

A Markov Decision Process Framework to Incorporate Network-Level Data in Motion Planning for Connected and Automated Vehicles

Xiangguo Liu, Neda Masoud, Qi Zhu, Anahita Khojandi

Abstract

Autonomy and connectivity are expected to enhance safety and improve fuel efficiency in transportation systems. While connected vehicle-enabled technologies, such as coordinated cruise control, can improve vehicle motion planning by incorporating information beyond the line of sight of vehicles, their benefits are limited by the current short-sighted planning strategies that only utilize local information. In this paper, we propose a framework that devises vehicle trajectories by coupling a locally-optimal motion planner with a Markov decision process (MDP) model that can capture network-level information. Our proposed framework can guarantee safety while minimizing a trip’s generalized cost, which comprises of its fuel and time costs. To showcase the benefits of incorporating network-level data when devising vehicle trajectories, we conduct a comprehensive simulation study in three experimental settings, namely a circular track, a highway with on- and off-ramps, and a small urban network. The simulation results indicate that statistically significant efficiency can be obtained for the subject vehicle and its surrounding vehicles in different traffic states under all experimental settings. This paper serves as a poof-of-concept to showcase how connectivity and autonomy can be leveraged to incorporate network-level information into motion planning.

Keywords: Connected and Automated Vehicles ,Trajectory planning

1. Introduction

Connected vehicle (CV) technology facilitates communication among vehicles, their surrounding infrastructure, and other road users. This connectivity is enabled through Dedicated Short Range Communication (DSRC) (Kenney 2011) or cellular technologies, and paints a more comprehensive picture of the transportation network than what could be observed by each individual road user. As such, it is expected that upon deployment, the CV technology would significantly improve mobility, enhance traffic flow stability, reduce congestion, and improve fuel economy, among other benefits. The CV technology has enabled several advanced driving assistance systems (ADAS), such as Cooperative Adaptive Cruise Control (CACC) (Shladover et al. 2015; Wang, Wu, and Barth 2018; Milanés and Shladover 2014), Connected Cruise Control (CCC) (Zhang, Sun, and Orosz 2017; Orosz 2016) and Platooning (Lioris et al. 2017; Maiti, Winter, and Kulik 2017; Z. Huang et al. 2018; Bhoopalam, Agatz, and Zuidwijk 2018). Although existing CV-enabled technologies are based on local communications, the CV technology can also provide granular data at the network level by strategically positioning road side units (RSUs) to ensure connectivity throughout an entire network.

Motion planning in transportation networks has been traditionally carried out using techniques that leverage local information to make locally-optimal decisions (González et al. 2015). In particular, optimal control-based models have been widely applied to traditional transportation networks for their ability to provide short-term efficient solutions. The CV technology can help improve these locally-optimal motion planners, as it allows vehicles to see beyond line of sight. More importantly, it enables vehicles to obtain network-level information through communication with other connected vehicles and RSUs. Such connectivity can be leveraged to enhance long-term safety and efficiency of planned trajectories; however, for this potential to be realized, the network-level information should be integrated into the decision making systems. This cannot be accomplished using existing techniques, as they are not scalable to utilize granular data collected from the entire network. Hence, new methods need to be developed that can (i) leverage network-level data, and (ii) provide fast and efficient trajectories that adapt to the stochasticity of traffic networks.

This paper introduces a general framework that combines high-level network-level information with granular local information to devise network-informed cruising, routing, lane-changing, and platoon-merging decisions for a CAVs in a mixed traffic scenario, as shown in Figure 1. As demonstrated in this figure, the proposed framework combines an optimal control (OC) trajectory planning model proposed in (X. Liu et al. 2021, in press) with a Markov decision process (MDP) model developed in this paper to devise an efficient trajectory for an entire trip. The proposed MDP model can capture the progression of traffic as a stochastic process at an aggregate level, thereby complementing the optimal-control-based motion planning model through incorporating network-level information. In this context, using the proposed MDP framework allows vehicles to skip near-sighted locally-optimal trajectories (X. Liu et al. 2021, in press), and make routing, lane-changing, and platoon-merging decisions with a long-term view so as to minimize a combination of short-term and long-term costs.

