Transportation Research Part B 110 (2018) 302–325

Contents lists available at ScienceDirect

Transportation Research Part B

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

Optimal traffic signal control under dynamic user equilibrium and link constraints in a general network

Hao Yu[a,b], Rui Ma[c], H. Michael Zhang[c,*]

[a] Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 210000, China
[b] Department of Civil and Environmental Engineering, University of Hawaii, Honolulu, HI 96822, United States
[c] Department of Civil and Environmental Engineering, University of California, Davis, CA 95616, United States

ARTICLE INFO

Article history:
Received 9 March 2017
Revised 13 February 2018
Accepted 13 February 2018

Keywords:
Optimal signal control
System optimal dynamic traffic assignment
Dynamic user equilibrium
Queue spillbacks
Genetic algorithm

ABSTRACT

In this study, an optimal traffic signal control framework is proposed for finding the signal control settings that minimize the total travel time in a road network with traffic lights. A novel aspect of this framework is its integration of the continuous-time double queue traffic flow model in a signal controlled traffic network to capture queue spillbacks in continuous-time. Furthermore, drivers' long term responses to changes in traffic signal control settings are captured by their route choices following Wardrop's first principle, which results in the dynamic user equilibrium state. Two signal control strategies, the fixed-timing control and the adaptive signal control, are considered. A continuous approximation method for the signal control is applied to eliminate integer variables and enhance the computational efficiency. A heuristic genetic algorithm based solution procedure is proposed to solve the proposed nonlinear programming problem with time-varying delay terms. Numerical tests are conducted in two testing networks and the results show that adaptive control with drivers taking into account signal timing on their route travel times performs best, and in some cases nearly as well as the benchmark performance derived from system optimal control without equilibrium constraints. The results also show that the advantage of adaptive over fixed-time signal control is more pronounced under UE than SO routing behavior.

1. Introduction

Traffic signals, by assigning the right-of-way temporally to separate conflicting movements at urban intersections, are essential to the safe and efficient operation of road networks. Yet, a poorly designed traffic signal timing plan would often yield the opposite results. An optimally designed traffic signal control system should consider travel demand, driver behavior, and traffic dynamics. Since traffic signal control and drivers' route choices can influence each other significantly (Ukkusuri et al., 2013; Yu et al., 2014), drivers' route choice behavior should be considered in optimal signal control.

In the past few decades, studies have been conducted in optimal signal control. A widely used signal timing model for isolated intersections was proposed in Webster (1958). Hunt et al. (1981) developed the adaptive signal control system SCOOT, which was beneficial for its ability to respond to fluctuating traffic conditions with real-time traffic. Intelligent algorithms, such as reinforcement learning, fuzzy logic, and neural networks, were also applied in this subject

* Corresponding author.
E-mail addresses: seudarwin@gmail.com (H. Yu), drma@ucdavis.edu (R. Ma), hmzhang@ucdavis.edu (H.M. Zhang).

https://doi.org/10.1016/j.trb.2018.02.009
0191-2615/

(Nakatsuji and Kaku, 1991; Trabia et al., 1999; El-Tantawy and Abdulhai, 2010). Detailed literature reviews can also be found in Papageorgiou et al. (2003) and Hamilton et al. (2013). As pointed out by Allsop (1974), drivers' route choice behavior in urban traffic networks shall be considered in signal control, and an iterative optimization assignment procedure was suggested to incorporate traffic assignment with signal control in that paper. Smith and his colleagues (Smith, 1979, 1981a; Smith and Ghali, 1990; Smith and Van Vuren, 1993) further provided a series of formulations integrating control and assignment in the static network setting. In Smith (1981b), an efficient signal control strategy was proposed, which maximized intersection capacity and ensured consistent and stable user equilibrium. Yang and Yagar (1995) formulated a bi-level model for optimal signal control with route choice in a saturated network. In their formulation, the lower level model presented the user equilibrium on a saturated network with signal settings provided by the upper level model, while the upper level model optimizes signal control with route choice provided by the lower level model. The assignment models in these formulations, however, are static in nature and do not fully capture queuing dynamics at intersections and in the networks.

