Transportation Research Part C 58 (2015) 487–501

Contents lists available at ScienceDirect

Transportation Research Part C

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

A junction-tree based learning algorithm to optimize network wide traffic control: A coordinated multi-agent framework

Feng Zhu, H.M. Abdul Aziz, Xinwu Qian, Satish V. Ukkusuri *

Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907-2051, USA

ARTICLE INFO

Article history:
Received 2 October 2013
Received in revised form 16 March 2014
Accepted 3 December 2014
Available online 31 January 2015

Keywords:
Reinforcement learning
Junction tree
Signal control coordination
VISSIM
Emissions
MOVES

ABSTRACT

This study develops a novel reinforcement learning algorithm for the challenging coordinated signal control problem. Traffic signals are modeled as intelligent agents interacting with the stochastic traffic environment. The model is built on the framework of coordinated reinforcement learning. The Junction Tree Algorithm (JTA) based reinforcement learning is proposed to obtain an exact inference of the best joint actions for all the coordinated intersections. The algorithm is implemented and tested with a network containing 18 signalized intersections in VISSIM. Results show that the JTA based algorithm outperforms independent learning (Q-learning), real-time adaptive learning, and fixed timing plans in terms of average delay, number of stops, and vehicular emissions at the network level.

1. Introduction

Vehicular traffic control on road networks is a complex decision making task in an inherently non-static environment. Heterogeneous agents (i.e., road users or vehicles, traffic controllers, pedestrians, system operators, and so on) interact with each other that shapes the dynamics of road traffic systems. Optimized traffic control systems directly contribute to travel time reduction, savings in fuel consumptions, and vehicular emissions reduction. Traffic signals are responsible for an estimated 5–10% of all traffic delays which is about 295 million vehicle-hours of delay on major roadways (FHWA, 2011) alone. The 2012 National Traffic Signal Report Card (NTOC, 2012) reports C grade for the current traffic signal operations and underscores the needs of optimizing traffic signals from system perspectives in a coordinated manner. Clearly there is a need for developing efficient algorithms for coordinating traffic signals to improve the operations of traffic systems.

Recent advances in connected vehicle (CV) environment offer useful technologies in detection and acquisition of high fidelity data that can be used for more efficient traffic control strategies. CV environment facilitates communication platform where vehicles can talk to each other (Vehicle-to-Vehicle, V2V), to the infrastructure components (Vehicle-to-Infrastructure, V2I), and also infrastructure to infrastructure communication (I2I) is possible. CV has received significant attention in Europe where it is known as Car to Car (C2C) and Car to X (C2X) technology. The intelligent transportation systems (ITS) program of the U.S. Department of Transportation (DOT) emphasizes the CV research in the ITS Strategic Plan (2010–2014). Using the accessible information from the surrounding environment to develop an efficient and robust traffic control systems is of key interest to many researchers and practitioners in the traffic engineering area.

* Corresponding author. Tel.: +1 (765) 494 2296; fax: +1 (765) 496 7996.
E-mail addresses: zhu214@purdue.edu (F. Zhu), haziz@purdue.edu (H.M.A. Aziz), qian39@purdue.edu (X. Qian), sukkusur@purdue.edu (S.V. Ukkusuri).

http://dx.doi.org/10.1016/j.trc.2014.12.009
0968-090X/

1.1. Literature review of the signal coordination problem

The signal control problem has been studies extensively in the literature. SCOOT (Hunt et al., 1982), SCATS (Lowrie, 1982), PRODYN (Farges et al., 1983), OPAC (Gartner, 1983), RHODES (Mirchandani and Head, 2001), UTOPIA (Mauro and Taranto, 1989), CRONOS (Boillot et al., 2006), and TUC (Diakaki et al., 2002) are among the first adaptive signal control systems developed by traffic engineering community. SCOOT and SCATS are centralized systems based on real time information. OPAC and RHODES use dynamic optimization to obtain the signal settings. Further, existing literature include (not limited to) rolling horizon type of control (Newell, 1998), model predictive control (Hegyi et al., 2005), store-and-forward models for traffic control (Aboudolas et al., 2009), mathematical programs with embedded traffic flow models (Lo, 2001; Lin and Wang, 2004; Beard et al., 2006; Pavlis and Recker, 2009; Aziz and Ukkusuri, 2012) and so on. Most optimization models are computationally expensive and large scale implementation is often challenging. Additionally, most control schemes do not account for dynamic feedback from the traffic environment to adjust the control scheme. In other words, control schemes do not use experience to optimize the decisions.