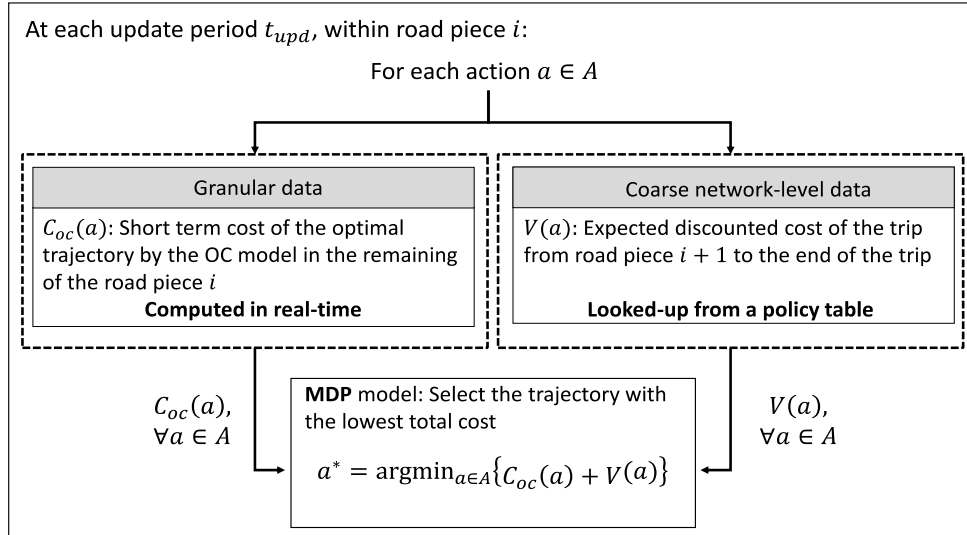


Figure 1: Structure of the proposed MDP framework. The optimal control (OC) model plans a number of trajectory to determine the short-term cost associated with every higher-layer action $a \in A$, which includes a combination of route choice, lane changing, and platoon merging. The MDP model assesses the long-term cost associated with each higher-level action $a \in A$. The MDP framework selects the action $a \in A$ that provides the minimum expected discounted cost of a trip, which is sum of the costs estimated by the OC and MDP models.

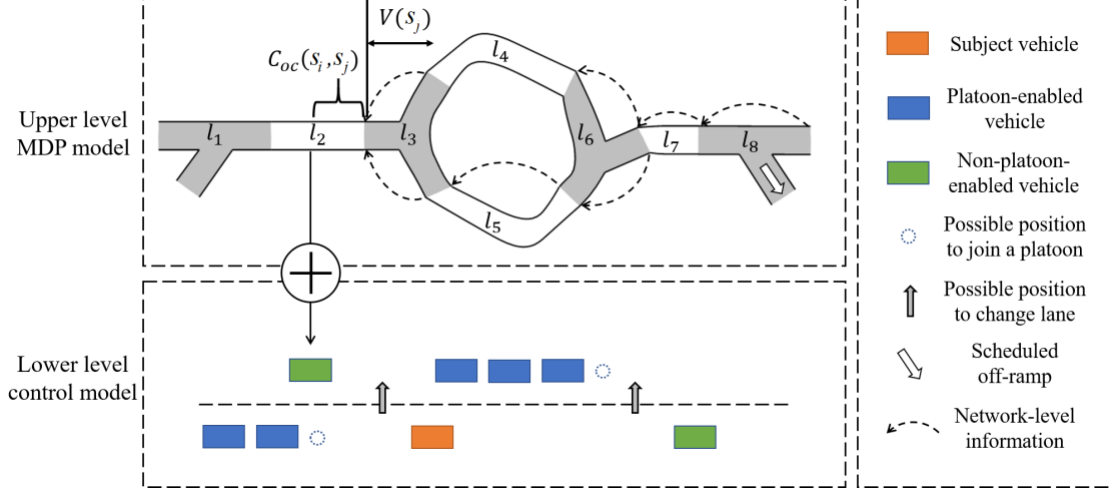


Figure 2: The upper figure displays a freeway stretch segmented into merge (on-ramp), diverge (off-ramp), and regular road pieces, where the MDP model operates. The lower figure displays a zoomed out view of a road piece, where the cost of each action (i.e., lane-changing, platoon-merging, and routing) is determined based on local information. Note that the cost of the optimal control model, C_{oc} , is computed for a given action, which is determined by the starting and ending states, s_i and s_j .

