Title: A Decentralized Network Level Adaptive Signal Control Algorithm by Deep Reinforcement Learning

Yaobang Gong*
Ph.D. student
Corresponding Author
Department of Civil, Environmental, and Construction Engineering
University of Central Florida
Orlando, Florida 32816-2450
Email: gongyaobang@knights.ucf.edu

Mohamed Abdel-Aty, Ph.D. PE
Professor and Chair
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Orlando, Florida 32816-2450
Email: M.Aty@ucf.edu

Qing Cai, Ph.D.
Post-Doc Associate
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Orlando, Florida 32816-2450
Email: qingcai@knights.ucf.edu

Md Sharikur Rahman
Ph.D. Candidate
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Orlando, Florida 32816-2450
Email: sharikur@knights.ucf.edu

ACKNOWLEDGMENTS
The authors are grateful to the Florida Department of Transportation (FDOT) for supporting this study. All opinions are those of the authors and not necessarily reflect FDOT’s position. The authors also appreciate the data provided by FDOT and Seminole County.
1 Introduction

Increasing travel demand which is often beyond the capacity of current roadway network leads to inevitable congestion. Thus, improving the efficiency of the traffic management, especially the control of traffic signals are necessary. Many adaptive traffic signal control systems (ATCS) were developed accordingly. Among them, many researchers proposed ATCSs utilizing one of the artificial intelligence (AI) algorithms, reinforcement learning (RL), to let “machines” to learn how to control traffic signals (1). Recently, thanks to the rapid development of deep learning, ATCSs with deep reinforcement learning (DRL) algorithms were developed by several researchers (2–8).

While those studies have shown the great potential, there are two major limitations:

(1) ATCSs based on DRL are shown be effective for isolated intersections, its effectiveness of network-wide signal control, especially for coordinated signals, is not proven.

(2) To the best of the author’s knowledge, all the studies about ATCSs based on DRL are evaluated by ideal simulated traffic scenarios. The intersection geometric design such as length of the turning storage bays is not considered. The traffic demand is hypothetical and the route choice behavior is simplified. The benchmark signal control system is naïve fixed-timing signal, signal controlled by conventional RL algorithm incorporating with a shallow neural network which is never be applied in the field, or even random signals.

To fill the gap, this study proposed a network-level decentralized adaptive signal control algorithm based on recent developed DRL algorithm, double dueling deep Q network. It allows the coordination among nearby intersections by information sharing. The proposed algorithm was trained by a simulated AM peak scenario of a suburban traffic corridor calibrated by real-world data. Finally, its performance was compared with the real-world coordinated actuated signal system whose configuration is provided by local jurisdiction.

2 Methodology

2.1 Decentralized Adaptive Signal Control Algorithm Based On Double Dueling Deep Q Network (3DQN)

Reinforcement learning (RL) is a goal-oriented machine learning algorithm, which learns to achieve a complex goal over many discrete steps by interacting with the environment. For a control problem, in every discrete control step, an RL control agent (e.g. signal controller) iteratively observes the state of the environment (e.g. roadway network), takes an action (e.g. directly change the signal phase) accordingly based on underlying policy π, receives a feedback reward r (e.g. waiting time) which will be accumulated to its long-run goal from the environment for the action taken, and adjusts its policy until it converges to the optimal mapping from the set of the all possible states S to the set of all possible actions A, which is the optimal policy π∗

In this study, a decentralized signal control problem is formulated into the standard RL problem. Each signal controller and its associated sensors acts as a RL agent. The agent observes the condition of the controlled intersection and its adjacent two intersections as the state. The state includes the detailed location of every vehicle occupying the roadway within a certain distance from the stop line and the current phase of the signal. The agent directly selects the appropriate phase every discrete step as its action. The discrete signal performance metrics, which is the waiting time of the vehicles in the queue, as the reward or penalty of the state-action pair. The long run goal of the system is reducing the cumulative delay/travel time.
One of the DRL algorithms, Double Dueling Deep Q Network (3DQN), is employed. The 3DQN has the capability to handle the very detailed representation of the states used in this study while the conventional RL does not. Convolutional neural network (CNN), one of the deep neural networks (DNN) was used to process the state representation.

2.2 Experiment
The proposed algorithm was implemented in a simulated traffic network using a commercial traffic simulator Aimsun Next 8.2.3. To better examine the performance of the algorithm, the simulation scenario was built based on a real-world suburban traffic corridor in Seminole County, Florida. There are one limit-access freeway (SR 417) and several parallel or connected arterials (SR 426, Red Bug Lake Road, Slavia Road, West Chapman Road and Dead road) within the corridor. Figure 1 shows the location of the corridor.

FIGURE 1 Location of the Simulation Test Site
Along with the corridor, there are eight signals in total under control of the proposed algorithm. All eight signals are under the control of coordinated actuated signals. These actuated signal controllers are used as the benchmark to evaluate the performance of the proposed algorithm. The demand data of the AM peak hours (7:00-9:00), which is extracted from Orlando Urban Area Transportation Study (OUATS) with base year 2009, were used as the input to simulate the recurrent congestion situation. And the actual turning movement counts and travel time is used to calibrate and validate the scenario.

The algorithm is firstly pre-trained by a supervised learning method using a fixed signal timing to speed up the training process. Then the algorithm is trained in episodes, which are simulation replication with the same aforementioned two-hour AM congested traffic demand by in different random states (The travel demands data only controls the trip generation. The randomness determine the vehicles’ departure times. And the route of a specific vehicle is determined by its route choice function).

There are totally 20 training episodes were conducted and the algorithm converged after around 10 episodes.

3 Findings

To evaluate the performance of the proposed ATSC algorithm, both the well-trained algorithm and the real-world benchmark signal control were implemented in a simulation replication with the same random state.

During the whole simulation episode, the average travel time and average delay of the simulated network controlled by the proposed algorithm are less than those of the network controlled by the benchmark signal. In average, the proposed algorithm reduced 10.27% (4.26 minutes versus 4.75 minutes) of travel time and 46.46% (11.27 second/mile versus 21.06 second/mile) of the total delay.

However, given the number of the stops remains at a low level, the proposed ATSC algorithm tends to increase the number of stops by 11.29% averagely (0.31 versus 0.28). The reason might be the algorithm tends to allow more green time for minor approaches and turning movements of the major approaches. On the one hand, this ensures the fair travelling rights between major and minor approaches as well as trough and turning movement. On the other hand, the vehicles travelling through the major approaches are forced to stop to yield the right of way. Since the algorithm does not take number of the stops as its goal, the practitioners could easily use number of stops as a part of the reward (in the RL setting) if the through movements of major approaches are preferred.

4. Conclusion

In this paper, a decentralized adaptive signal control algorithm using double dueling deep Q network is proposed for network level signal control. The proposed algorithm captures detailed locations of vehicles, extracts relevant information by convolutional neural networks, and selects appropriate signal phases every second to reduce vehicle’s waiting time. The algorithm was evaluated by the real-world coordinated actuated signals in a simulated real-world suburban traffic corridor in Seminole County, Florida. The evaluation results showed that the algorithm reduces 10.27% of travel time and 46.46% of delay compared with the real-world benchmark. Meanwhile, it ensures the fair travelling right for all movements.
Reference


