six intersections. Finally, Section 5 gives some conclusions and points out lines of future work.

2 Related Work

MAS involve using distributed intelligence, often autonomous, to develop problem solutions. For example Choy et al. [8] present a hierarchical MAS that consist of three layers of agents: controller agents, zone controller agents, and regional controller agents. The implementation of agents is based on feed-forward neural network and fuzzy logic theories.

MAS often use ML to adapt to the evolving traffic conditions. For example, Steingrover et al. [9], Weiring [10] employ a reinforcement learning technique to minimize the overall waiting time of the vehicles. Here the learning task is represented as a feedback loop focused on the aggregated waiting times for individual vehicles. Another example is the work done by Tantawy et al. [11] in which a MAS is proposed for adaptive traffic signal control. In their proposed approach each agent (which controls one intersection) plays a game with its immediate neighbors and learns and converges to a response policy to all neighbors' policies using a reinforcement learning technique.

Some researchers suggested adapting market-based ideas to traffic signal control. Isukapti and List [12] have demonstrated that auction theory can be used as a paradigm for modeling signal control. Vasirani and Ossowski have also demonstrated that a distributed, market-inspired, mechanism can be developed for the management of an urban road network, where drivers trade with the infrastructure agents in a virtual marketplace, purchasing reservations to cross intersections [13]. Carlino et al. [14] have shown that auctions can be run at each intersection to determine the order in which drivers perform conflicting movements. In their approach autonomous vehicles (which are considered as agents) bid on behalf of the travelers. This approach has been used for an isolated intersection. But how the agents (drivers) bid is not clear and well explained. More importantly no strategy for optimizing the bids is proposed or implemented.

A comprehensive literature review of agent-based technology for transportation systems can be found in [15].

3 Methodological Approach

Following the lead of Isukapati and List [12], the intersection space is viewed as a spatially complex set of scarce commodities whose temporal non-overlapping use is adjudicated through an interactive bidding process that involves multiple agents. To further ensure safety, change intervals are included (i.e., a yellow interval followed by an all red) and minimum greens are used. No maximum greens are employed. FIFO-based discharge strategy is employed to ensure that realistic separations are maintained between vehicles as they are discharged from the stopbar.

Each movement is managed by an agent called “movement manager”. In a typical intersection this means there are eight movement managers operating in parallel as shown in Fig. 1. There is another agent called “intersection manager” that control the behavior of movement managers.

Bids occur each time the use of the intersection is contested. This arises when agents for conflicting movements have non-zero queues and transitioning to a new stage is possible. Transitions are possible at the end of the minimum green for the current stage or at the end of every ensuing discharge headway or at the end of each maximum gap time if there is no discharge headway.

Stages (phases) are formed by combinations of winning bidders. Eight stages (movement combinations) are possible for the intersection shown in Fig. 1. The stages are shown in the table in the left-hand bottom corner of the Fig. 1. First column shows the stages, and the second column shows the movement combinations. For example, stage 3 is formed by combining movement 1 and 5.

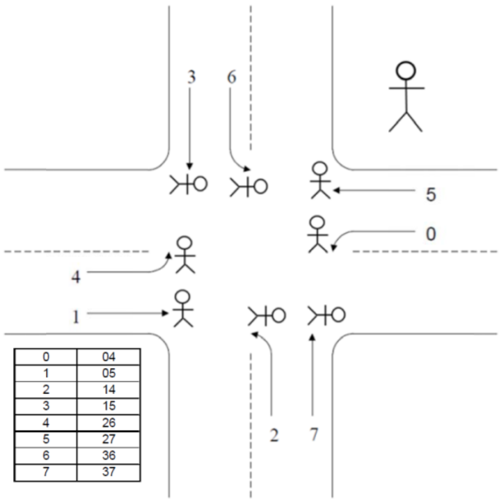

Fig. 1. Structure of the agent based modeling approach

The bidding process is as follows. Each movement manager (with a non-zero queue) submits a bid and the intersection manager sums these bids for all possible, safe, movement combinations (eight in this case). The pair of movements with the highest combined bid win. For example, if the bids are 1, 6, 2, 5, 3, 2, 4, 5 respectively for movements 0, 1, 2, 3, 4, 5, 6, 7, the combined bids would be 4 (04), 3 (05), 9 (14), 8 (15), 6 (26), 7 (27), 9 (36), 10 (37). Movements 3 and 7 would win and pays the intersection manager what was bid (first-price bidding).