2. Related Works

2.1. Motion Planning

Motion planning for automated vehicles has been an active research topic (González et al. 2015; Zheng 2014; Larsson, Sennton, and Larson 2015; Cheng et al. 2015). With the advancement of communication, computation and sensing technologies, various planning and control techniques have been proposed, developed, and applied in complex traffic environments. Paden et al. (2016) reviewed the planning and control techniques in an urban environment, Claussmann et al. (2019) reviewed motion planning techniques for highway driving, Gritschneider, Graichen, and Dietmayer (2018; Katrakazas et al. 2015) emphasized the real time performance of planning techniques, and Zeng and Wang (2018; De Nunzio et al. 2016; H. Rakha and Kamalanathsharma 2011) focused on the efficiency of the proposed methods. B. Li et al. (2018; C. Liu, Lin, and Tomizuka 2018) attempted to balance computational performance and solution quality. Zhang, Sun, and Orosz (2017; Orosz 2016) considered communication delay and reaction time in designing motion planners, and Hardy and Campbell (2013; Galceran et al. 2015; Brechtel, Gindele, and Dillmann 2014) addressed the uncertainty in the driving behaviour of vehicles surrounding autonomous vehicles.

The motivation of this large body of research on motion planning has been to improve safety and comfort as well as reduce travel time and fuel consumption. Safety and collision avoidance have been discussed in many studies (H. Guo et al. 2018; X. Li et al. 2017; Alia et al. 2015; J. Zhou et al. 2019), some of which have considered the uncertainty of the surrounding vehicles' motion (Hardy and Campbell 2013; Bandyopadhyay et al. 2013). Besides safety guarantee, efficiency manifested in the form of reducing travel time (H. A. Rakha, Ahn, and Moran 2012; Paikari et al. 2013) and fuel consumption (H. A. Rakha, Ahn, and Moran 2012; Ahn and Rakha 2013; Boriboonsomsin et al. 2012) or increasing traffic flow (Duret, Wang, and Ladino 2019; Wei et al. 2017) has been one of the driving objectives in developing motion planners. Despite the proven short-term capability of the

proposed methods to increase efficiency, they cannot guarantee long-term efficiency due to the limited captured horizon.

Several attempts have been made in the literature to devise trajectories that account for beyond the local neighbourhood of the subject vehicle. One such approach is hierarchical design, which is sometimes referred to as the combination of trajectory planning and tracking (H. Guo et al. 2018; X. Li et al. 2017; Alia et al. 2015; Qian, De La Fortelle, and Moutarde 2016; Neunert et al. 2016; C. Huang, Naghdy, and Du 2016), and sometimes as the combination of long- and short-horizon planning. To avoid confusion, in this study, we use the term hierarchical design to denote the long- and short-horizon planning, where higher- and lower-layer decisions are made, respectively.

There have been several attempts in the literature to conduct longer-term horizon planning using hierarchical design, under specific assumptions. Zeng and Wang (2018) proposed a dynamic programming algorithm under the assumption that the speed profile of the subject vehicle's immediate leading vehicle is fixed and known. Similarly, Qian et al. (2016) assumed the surrounding vehicles' future motions to be given. Studies that assume the surrounding traffic environment to be fixed and known can compute the optimal speed profile of the subject vehicle and have the subject vehicle follow this profile (K. Huang et al. 2018; G. Guo and Wang 2018; Zeng and Wang 2018). However, due to the assumptions on the motion profiles of the surrounding vehicles, these higher-layer plans are not guaranteed to be well-executed or feasible to navigate by lower-layer planners. Because the lower-layer planners need to ensure safety and comfort and follow traffic rules, sometimes they cannot follow the suggested speed or the planned route due to not finding the opportunity to change lane, etc. On the other hand, the hierarchical layered design cannot simply be replaced with a one-time optimization problem to make both higher- and lower-layer decisions, due to its high computational complexity (Brechtel, Gindele, and Dillmann 2011). In this paper, We aim to bridge the gap between the hierarchical but non-efficient trajectory planning and the optimal but computationally-complex planning by establishing a feedback loop between higher- and lower-layer decisions in hierarchical schemes. In our proposed method, while the lower-layer planner attempts to follow the plan provided by the higher-layer planner, the higher-layer plan can also be adjusted according to the real-time execution status in the lower-layer.

