Transfer Learning Acceleration for Reinforcement Learning of Traffic Light Control Policy

Reid Christopher, Woo Maeng, and Masafumi Endo

I. Introduction

The first modern electric traffic light was introduced in early 20th century, and it was operated manually by human from a tower in the street [2]. It was developed due to the chaos on roads caused by the disorder of cars, bicycles, horses, wagons and pedestrians without any regulations. There are no horses or wagons on the road anymore, but instead the number of vehicles on roads has dramatically increased. Traffic light systems have helped to alleviate this disorders on roads. However, conventional traffic light systems still have significant room for improvement, as many are have static control schemes regardless of the situation. As a result, traffic congestion negatively affects modern civilization. For example, traffic congestion cost the British, French, German and American economies over $200 billion in 2013 [3]. Instead of the conventional system that either deploys fixed time/sequence without considering the real-time situation, more adaptable, intelligent traffic light control strategies are necessary for optimizing complex traffic flows.

One of the ways to solve this problem is using methods in the field of artificial intelligence, particularly reinforcement learning. This project aims to produce a reinforcement learning (RL) based approach that optimizes traffic light control at an intersection. Especially, we plan to devise a time-efficient optimization strategy by combining transfer learning and RL. The main motivation is that earlier adaptation of the optimization policy to the real world can avoid traffic congestion caused by poor decisions made early in the learning procedure. In order to solve the problem that current methods are facing, we plan to generate a traffic light system, specifically for a 4-way intersection using transfer learning and reinforcement learning that is more efficient than current solutions in practicality. Data for training our system will be collected by running simulations adjusting parameters using Simulation of Urban Mobility (SUMO) software [4] . The system will first be trained to mimic typical traffic light behavior before being allowed to explore other option with reinforcement learning. Inputs for our system will be the position and speed of cars in the system as well as information about the current state of the traffic light. Using these inputs, our system will output the next state the traffic system should transition to.

II. Background

Traffic lights have been around us for almost 100 years and it is well known that traffic lights increase the safety and improves the traffic flow [1] . However, people still spend a lot of unnecessary time on roads waiting for green lights. It was estimated that congestion caused urban residents to travel an extra 6.9 billion hours and consequently 3.1 billions of gallons of fuel in United States in 2014 [3]. One effective way to ease the traffic congestion is to expand the transportation capacity and build higher quality road to deal with higher traffic flow. However, this solution imposes us to spend a significant amount money and time. Hence, using existing road that improves adaptability is a more efficient way for our society. Two of the most popular traffic light control systems currently being used are fixed time traffic light control systems and basic dynamic control systems. Fixed time traffic light uses a timer and changes the light color when the set time is up. This has people stop and wait even if there are no other cars on the road, making it especially ineffective during light traffic flow. Dynamic control system uses detectors/sensors to detect cars on roads or intersections and changes the color with respect to the information acquired from detectors to adjust the traffic flow [8]. It is much more efficient than the fixed time system, however, it still lacks adaptability to changes in traffic. Traffic is a generally unpredictable phenomena. While certain general statements about the levels of traffic at various times tend to hold true, minor differences are always present from moment to moment. It is affected by many variables, such as the general stochastic nature of human decision making, events in the area, accidents, or even the weather. There has been many cases that tried to overcome these challenges by using neural networks, reinforcement learning, and other artificial intelligence methods [9]. Among them, reinforcement learning (RL) has been attracting attention to solve traffic light optimization problem. RL has abilities to 1) adapt to different situations that are caused by weather conditions, car accidents, and special events 2) learn to optimize traffic flow that is difficult to give a correct answer without supervision by human experts 3) model environment on its own, since an agent learns every variable of the environment through learning procedure. Although these advantages have been pushing forward to use RL for traffic light control problems and research works so far successfully eased traffic congestion in simulation environments, these methodologies still do not apply to the actual traffic lights. One of the problems to practically use RL for traffic light optimization is that these methods require huge amount of data needed for training making effective implementation in reality difficult [9].

III. Related Work

A. Neural Networks

Neural networks are a representation of functional mappings between inputs and outputs that have gained increasing popularity in the last decade. They propagate inputs through some number of computational layers, with the output of each layer fed into the input of the next before eventually generating an output. A key feature of these networks is their ability to represent a function of arbitrary complexity. The universal approximator theorem states that a neural network can be constructed to approximate *any* function within a given finite error [10]. This ability to represent any function removes the need to make assumptions about the form of the mapping between the inputs and outputs that is desired to be learned. Given that the mapping between the traffic state and best traffic control choices is highly complex, neural networks pose a convenient tool to use for optimal traffic control. Furthermore, algorithms have been developed that enable the parameters of neural networks to be tuned to best represent the function we are attempting to discover. We can take advantage of these algorithms as we teach our neural network controller.