To minimize the vehicle to infrastructure (V2I) dependence, consistent with the current state of the practice in rich intelligent transportation systems (ITS), movement managers receive tokens when vehicles join their queues. They use these tokens as the basis for submitting their bids and they "pay" tokens to the intersection manager when they win bids. In the simulation-based realizations presented here, the movement managers have access to information about their respective movements and limited information about the other movement managers. They know how many drivers are in queue, how many tokens they have, the bid that they submit, and what the winning bid was.

As indicated, when movement managers win following one or more prior losses, they receive use of the intersection for a minimum green. When the minimum green expires, they have to bid again to retain control at the end of each discharge headway (assuming they have a queue) or minimum gap time to see if they can continue to retain control of the green. When bids are to be submitted, those movement managers with non-zero queues are allowed to do so. Those without queues are excluded. If the movement managers currently holding control of the intersection submit winning bids, they are allowed to continue discharging vehicles for one more discharge headway or the minimum gap time (i.e., 3 seconds), whichever is smaller. If they lose control, then, then a clearance interval is imposed and control shifts to the new winners for a minimum green.

Reinforcement learning has been used to help the movement managers improve their bidding strategy based only on the knowledge of their own past received payoffs. Specifically, a Q-learning approach is used to create an optimal bidding strategy for each movement manager.

The core of the Q-learning algorithm is a Q-table and an algorithm for updating the Q-table and choosing actions. A Q-table *Q(s, a)* is a matrix indexed by state *s* and action *a*, which is the expected discounted reinforcement of taking action *a* in state *s*. At each time, an agent is assumed to be in a certain state *s*, and it chooses an action *a* according to the Q-table and other algorithms to interact with the environment. Then the agent receives a reward *r* from performing action *a* and observes a new state *s'*. After that, the Q-table is updated by the following equation:

$$Q(s,a) = (1-\alpha)Q(s,a) + \alpha(\mathrm{r} + \gamma \max_{a'}(\mathrm{s}', \mathrm{a}')) \quad (1)$$

Where α is the learning rate, and γ is the discounting factor. Q-Learning can be summarized in following steps:

1. Let the current state be *s*
2. Choose an action *a* to perform
3. Receive a reward *r* from performing action *a* and the resulting state *s'*.
4. Update *Q(s,a)* using equation (1).
5. Go to step 1).

The design elements of the Q-learning in terms of the typical structure of it for each movement agent are:

- *State*: The current queue
- *Action*: the bid amount to submit

- *Reward*: the immediate reward is composed of two terms. The first term is a function of the delay and is defined as the change (saving) in the total cumulative delay, i.e., the difference between the total cumulative delays of two successive decision points (bidding cycles). The second term is a function of winning the bid. It is zero if the movement agent loses the bid. On the other hand, if the movement agent wins the bid, the reward is based on the difference between the number of vehicles discharged and the submitted bid. For example, if the movement agent wins the bid and receives a minimum green (e.g., 6 seconds), the average headway is 2 seconds, and the submitted bid was 1.5 tokens, then the movement agent is able to discharge 3 vehicles (6/2) and the reward would be 1.5 (3-1.5).

Movement managers use an ϵ-greedy selection approach in order to explore the state and action environment. That is, random actions are selected with probability ϵ and the actions with the highest Q-values are chosen with a probability of (1-ϵ). The exponential function (e^{-En}) is used for ϵ-greedy approach in which E is a constant and n is the number of iteration. The exponential function is a simple ϵ-greedy exploration approach with a gradually decreasing rate of exploration. By using this approach, at the starting the agent mainly explores, as it has no previous information to exploit, and the agent moderately increases the degree of exploitation towards the end of the learning process. Gradual shifting is necessary to ensure that the entire state space is covered during the learning process. As suggested in the literature [16] the learning rate is considered as follows:

$$\alpha^k = \frac{E}{v^k(s,a)} \qquad (2)$$

Where α^k is the learning rate at time k, E is a constant, and $v^k(s, a)$ is the number of visits to a particular state-action pair (s, a). This learning rate allows to start with high learning rate at first to allow for fast modifications then use lower rates as time progresses.

The overall pseudo code for the model is summarized in Fig. 2.