In addition to the possibility that the higher-layer plan may not be executable, the plans at the lower-layer, the higher-layer, or both layers may be outdated at the time of execution in a fast-changing traffic environment. To combat outdated decisions, Paikari et al. (2013; Boriboonsomsin et al. 2012) proposed to update the higher-layer plan, while H. Guo et al. (2018; X. Li et al. 2017; Alia et al. 2015) considered updating the lower-layer plan. K. Huang et al. (2018) utilized a genetic algorithm for higher-layer planning, and a quadratic program for lower-layer adaptation, where plans on both layers are updated periodically. The two layers of decision making in our proposed hierarchical design are also closely coupled, as the lower-layer plan is devised based on higher-layer decisions, and the higher-layer plan can also be adapted based on the lower-layer execution status. Moreover, the higher-layer plan in our work is updated not only based on real-time state of the downstream traffic, but also based on the network-level evolution of traffic. Additionally, our framework is more comprehensive as it includes decision making for routing, lane-changing, platooning, and cruising.

2.2. Markov Decision Processes in Transportation

A Markov decision process (MDP) is a stochastic control process that is used extensively in many fields, including transportation, robotics and economics. MDPs can model the interaction between agents and the stochastic environment. The goal of an MDP model is to find a policy that

maximizes the total expected cumulative reward in a stochastic environment (Bellman 1957; Sutton and Barto 2018).

In the transportation field, MDPs have been utilized to plan local trajectories by modeling the uncertainty of driver behavior (Bandyopadhyay et al. 2013; Mouhagir et al. 2016). MDP and its variant, partially observable Markov decision process (POMDP), have also been applied for vehicle behavior analysis and prediction (Kamrani et al. 2020; Galceran et al. 2015; Brechtel, Gindele, and Dillmann 2014) and driving entity switching policy (Wyk, Khojandi, and Masoud 2020). Brechtel, Gindele, and Dillmann (2011) proposed an MDP-based motion planning model to devise a vehicle’s target position and velocity. The authors identified the scalability of their proposed method with respect to the number of vehicles as an open problem. To tackle the computational complexity of the problem, the authors adopted a fixed discretization of the action space to formulate the problem, which could render their methodology inefficient. You et al. (2018) designed a reward function for MDP with the objective of obtaining expert-like driving behavior. This model determines the velocity of the subject vehicle and whether the vehicle should change lanes, considering the relative position of the subject vehicle and its surrounding vehicles.

The studies above mostly employ MDPs to determine the velocity of the subject vehicle, leaving out higher-layer decisions. A recent work ((S. Zhou et al. 2017)) developed a hierarchical framework in which an MDP model was employed to make lane-changing decisions in the higher layer. They introduced three models, namely a trajectory smoother, a longitudinal controller, and a lateral controller to address the detailed execution in the lower layer. In our work, we further consider the long-term efficiency of a trajectory by extending the MDP model to a more general motion planner, which includes routing, lane-changing, and platoon-merging. In our proposed work, safety and comfort are ensured by the planner in the lower layer, while the MDP model explores the long term benefits of the planned trajectory by considering the stochastic changes in the downstream traffic environment. We use simulations to demonstrate that our proposed method results in statistically significant reductions in the long-term generalized trip cost.

2.3. *Our Contributions*

This paper introduces a framework that facilitates making trajectory planning decisions (namely, cruising, lane-changing, platoon-merging, and route choice) based on both local and network-level data. More specifically, our framework makes joint cruising, lane-changing, platoon-merging, and routing decisions to minimize the total expected discounted cost of a (leg of a) trip in a dynamic environment. This is accomplished through two main modules within an MDP framework: (1) an optimal-control-based trajectory planning model that provides the vehicle’s acceleration profile with the goal of maximizing safety and comfort locally (X. Liu et al. 2021, in press); and (2) an MDP model that enables incorporating network-level information into the decision making process.

The contributions of this paper are as follows. This work is the first to advance the traditional local motion planning models by incorporating strategically-condensed high volume of network-level data using a Markov Decision Process (MDP) modeling framework, hence devising entire efficient trajectories in dynamic traffic streams. In this general framework, cruising, routing, lane-changing, and platoon-merging decisions are made concurrently. We conduct comprehensive simulation experiments to demonstrate the benefits of augmenting traditional trajectory planning models with an MDP model for both the subject vehicle and its surrounding vehicles. We demonstrate that not only does a CV benefit from utilizing network-level information in devising its own trajectory, but its surrounding vehicles, which may be CAVs or legacy vehicles, also experience

second-hand cost-reduction benefits. These results could have great policy implications, as they demonstrate that only a handful of CAVs in a traffic stream could serve as traffic regulators.

