Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Hierarchical multi-agent control of traffic lights based on collective learning

Junchen Jin [a,*], Xiaoliang Ma [a,b]

[a] System Simulation & Control (S2CLab), Department of Transport Science, KTH Royal Institute of Technology, Teknikringen 10, Stockholm 10044, Sweden
[b] iTekn Solutions, Stockholm, Sweden

ARTICLE INFO

Keywords:
Hierarchical model of traffic system
Multi-agent traffic light control
Decentralized system
Learning-based control
Collective machine learning

ABSTRACT

Increasing traffic congestion poses significant challenges for urban planning and management in metropolitan areas around the world. One way to tackle the problem is to resort to the emerging technologies in artificial intelligence. Traffic light control is one of the most traditional and important instruments for urban traffic management. The present study proposes a traffic light control system enabled by a hierarchical multi-agent modeling framework in a decentralized manner. In the framework, a traffic network is decomposed into regions represented by region agents. Each region consists of intersections, modeled by intersection agents who coordinate with neighboring intersection agents through communication. For each intersection, a collection of turning movement agents operate individually and implement optimal actions according to local control policies. By employing a reinforcement learning algorithm for each turning movement agent, the intersection controllers are enabled with the capability to make their timing decisions in a complex and dynamic environment. In addition, the traffic light control operates with an advanced phase composition process dynamically combining compatible turning movements. Moreover, the collective operations performed by the agents in a road network are further coordinated by varying priority settings for relevant turning movements. A case study was carried out by simulations to evaluate the performance of the proposed control system while comparing it with an optimized vehicle-actuated control system. The results show that the proposed traffic light system, after a collective machine learning process, not only improves the local signal operations at individual intersections but also enhances the traffic performance at the regional level through coordination of specific turning movements.

1. Introduction

Most people living in a populated area suffer from traffic congestion problems as traffic consumes time, energy, and patience. Traffic light control (TLC) is a crucial traffic management instrument in urban areas. For several decades, researchers have been using mathematical and computational methods to facilitate efficient traffic signal operations. A conventional approach in TLC planning applies an off-line optimization model to find the most appropriate signal parameters according to historical traffic observations (e.g., Ma et al., 2014). Although it is still a common practice in traffic engineering, the off-line approaches are limited because overall traffic patterns are stochastic and time-varying and historical data cannot adequately capture real-time traffic situations.

Along with the evolution of new concepts and technologies, some adaptive TLC systems have been proposed to address the issues in conventional TLC systems. Adaptive TLC systems normally adjust control parameters in accordance with real-time traffic patterns. For instance, SCOOT (Hunt et al., 1982), SCATS (Sims and Dobinson, 1980), RHODES (Mirchandani and Head, 2001) and TUC (Boillot et al., 2006) are examples of adaptive traffic signal systems that have been implemented in real applications. An adaptive TLC system can be further classified as a centralized system or a decentralized system.

In a centralized system, traffic light controllers are managed through a traffic control center that monitors traffic networks and performs optimization techniques to utilize the existing infrastructure better. Despite the fact that the ultimate goal of the centralized approach is to optimize the system-wide performance measures, its efficiency is questionable. Three issues of robustness, scalability, and efficiency were raised in previous studies for large-scale systems or networks with complex structures (e.g., El-Tantawy and Abdulhai, 2013). Conversely, each intersection in a decentralized system operates individually and autonomously. Due to the relative ease of implementation

* Corresponding author.
E-mail address: junchen@kth.se (J. Jin).

https://doi.org/10.1016/j.engappai.2017.10.013
Received 8 January 2017; Received in revised form 23 August 2017; Accepted 15 October 2017
Available online xxxx
0952-1976/

Table 1
A summary of typical studies that apply RL-based multi-agent modeling frameworks to TLC systems.