Input:

- δt = 0.1 (initialize time step)
- t = 0 (simulation clock)
- g_{min} = 6.0 (minimum green)
- c = 4.0 (change interval, i.e., yellow and all red intervals)
- h_d= 2.0 (headway)
- $f_{\min}$ = 1.0 token (initial fee for each vehicle)

Initialize: for each movement manager initialize Q values as zeros

The algorithm:

```
while t < T (simulation time period):
  t = t + δt
  for each movement manager
     if <new arrivals = True>
        collect f_min
        add new arrivals to service queue
  if the use of intersection is contested
     for each movement manager who submitted bid in previous bidding cycle
        compute the total cumulative delay
        compute the reward
        update the Q values
     for each movement manager with non-zero queues
        if explore (ϵ) then
           submit random bid (action) for the current queue (state)
           (bid is subject to { bid_min = 0.5*queuLength , bid_max = (bankBalance/queueLength) * (g_min/h_d) })
        else
           submit bid according to the policy
     intersection manager sums submitted bids for all possible, safe, movement combinations (eight)
     the pair of movements with the highest combined bid will be selected
     wining movement managers pays the intersection agent amount equal to winning bid
     if < current winner = previous winner >
        intersection manager extend green by 3 seconds
     else:
        intersection manager impose a clearance time of 4 seconds
        intersection manager allocate intersection control to new winners for t = t + g_min
```

Fig. 2. Pseudo Code for control structure of simulation

4 Case Study

Fig. 2 shows the network of intersections which has been modeled (each intersection simply applies the control strategy presented in the previous section). The east-west arterial has four intersections (0, 1, 2, and 3) while the north-south arterial has three (4, 2, and 5). Intersection 2 is common between them. Left turn bays exist everywhere. The figure also shows the travel distances between the intersections. The longest distance is 800 feet and the shortest is 550 feet with the nominal travel speed at 30 mph.

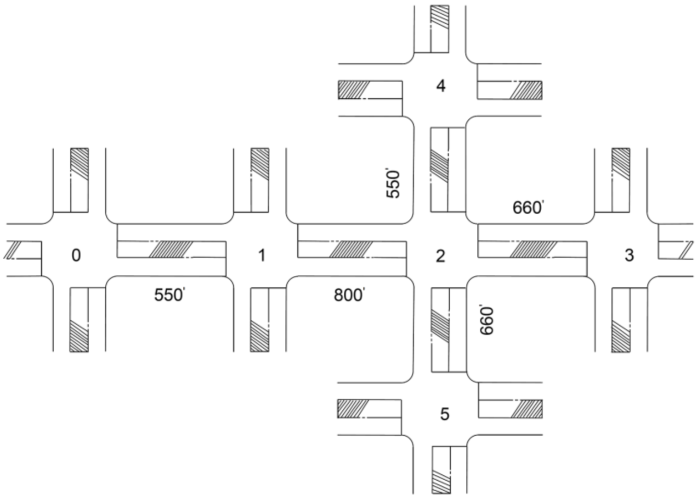

Fig. 3. Modeled network of intersections

Two different combinations of traffic volume shown in Table 1 are considered in this paper. Case 1 represents a “nominal” or baseline condition for the network. The heavier flows are east-west, with 600 vehicles per hour (vph) entering the network in these two directions. Side street volumes along the east-west arterial are 200 vph. The north-south arterial has slightly less traffic, with main flows at 400vph, both north and southbound. The side street volumes are 150 vph. In all cases, for simplicity, the turning percentages are 15% left turns and 10% right turns.

The other case makes adjustments to the baseline values. All of the volumes increase by 25%. Thus, the street volumes grow to 750 vph east-west and 500 vph north-south.

Table 1. Volume combination for each intersection.

Approach	Intersection 0	1	2	3	4	5
			Case 1			
North	200	200	-	200	-	**400**
South	200	200	-	200	**400**	-
East	**600**	-	-	-	150	150
West	-	-	-	**600**	150	150
			Case 2			
North	250	250	-	250	-	**500**
South	250	250	-	250	**500**	-
East	**750**	-	-	-	188	188
West	-	-	-	**750**	188	188

Three signal control strategies are explored. These are: pre-timed, fully actuated, and the proposed control strategy. The three of these control strategies are implemented in the custom-built traffic simulator developed by the authors.

For the pre-timed control strategy, Synchro7 [17] which is an analysis and optimization program for optimizing signal timing for arterials, was used to develop optimal signal timings. The obtained signal timing plans including cycle lengths, splits, phase sequences and offsets for each intersection and in each case were implemented in the custom-built traffic simulator.

