Hybrid model predictive control based dynamic pricing of managed lanes with multiple accesses

Zhen Tan, H. Oliver Gao*

School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853, USA

A R T I C L E   I N F O

A B S T R A C T

We propose a hybrid model predictive control (MPC) based dynamic tolling strategy for high-occupancy toll (HOT) lanes with multiple accesses. This approach preplans and coordinates the prices for different OD pairs and enables adaptive utilization of HOT lanes by considering available demand information and boundary conditions. It also addresses such practical issues as prevention of recurrent congestion in HOT lanes, ensuring no higher toll for a closer toll exit and fairness among different OD groups at each toll entry, as well as the fact that high occupancy vehicles (HOVs) have free access to the HOT lanes. Taking the inflows at each toll entry as the control, traffic densities and vehicle queue length as observed system states, and boundary traffic as predicted exogenous input, we formulate a discrete-time piecewise affine traffic model. Optimal tolls are then derived from a one-to-one mapping based on the optimal toll entry flows. By properly formulating the constraints, we show that the MPC problem at each stage is a mixed-integer linear program and admits an explicit control law derived by multi-parametric programing techniques. A numerical experiment is presented for a representative freeway segment to validate the effectiveness of the proposed approach. The results show that our control model can react to demand and boundary condition changes by adjusting and coordinating tolls smoothly at adjacent toll entries and drive the system to a new equilibrium that minimizes the total person delay. Under the optimal prediction horizon, the on-line computational cost of the proposed control model is only about 4% and 8% of the modeling cycle of 30 s, respectively, for two typical traffic scenarios, which implies a potential of real-time implementation.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Literature review and motivation

High-occupancy toll (HOT) lanes have been recognized as one of the most applicable and cost-effective measures for reducing freeway congestion (Wang and Zhang, 2009; Liu et al., 2010). By allowing single-occupancy vehicles (SOVs) to use high-occupancy vehicle (HOV)/carpool lanes by paying a toll, excess capacity of the HOV lanes can be utilized (Wang and Zhang, 2009). HOT policy can also help alleviate traffic-related environmental problems and support sustainable urban development (Bento et al., 2014). There are three major tolling schemes used in HOT management: static pricing, time-of-day pricing, and dynamic pricing (Yang et al., 2012; Chung, 2013). Static and time-of-day tolls do not reflect or respond to real-

---

* Corresponding author.
  E-mail address: hg55@cornell.edu (H.O. Gao).

time traffic conditions, while dynamic pricing is designed to adjust toll rates according to traffic conditions (Chung, 2013). However, the performance of dynamic tolling relies on the toll-control algorithm used, which should properly account for traffic information and be implemented efficiently in an online fashion. There are various types of toll mechanisms: pass based (i.e., vehicles with a prepaid toll pass can enter the toll lane at any time), per-use based, per-mile (distance) based, zone (section) based, origin–destination (OD) based, and toll-entry based, and the combination of any of these mechanisms. The first two mechanisms are for single entry/exit managed lanes, while the rest are for systems with multiple toll entries/exits (details can be found in Yang et al., 2012 and Michalaka et al., 2013). The algorithm we propose in this study is for OD based tolling, which is an ideal toll mechanism for full utilization of the HOT lanes without creating excessive inequality across different OD pairs (Michalaka et al., 2013).

A number of HOT lane facilities have been implemented in the U.S., such as on routes I-5, I-10 W, I-15, I-95, I-394, and SR-167 (Wang and Zhang, 2009; Yang et al., 2012; Chung, 2013), and others are in progress. According to reviews (e.g., Yang et al., 2012; Chung, 2013), most of these projects use dynamic tolling, and several use distance-based rates (Chung, 2013). For example, tolls for the I-394 HOT lanes are adjusted every 3 minutes according to the detected traffic density; the toll rates, which are given in a "delta density–toll increment lookup table" (Halvorson et al., 2006; Yin and Lou, 2009; Chung, 2013) vary from $0.25 to $8.00 across different toll sections and are prescribed to maintain a free-flow traffic speed in the HOT lanes. Similarly, tolls for I-15 change every 6 minutes and vary from $0.5 to $8.00, as given in a lookup table that specifies tolls for different traffic volumes and levels of service (LOS) (Yang et al., 2012; Chung, 2013; Jou and Yeh, 2013). Further details of tolling methods can be found in studies such as Yang et al. (2012) and Chung (2013). These predefined toll adjustment rules often have many parameters to be decided (Chung, 2013). The introductory period for settling on the rules and determining the parameters can extend to years (FHWA, 2008), and the performance of the resulting schemes is still uncertain in future scenarios. Most studies on dynamic tolling have proposed similar elementary rule based models (Fu and Kulkarni, 2013). For example, Wang and Zhang (2009) proposed a feedback control dynamic tolling algorithm that uses traffic speed as the feedback variable to maintain a high LOS in the HOT lanes and maximize the total throughput; Yin and Lou (2009) also proposed a feedback approach that adjusts the toll based on the measured traffic density to maintain the desired traffic density via a linear regulator. Gardner et al. (2013) designed a more complicated approach that optimizes the current toll for full utilization of the HOT lanes based on current demand and the value of time (VOT) distribution.