3. Problem Statement

Consider a CAV, to which we refer as the *subject vehicle*, who is making a trip from a known origin to a known destination. The subject vehicle is able to directly *observe* its surroundings using its onboard sensor systems as well as basic safety messages (BSMs) obtained from other vehicles or RSUs within its communication range. Owing to its connectivity, the subject vehicle can also obtain network-level information about the state of traffic. The objective of the subject vehicle is to navigate the network safely and comfortably, while at the same time minimizing its travel cost, which is composed of time cost and energy cost, by utilizing both granular local data and coarse network-level information.

4. Methodology

4.1. The MDP Framework

The proposed framework determines the trajectory of a subject vehicle, including *fine-grained* decisions (i.e., the acceleration profile) and *coarse* decisions (i.e., routing, lane changing, and platoon merging). In this framework, fine-grained decisions are made by a local optimal control trajectory planning model using only local information, and coarse decisions are made by an MDP model using network-level information. The MDP framework combines the two models to make a final decision about the trajectory of the subject vehicle: For each coarse action (where a coarse action is a unique vector of route choice, platoon merging, and lane changing), the MDP framework uses the optimal control model to obtain the lowest short-term cost of completing the action, and the MDP model to obtain the long-term expected discounted cost of completing the same action. Finally, the action that provides the the lowest total cost will be selected and pursued by the vehicle. This framework is demonstrated in Figure 1.

An example network is displayed in Figure 1, where the subject vehicle is located on the right lane, planning to take the off-ramp marked by an arrow. The general travel cost incurred by the vehicle is a linear combination of the route travel time and fuel cost. To optimize its trajectory, in addition to determining the exact position, speed, and acceleration of the subject vehicle at each point in time, we need to make three sets of higher-level decisions with long-term implications: whether (and where) to change lanes, whether to join (or split from) a platoon, and which route to take.

Each action can have conflicting implications in terms of energy efficiency, travel time, and passenger safety. For example, the vehicle would be able to travel at a higher speed on the left lane, but may have more opportunities to join a platoon and increase its fuel economy on the right lane. The trade-offs between these actions can be captured by an optimal-control-based trajectory planning model that uses local information (i.e., the speed and availability of platoons at both lanes). As another example, while joining a platoon would provide fuel efficiency, changing platoon membership frequently could pose safety risks on the vehicle occupants and create instability in the traffic stream. This example highlights the importance of not making decisions based solely on minimizing the short-term vehicle-specific costs, and taking a longer-term, futuristic view of the cost that requires incorporating network-level information into the decision making process. As such, the

proposed MDP framework is designed to capture the expected long-term cost of each action, allowing the vehicle to make informed decisions based on both local and network-level information.

In order to model the system with a view on facilitating the incorporation of both granular and network-level information, we make a number of assumptions. First, we divide the network into a number of relatively large cells, to which we refer as *road pieces*. Road pieces are constructed such that (i) the macroscopic-level traffic dynamics are homogeneous within each piece at each point in time; and (ii) all vehicles within a road piece are within a reliable communication range of one-another. As such, we introduce three types of road pieces, namely, merge (which includes a single on-ramp/road), diverge (which includes a single off-ramp/road), and regular (which does not include any on- or off-ramps). In Figure 1, for example, l_1 is an on-ramp or merge piece, while l_4 and l_5 are regular pieces.

The trajectory planning model is re-optimized dynamically as the immediate neighborhood of the subject vehicle evolves. This re-optimization occurs after a time period t_{upa} has lapsed, which is set to 0.4 sec following (X. Liu et al. 2021, in press). The MDP model is solved off-line, and its resulting optimal policies are stored in a look-up policy table that can be accessed at any time. In the rest of this section, we elaborate on the MDP model in subsection 4.2, and provide a brief overview of the optimal control model in subsection 4.3.