B. Transfer Learning

Transfer learning is a subsection of machine learning where information learned through a similar task is then applied to a different task. If the two tasks are similar enough, then the speed of learning the second task can be increased over starting from scratch. We hope to utilize transfer learning by starting with supervised learning to encode typical traffic light behavior,before using that starting point to explore for an optimal control policy using reinforcement learning. Supervised learning utilizes the training of a learning agent with examples to which the desired output is known. This means we can encode a specific behavior in a neural network with supervised learning before allowing the network to explore alternative behaviors with reinforcement learning.

C. Reinforcement Learning

Reinforcement learning is a branch of machine learning that attempts to have the learning agent learn from experience and feedback from the system. This kind of learning works in problems where the desired output for a given input is not known, and must be found through experience. This differs from supervised learning, where the desired outputs are known, allowing an error to be used in adjusting the learner. A agent learning under reinforcement learning is given a reward for the actions it takes. From these rewards, the learner attempts to learn an overall value of the states (or (state, action) pairs) within its state space. It then takes the actions that it believes will generate the maximal reward. Reinforcement learning is highly applicable to the problem of optimal traffic control because the optimal solution is unknown, and reinforcement learning provides an option to explore avenues of control that may not be looked at in human generated optimizations.

D. Literature Review

Intelligent traffic light control methods have been studied for improving conventional traffic systems to be more adaptable to dynamically-changing traffic flows in real world. Chiu et al. [11] presented an optimization method for traffic signal timing using fuzzy logic. Schutter et al. [12] proposed a model that describes the evolution of the queue length at an intersection for solving traffic light switching problems. These studies, however, model road traffic by limited information so that these approaches cannot be applied in complex, realistic traffic control problems.

Learning-based traffic light control has been actively studied to optimize dynamic, uncertain traffic situations in the real world. Liu [13] summarized intelligent traffic signal control methods until 2006, and RL is introduced as one of the effective optimization technologies. During this period, however,RL methods were still under development, such as using a linear function being usually used to estimate the Q value. In that period, a small-sized state space was mainly applied for RL. Abdulhai et al. [14] presented an intelligent transportation system by Q-learning that represents state space based on the number of waiting vehicles. Balaji et al. proposed multi-agent traffic signal control by RL that makes state space using vehicle occupancy and the statics obtained from historical traffic data. The difficulty of traffic light control by RL in these methods is that the dynamic, complex urban traffic systems cannot be represented by the limited information.

Recent technological advancements in the field of artificial intelligence have been pushing forward to use both deep learning and reinforcement learning as deep-reinforcement learning to estimate Q-value. Several research works apply deep-reinforcement learning to control traffic light in uncertain, dynamic environments. Li et al. [16] used deep reinforcement learning to find appropriate signal timing policies. Pol et al. [17] presented a traffic light control approach by deep Q-learning to optimize traffic flow in the cross-shape intersection. The state of traffic situation is represented as an image matrix, and reward is designed by combining vehicles' and traffic lights' information such as waiting time and changing light. Liang et al. [18] applied a deep-reinforcement learning model consisting of several components such as dueling network and double Q-learning network to improve optimization performance. Gao et al. [19] proposed an algorithm that automatically extracts useful features for deep-reinforcement learning to learn optimal policy, instead of using human-created features. The main motivation of this approach is to gain useful information from raw data that is sometimes ignored when human experts arbitrarily select information. Above all the research works based on deep reinforcement learning show better performance in terms of optimizing traffic flow with complexity.

However, there are still problems in accomplishing optimized traffic light control using reinforcement learning. The first problem is the design of a reward that is suitable for optimizing traffic flow. The objective of the traffic light to optimize traffic flow is to minimize total travel times of vehicles on roads. However, the travel time of vehicles cannot be obtained until they pass the roads; hence, it leads to the problem of delayed rewards. Instead of this rewards, the change of the cumulative waiting time and combination of cars' rewards such as sudden braking and de-acceleration and the lights' rewards [18] [17]. Appropriate design of rewards is one essential aspect of the traffic light control learning that is not fully solved.