In reality, speed fluctuations and recurrent congestion are common in managed-lanes, even when there is decent capacity in the system (Liu et al., 2010; Fu and Kulkarni, 2013). One important reason for this is a lack of good use of available upstream traffic demand information and measured boundary traffic conditions; toll adjustments are based only on current traffic conditions (Fu and Kulkarni, 2013), as is the case in the aforementioned studies. Yin and Lou (2009) proposed an approach that uses "self-learned" willingness to pay (WTP) parameters to derive the toll rate for the next step by solving a one-stage nonlinear optimization problem. Their model uses upstream demand information to predict the traffic in the next step. Fu and Kulkarni (2013) proposed a feedback control approach that also incorporates upstream traffic information and was shown to have a faster response to real-time traffic changes than the simple feedback approach. In both studies, traffic evolution beyond the next step was not taken into account in deriving the tolls; as a result, the performance over a longer period could be unsatisfactory. Lou et al. (2011) extended the "self-learning" approach (Yin and Lou, 2009) to a rolling-horizon setting that considers a longer period of future traffic evolution. However, it optimizes only one future toll input at each stage, which limits its pre-planning capability. In addition, the optimization model is hard to solve, although it considers only a toll range constraint. Toledo et al. (2015) proposed a rolling horizon approach that optimizes a sequence of future tolls with predicted demand from traffic simulation. It solves the nonconvex control problem by exhaustive search. Since the number of feasible tolls is exponential in the horizon length, the real-time applicability of this approach is limited.

All of the aforementioned studies focus on a single toll section. However, real-world HOT projects often have multiple toll entries and exits (Yang et al., 2012; FHWA, 2012), and new projects tend to have multiple access points with more complicated pricing schemes (FHWA, 2012). For example, there are six main toll entries and exits on the northbound I-15 in San Diego, CA (How to Use the I-15 Express Lanes, 2016), and five main toll entries and exits on the Northeast Loop 820 in Farmers Branch, TX (North Texas Express Lane Project, 2016). Research on control in the case of multi-access managed lanes is relatively limited. Fu and Kulkarni (2013) extended their single-toll-entry control method to a multi-access system but proposed only a very simple heuristic with no practical constraints. Michalaka et al. (2013) developed a general simulation model for HOT lanes with multiple toll sections that can be tailored to various toll mechanisms. Dorogush and Kurzhanskiy (2015) proposed a dynamic tolling model that first designs the flow split ratios to the two types of lanes and then sets the tolls based on VOT distributions or auctions. However, it uses no information about incoming flows. Yang et al. (2012) developed a distance-based dynamic tolling model for managed lanes with multiple entries and exits. It uses a quite realistic nonlinear continuous-time traffic-flow model but makes the control problem nonconvex and highly intractable. Zhu and Ukkusuri (2015) proposed an interesting reinforcement-learning (RL) based approach for dynamic per-mile tolling of managed lanes with flexible accesses. Real-time computation seems to be a limiting factor and the model considers only a few traffic densities and toll rates due to "curse of dimensionality".

The aforementioned issues need to be considered and addressed in the design of practical and more effective HOT-lane management strategies. In this study we formulate the managed lane system as a tractable hybrid system together with a model predictive control (MPC) based tolling strategy. Our model has two key features: 1) it can handle systems with multiple HOT-lane accesses by optimizing tolls for each toll entry–exit pair utilizing the available upstream demand information; 2) it has the flexibility to accommodate various constraints (such as free-flow on HOT lanes) while admits convenient real-time implementation. The objective of the control model is to minimize the total person delay in the system, thus improving