Literature	Agent	RL model	State	Action	Reward
Bazzan et al. (2010)	Intersection	Q-learning	Three-value state according to vehicle loading	One of three pre-defined signal plans	Average queue length in all links
Arel et al. (2010)	Intersection	Q-Learning with neural network for function approximation	Total delay of vehicles in a lane divided by average delay at all lanes	One of pre-defined available phases	Variation in travel delay
Balaji et al. (2010)	Intersection	Q-learning	Occupancy ratio, local traffic variations, and neighboring states	Green time for each phase	Variation in queue length
El-Tantawy et al. (2013)	Intersection	Model-based Q-learning (MARLIN)	Index of current phase, elapsed time, queue length associated with each lane	One of pre-defined available phases	Variation in travel delay
Abdoos et al. (2014)	Bottom level: intersection; Top level: region	Bottom level: Q-learning; Top level: Q-learning with tile coding for function approximation	Bottom level: ranks determined by average queue lengths; Top level: average queue length of all links inside the region	Bottom level: portion of green time for phases Top level: one of three restrictions for agents at the bottom level	Average queue length in all links
Khamis and Gomaa (2014)	Vehicle	Model-based Q-learning	Status of traffic light of the lane in which the vehicle is moving or waiting, vehicle position, and vehicle traveling destination	Green or red indication for the traffic lights that the vehicle agent is associated with	Average trip waiting time and average travel time
Jin and Ma (2017)	Signal group	SARSA with multiple-step backups	Vehicle arrival gap, occupancy ratio, elapsed green time, phase status, and neighboring states	Green extension between 0 and 4 seconds	Variation in travel delay

and other advantages mentioned above, there have been increasing efforts in developing decentralized systems (e.g., Cools et al., 2013). Simultaneously, emerging technologies for inexpensive computing and communication devices provide accessible opportunities to introduce decentralized schemes into the TLC system in more markets (Jin et al., 2017a).

In recent studies, a multi-agent framework is widely used in modeling TLC operations. Meanwhile, emerging methods in machine learning enable agents in such framework to build their knowledge and to operate guidelines based on feedback information concerning mobility performance and other measures like energy efficiency and environmental impacts. Approximate dynamic programming (ADP) or reinforcement learning (RL) (Barto, 1998), rooted in the perspectives of mimicking human-level intelligence, provides an insightful approach on how intelligent agents optimize their control within an application context, such as in games with high-complexity (Mnih et al., 2015).

This paper extends the RL-based adaptive intersection control approach proposed in Jin and Ma (2017) for operating a network of signalized intersections. The rest of this paper is organized as follows. Section 2 reviews several state-of-the-art TLC systems based on multi-agent modeling framework and ADP- or RL-based approaches. In Section 3, a hierarchical modeling framework is introduced to represent a decentralized TLC system. This is followed by a detailed presentation of the intersection control (in Section 4) and a description of a proposed collective learning process (in Section 5). A case study is finally carried out with experimental setup, agent design, traffic simulations, result analysis, and discussions elaborated in Section 6. The last section concludes this paper by summarizing the main findings together with an outlook on future research.

2. Relevant studies

The application of RL to TLC systems was first introduced in 1996 when Thorpe and Anderson (1996) proposed an adaptive intersection control scheme capable of modifying signal timing according to traffic conditions. The authors claimed that the proposed controller outperformed a fixed-time controller by reducing average waiting time of vehicles at an intersection. Since then, research in this area has moved towards integrating agent-based modeling technology with advanced RL or ADP methods. Table 1 summarizes the recent developments in RL-based TLC systems, and their system designs, concerning state, action, reward function and learning algorithm, are compared in details.

For intersection control, the major difference between the proposed systems lies in their agent design approach. They can be categorized by taking different entities as agents, including vehicle, intersection, and component of a signal controller. According to the design level of detail, the component of a signal controller could refer to:

- A turning movement;
- A signal group representing a group of turning movements;
- A signal phase composed of a collection of signal groups.

Among the agent designs above, several recent approaches model traffic light controller by intersection agents that determine signal

Please cite this article in press as: Jin J., Ma X., Hierarchical multi-agent control of traffic lights based on collective learning. Engineering Applications of Artificial Intelligence (2017), https://doi.org/10.1016/j.engappai.2017.10.013.