The next problem comes from the difficulty to deal with large-scale traffic flows such as multiple intersections where is more complex and difficult to derive a traffic light policy. Although many studies aim to learn a traffic light policy for one intersection, it is necessary to apply these control methods in large-scale traffic systems to use them practically. In such a cases, multi-agent learning is used to optimize multiple intersections. However, undesired learning will occur due to other intersections' learning and optimization results: if these traffic lights learn and update at the same time, each policy during other traffic lights' learning processes negatively affects the consequence of global optimization. One approach to this problem is to use transfer learning to enable stable learning by copying one traffic light's learning results in other ones [17]. In order to implement RL-based traffic light control into the real world, scalability should be investigated in detail.

The third problem is that current RL-based approaches do not consider efficiency of learning processes. While these methods perform better results to solve traffic congestion, the learning cost such as too long running time is another undesirable factor for optimization in the real world problem. During trial-and-error process of RL, vehicles in real world are caught in traffic jam in the traffic light control problem. Hence, traffic light control policies are required to quickly adapt the real world situations. In order to solve this problem, time-efficient optimization method by RL is required such as learning from limited data samples, efficient exploration methods, and implementing it with transfer learning [9] .

REFERENCES

[1] Nelson, M. K. (2018, May 1). A Brief History of the Stoplight. Retrieved from https://www.smithsonianmag.com/innovation/brief-history-stoplight-180968734/.

[2] History.com Editors, (2009, November 13) "First electric traffic signal installed," *A & E Television Networks*, Retrivied from https://www.history.com/this-day-in-history/first-electric-traffic-signal-installed.

[3] The Economist. 2014. The cost of traffic jams. https://www.economist.com/blogs/economist-explains/2014/11/economist-explains-1, November, 2014.

[4] K. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of sumo-simulation of urban mobility," *International Journal on Advances in Systems and Measurements*, Vol. **5**, No. 3&4, pp. 128–138, 2012.

[5] D. Schrank, B. Eisele, T. Lomax, and J. Bak, "2015 urban mobility scorecard," 2015.

[6]

[7] P. Lowrie, "SCATS–a traffic responsive method of controlling urban traffic," *Roads and Traffic Authority*, Sydney, Australia, 1990.

[8] P. B. Hunt, D. I. Robertson, R. D. Bretherton, and M. C. Royle, "The SCOOT on-line traffic signal optimization technique," *Traffic Engineering & Control*, Vol. **23**, No. 4, pp. 190–192, 1982.

[9] H. Wei, G. Zheng, V. Gayah, and Z. Li, "A survey on traffic signal control methods," *arXiv preprint arXiv: 1904.08117v2*, 2019.

[10] K. Hornik, M. Stinchcombe, and H. White "Multilayer feedforward networks are universal approximators," *Neural Networks*, 1989.

[11] S. Chiu and S. Chand, "Adaptive traffic signal control using fuzzy logic," *Proceedings. The First IEEE Regional Conference on Aerospace Control Systems*, Westlake Village, CA, USA, 1993, pp. 122–126.

[12] B. De Schutter and B. De Moor, "Optimal traffic light control for a single intersection," *European Journal of Control*, Vol. **4**, No. 3, pp. 260–276, 1998.

[13] Z. Liu, "A survey of intelligence methods in urban traffic signal control," *IJCSNS International Journal of Computer Science and Network Security*, Vol. **7**, No. 7, pp. 105–112, 2007.

[14] B. Abdulhai, R. Pringle, and G. J. Karakoulas, "Reinforcement learning for true adaptive traffic signal control," *Journal of Transportation Engineering*, Vol. **129**, No. 3, pp. 278–285, 2003.

[15] P. G. Balaji, X. German, and D. Srinivasan, " Urban traffic signal control using reinforcement learning agents," *IET Intelligent Transport Systems*, Vol. **4**, No. 3, pp. 177–188, 2010.

[16] L. Li, Y. Lv., and F. Y.Wang, "Traffic signal timing via deep reinforcement learning," *IEEE/CAA Journal of Automatica Sinica*, Vol. **3**, No. 3, pp.247–454, 2016.

[17] E. van del Pol and F. A. Oliehoek, "Coordinated deep reinforcement learners for traffic light control," *NIP's Workshop on Learning, Interfaces and Control of Multi-Agent Systems*, 2016.

[18] X. Liang, X. Du, G. Wang, Z. Han, "A deep reinforcement learning network for traffic light cycle control," *IEEE Transactions on Vehicular Technology*, Vol. **68**, No. 2. pp. 1243–1253, 2019.

[19] J. Gao, Y. Shen, J. Liu, M. Ito, and N. Shiratori, "Adaptive traffic signal control: deep reinforcement learning algorithm with experience replay and target network," *arXiv preprint arXiv: 1705.02755*, 2017.