Identifying traffic control as a fundamental sequential decision making problem, researchers (Abdulhai et al., 2003; Bazzan, 2005; Bazzan et al., 2010; Medina and Benekohal, 2012a; El-Tantawy and Abdulhai, 2010) applied the framework of Markov Decision Processes (MDP) and deployed approximate dynamic programming (ADP) or reinforcement learning techniques to solve the problem. RL based techniques are well suited for dynamic environment like the road traffic networks. A major advantage can be gained in terms of computational complexity because no optimization is necessary in real-time. In addition, the implementation of RL-based algorithms can be paired up with connected vehicle (CV) paradigm which is expected to play a significant role in the next generation intelligent transportation systems.

The coordinated signal control problem well fits into the coordinated multi agent system framework and researchers from diverse areas have studied the potential and applicability of RL algorithms to solve the traffic control problem. Mikami and Kakazu (1994) proposed cooperative signal control scheme with a combination of evolutionary algorithm and reinforcement learning techniques. France and Ghorbani (2003) introduced a hierarchical multi agent system to design a coordinated traffic light system. Wiering et al. (2004) proposed co-learning algorithms at network level to minimize the waiting time for the vehicles. The concept of co-learning was introduced that allows both cars and traffic lights to learn from the environment. Bazzan (2005) proposed a single stage coordination game for the synchronization of traffic signals. The concepts of evolutionary game theory are applied and the analyst has to define the payoff matrices. El-Tantawy and Abdulhai (2010) recently developed a neighborhood coordinated RL based signal control that applies a joint decision framework. Other approaches include distributed constrained optimization with centralized cooperative (Oliveira et al., 2005), decentralized swarm based models (Oliveira and Bazzan, 2006), Tabu search (Hu and Chen, 2012), self organizing maps (Li et al., 2011), mixed approach of RL and supervised learning (Bazzan et al., 2010).

Application of graphical models in the area of multi-agent coordination (especially to compute the best joint actions for multi-agents) is not common. Recently, max-plus algorithm (Kok and Vlassis, 2006; Kuyer et al., 2008) has drawn the attention of a handful of researchers to solve traffic control problem. Max-plus algorithm origins from the max-product or the max-sum algorithms (Wainwright et al., 2004) which is common in graphical models. Medina and Benekohal (2012b) applied the algorithm as a coordinating strategy in the network-wide signal control problem. However, the max-plus algorithm has two key limitations.

First, it is only applicable to tree-structured networks. For general cyclic networks it cannot guarantee the convergence to an optimal solution, because the message passing in max-plus algorithm is directional. For cyclic graphs, the message passing can visit some node for multiple times. For some application it may converge, where as for others, it may not. Since cyclic structures are not uncommon in real world road networks, the quality of solution is compromised. Second, the same as the max-sum algorithm, the max-plus algorithm only provides a loopy brief propagation. Loopy brief propagation refers to the inexact messages received at a node. As there is loop in the graph, the algorithm may stop according to some criterion even if the convergence is not met. The message of a node is calculated using the most recently received incoming message. Hence the algorithm only provides an approximate inference of the exact message passing.

1.2. Motivations and contributions of the paper

A potential alternative to max-plus is the junction tree based algorithm. In this study, we extended the Junction Tree Algorithm (JTA) to obtain the best joint actions for the entire traffic network. Compared with the max-plus algorithm, JTA is an exact inference procedure capable of dealing with graphs having loops. However, JTA is originally developed to solve the general inference problem in graphical models. Accordingly, it is not readily applicable to the coordination problems in the context of traffic signal control.

To the best of our knowledge, JTA has not been applied to address the coordinated signal control problem. The advantages of proposing the JTA based RL algorithm to solve the coordinated signal control problem are as follows:

(a) it is computationally efficient,
(b) it is applicable to general cyclic or acyclic networks,
(c) it provides an exact inference of best joint selection,