For fully actuated control strategy, the implemented logic is as follows. Normal stage rotation is followed including the feature that in the absence of any requests for service, the controller returns to the main street stage. This means the user needs to specify a main street stage and minimum and maximum greens for each movement. In the implemented logic here, for intersections 0, 1, 2, and 3, stage 3 is considered the main street stage, and for intersections 4 and 5, stage 7 is considered as the main street stage. Minimum green is 6 seconds and maximum green is 40 seconds.

Fig. 3 and Fig. 4 show the results of the simulation for two traffic volume cases. As can be seen from the figures, the proposed control strategy improves as time passes which make sense. Initially the proposed control strategy does a lot of exploration and as the time passes it learns the best control policy and does more exploitation.

In case 1 (Fig. 3) which represents a low traffic volume, pre-timed control strategy is the poorest and the proposed control strategy is the best and actuated control strategy performs in the middle. Again, the results make sense, imposing a pre-defined control strategy (pre-timed) when the traffic volume is low is not a good strategy.

In case 2 (Fig. 4) which represent a heavy volume of traffic, pre-timed control strategy is performing better than the other two control strategies, the proposed strategy is the second and fully actuated has the poorest performance. Again, the results make sense since when there are heavy traffic, it is better to impose some pre-defined timing plan to regulate the traffic.

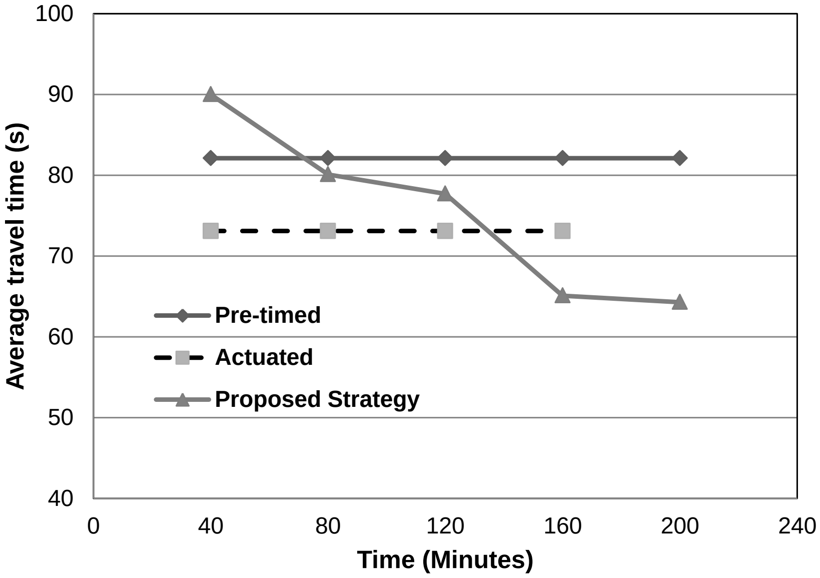

Fig. 4. Results of case 1

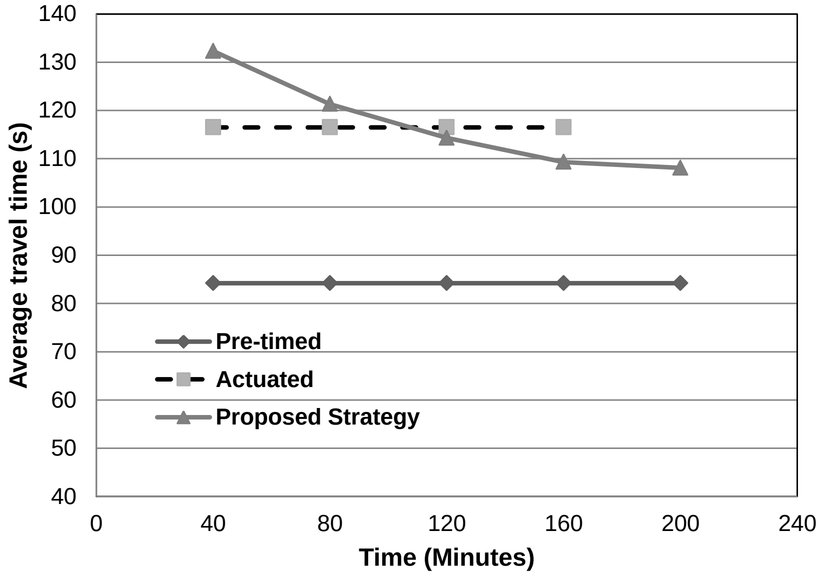

Fig. 5. Results of case 2

5 Conclusion and Recommendation

In this paper a new control strategy is presented which integrates MAS, ML technique, and auction theory to regulate the movement of cars at an intersection. This multi-tiered agent-based framework provides a way to decompose the intersection into smaller sub-parts. It increases the efficiency of the model, provide a way to create a parallel computing realization, and enhance scalability.