Studies on dynamic optimal signal control problems started without dynamic user equilibrium constraints. Abdelfatah and Mahmassani (1998) presented a simulation-based solution algorithm for the system optimal signal control problem, where traffic dynamics are simulated in the package DYNASMART, while the flow assignments do not follow the user equilibrium. Sun et al. (2006) proposed a bi-level formulation with the cell transmission model (CTM, Daganzo (1995)). In their paper, drivers' path choices were modeled as a stochastic user equilibrium procedure, e.g. traffic flow was loaded to the network by following the logit choice model; and in the upper level model, signal lengths were optimized by a genetic algorithm. Recent advances in dynamic traffic assignment (DTA) provided an opportunity to formulate optimal traffic signal control problems that explicitly consider traffic evolution and queuing at intersections and in the network. DTA has been extensively studied in transportation research (see reviews in Peeta and Ziliaskopoulos (2001), Friesz et al. (2010) and Ban et al. (2012b)). Optimal control and variational inequalities frameworks have been extensively studied for DTA models (Friesz et al., 1993; Ran and Boyce, 1996; Xu et al., 1999; Zhu and Marcotte, 2000; Perakis and Roels, 2006). The recently developed differential variational inequality (DVI) paradigm (Pang and Stewart, 2008; Friesz, 2010) provides an alternative formulation that puts flow dynamics in the center and treat the equilibrium conditions in the form of Wardrop's first principle (Wardrop, 1952) as constraints. Ukkusuri et al. (2013) extended their previous CTM-embedded DTA studies (Ukkusuri et al., 2012; Ukkusuri and Waller, 2006) to a bi-level formulation of combined dynamic traffic assignment and traffic signal control. The lower-level is the dynamic user equilibrium (DUE) constraints as a variational inequality (VI), and the upper-level is the optimization of signal timing. The bi-level problem was then divided into these two levels, and was solved iteratively in their heuristic algorithm. More specifically, the lower level (VI for DUE) was solved by a projection algorithm proposed in Ukkusuri et al. (2012), and the upper level (binary signal settings) was solved directly by a MIP solver.

Formulating and solving the optimal signal control problem under DUE constraints is challenging for several reasons. First, the choice behaviors in DUE usually involve time-varying delays in the exact description of users' route choice under dynamic traffic modeling framework. To avoid directly dealing with these delays, approximate or simpler treatments are often used instead. Second, the queue spillback phenomenon needs to be captured properly. At signalized intersections, queues can grow from one intersection to block other intersections, which is a process known as queue spillbacks. The delay-function model (Chen and Ben-Akiva, 1998; Nie and Zhang, 2005b) and the point-queue model (Ban et al., 2012a) are relatively easy to implement, but cannot capture the spillback phenomena due to their assumption of unlimited link storage capacity. The CTM (Daganzo, 1995), which is a discrete version of the LWR kinematic wave model of Lighthill and Whitham (1955) and Richards (1956), can capture spillbacks. However, the CTM requires the discretization in both time and space. As a result, the dimension of a DTA problem with CTM is potentially much larger than that of models based on other traffic flow models (Nie and Zhang, 2005a). Similar to the CTM, the link transmission model (LTM) is derived from the kinematic wave model and can properly capture the realistic spillback phenomena (Yperman, 2007). Based on the link transmission model, a recent proposed link model named double-queue model, in which each link has a set of two queues, was developed. The concept of 'double-queue' was firstly proposed by Osorio et al. (2011), which involves an upstream and a downstream queue. The double queue (DQ) model simplifies the mathematical expression of the LTM from a partial differential equation to an ordinary differential equation (Ma et al., 2014a; Han et al., 2015), while preserves the LWR-equivalence. The DQ model uses dynamics of two dependent queues to approximate the within-link dynamics, so that the space of each link does not need to be discretized, which reduces problem size from the CTM. Another dynamic loading model named continuous-time LTM is proposed recently in Jin (2015), which has similar properties as the DQ model. Third, the DUE constraints are usually non-linear in DTA models, and few general solution methods have been developed to deal with such constraints. Moreover, the DUE states in an interrupted flow condition are hardly properly defined and solved (Han et al., 2014). In the literature, in order to formulate an optimization problem of signal control, heuristic solution methods are frequently applied but the existence of solution is not always guaranteed. Last but not the least, the integer variables involved in the formulation of signal control problem increases the computational complexity (Chen and Ben-Akiva, 1998; Lo, 2001; Sun et al., 2006; Yang and Yagar, 1995).

In this study, we propose an optimal control framework to find the optimal signal control settings that minimize the total network travel time with the consideration of spillbacks and drivers' route choice behaviors. We apply the DQ model for the dynamic loading, so that traffic dynamics and possible queue spillbacks can be captured. Drivers route choices are captured by the complementarity DUE constraints. A continuous approximation method for the signal control is applied, so that integer variables are eliminated. This is done to increase the computational efficiency while still capturing the main