The authors believe that this area of research is one that has significant promise for the future, especially in light of the increasing demands for more features and capabilities, especially real-time control, being placed on advanced traffic management systems.

There are various venues for further research. Our own future research will concentrate on:

- applying other exploration functions other than ϵ-greedy such as softmax and ϵ-softmax
- second-price bidding is another area of future research. In this bidding strategy, truthful bidding is a dominant strategy. The fact that truthfulness is a dominant strategy also makes second-price auctions conceptually very attractive.
- in the current model, the intersection managers are not cooperating, so the other area of future research is to make intersection mangers cooperative and see how it affects the performance of the whole network.
- linkage of the developed signal control logic to a commercially available traffic simulator (e.g., VISSIM [18], Paramics [19]) to test the performance of the developed control logic for traffic scenarios similar to real world cases and compare its results with other adaptive controller such as SCOOT and SCATS.

References

1. Robertson, D.: TRANSYT method for area traffic control. In: Traffic Eng. Control, vol. 10, pp. 276–281 (1969)
2. Robertson, R., Bretherton, R. D.: Optimizing networks of traffic signals in real time - the SCOOT method. In: IEEE Trans. Veh. Technol., Vol. 40, pp. 11-15 (1991)
3. Sims, A.G., Dobinson, K.W.: The sydney coordinated adaptive traffic (SCAT) system philosophy and benefits. In: IEEE Trans. On Vehicular Technology, 29(2), pp. 130-137 (1980)
4. Henry, J., Farges, J. L., Tuffal, J.: The PRODYN real time traffic algorithm. In Proceedings of the International Federation of Automatic Control (IFAC) Conference, Baden-Baden: IFAC (1983)
5. Gartner, N. H.: OPAC—A demand-responsive strategy for traffic signal control. Transportation Research Record, 906, pp. 75–81 (1983)
6. Di Taranto, M.: UTOPIA. In Proceedings of the IFAC-IFIP-IFORS Conference on Control, Computers, Communication in Transportation, Paris. ifac (1989)
7. Bazzan, A.L.C.: Opportunities for multiagent systems and multiagent reinforcement learning in traffic control. Autonomous Agents Multi-Agent Syst., vol. 18, no. 3, pp. 342–375, Jun (2009)

8. Choy, M.C., Srinivasan, D., Cheu, R. L.: Cooperative, hybrid agent architecture for real-time traffic signal control. In: Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on 33.5, pp. 597-607 (2003)
9. Steingrover, M., Schouten, R., Peelen, S., Nijhuis, E., Bakker, B.: Reinforcement Learning of Traffic Light Controllers Adapting to Traffic Congestion. In: BNAIC (2005)
10. Wiering, M.A., Veenen, J., Vreeken, J., Koopman, A.: Intelligent traffic light control." Institute of Information and Computing Sciences. Utrecht University (2004)
11. El-Tantawy, S., Abdulhai, B., Abdelgawad, H.: Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (marlin-atsc): Methodology and largescale application on downtown toronto. In: Intelligent Transportation Systems, IEEE Transactions on 14(3), pp. 1140–1150 (2013)
12. Isukapati, I.K., List, G. F.: Agent Based Framework for Modeling Operations at Isolated Signalized Intersections. In: Elsevier, Procedia - Social and Behavioral Sciences (2013)
13. Vasirani, M., Ossowski, S.: A computational market for distributed control of urban road traffic systems. In: IEEE Transactions on Intelligent Transportation Systems 12(2), pp. 313–321 (2011)
14. Carlino, D., Boyles, S.D., Stone, P.: Auction-based autonomous intersection management. In: IEEE ITSC'13 (2013)
15. Bazzan, A.L.C., KLUGL, F.: A review on agent-based technology for traffic and transportation. In: The Knowledge Engineering Review, vol. 29:3 (2013)
16. Gosavi, A.: Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning. In: Springer, Netherlands (2003)
17. Husch, D., Albeck, J.: Trafficware SYNCHRO 7 User Guide, Trafficware, Albany, Calif, USA (2006)
18. PTV Plannung Transport Verkher AG.: PTV, VISSIM 5.40-03 User Manual, Karlsruhe, (2012)
19. Quadstone Paramics.: Paramics Microscopic Traffic Simulation Software, http://www.paramics-online.com