4.2. The MDP Model

The MDP framework considers three traffic states, namely, free-flow, onset-of-congestion, and congested traffic. The traveling speed of the subject vehicle is determined based on the traffic state of the road piece the vehicle is traversing. When the subject vehicle enters a new road piece l_i , a decision is made as to whether the vehicle should change lanes and whether to join a platoon. It is assumed that the vehicle can finish the lane changing and platoon merging processes within the same road piece l_i . If there are more than one road pieces following l_i , the subject vehicle also has to make a route choice decision by selecting one of the candidate road pieces, $l'_i \in S_l(l_i)$, where $S_l(l_i) = \{l'_{i1}, l'_{i2}, \dots\}$ is the set of road pieces connected to l_i , and therefore depends on the network structure.

Table 1: Table of Notation

Notation	Definition
Le	Left lane
Ri	Right lane
l_i	Road piece i
l	A generic road piece
l'_{ij}	The j th road piece directly connected to road piece l_i
$S_l(l_i) = \{l'_{i1}, l'_{i2}, \dots\}$	Set of road pieces directly connected to road piece l_i
$l'_i \in S_l(l_i)$	The selected road piece among the set of road pieces connected to l_i
$L = \{l_i\}$	Set of road pieces
l_o	The road piece at the origin of the trip
l_d	The road piece at the destination of the trip
ξ_{tr}^{Le}	Macroscopic state of traffic in the left lane
ξ_{tr}^{Ri}	Macroscopic state of traffic in the right lane
$\xi_{tr} = [\xi_{tr}^{Le}, \xi_{tr}^{Ri}]$	Vector specifying the macroscopic state of traffic

Notation	Definition
ξ_p^{Le}	Percentage of platoon-enabled vehicles in the left lane
ξ_p^{Ri}	Percentage of platoon-enabled vehicles in the right lane
$\xi_p = [\xi_p^{Le}, \xi_p^{Ri}]$	Vector specifying the percentage of platoon-enabled vehicles
$\mu = [l_i, \xi_{tr}, \xi_p]$	The environment state vector
$\phi_l \in \{Le, Ri\}$	The lateral position of the subject vehicle
$\phi_p \in \{0,1\}$	Platoon membership status of the vehicle
$\phi = [\phi_l, \phi_p, d]$	The vehicle state vector
d	The number of road pieces to the scheduled splitting of the platoon, where $d = -1$ if the subject vehicle is not in a platoon
$s = (\mu, \phi) \in S$	State of the traffic dynamics process
$S = \{s\}$	Set of all possible states of the traffic dynamics process
$c^f(s)$	The fuel cost of the subject vehicle at state s
$c^t(s)$	The time cost of the subject vehicle at state s
$c^{di}(s_1, s_2)$	Passenger discomfort/safety risk cost for a vehicle transitioning from state s_1 to s_2
N_{lc}	Number of lane changes
$\Lambda = [\lambda_f, \lambda_t, \lambda_{di}]$	Vector of cost component coefficients, containing elements for fuel, time, and discomfort/safety
$C(s_1, s_2)$	Cost vector for a vehicle transitioning from state s_1 to s_2
$= \left[\left(c^f(s_1) + c^f(s_2) \right) / 2, \left(c^t(s_1) + c^t(s_2) \right) / 2, c^{di}(s_1, s_2) \right]^T$	
$C_{s_1}^{s_2} = \Lambda C(s_1, s_2)$	Sum of fuel, time, and comfort/safety costs
$V([l, \xi_{tr}, \xi_p], [\phi_l, \phi_p, d])$	The minimum total expected discounted cost-to-go starting from state $s = (\mu, \phi)$
c_{fl}	Cost of missing the trip destination
Probability distributions	
$q_\phi^f(\mu)$	Probability that the subject vehicle fails to change lanes if such a decision has been made
$g_l^1(\xi_p)$	Probability of successful platoon merging with lane changing
$g_l^0(\xi_p)$	Probability of successful platoon merging without lane changing
$w(k)$	Probability distribution for the number of road pieces, k , for which the subject vehicle can stay with a platoon it has met
Transition matrices	
$p_l^{Le}((\xi_{tr}^{Le})' \xi_{tr}, \xi_p)$	Probability that the traffic state transitions to $(\xi_{tr}^{Le})'$ in the left lane, given ξ_{tr} and ξ_p
$p_l^{Ri}((\xi_{tr}^{Ri})' \xi_{tr}, \xi_p)$	Probability that the traffic state transitions to $(\xi_{tr}^{Ri})'$ in the right lane, given ξ_{tr} and ξ_p
$h_l^{Le}((\xi_p^{Le})' \xi_{tr}, \xi_p)$	Probability that the platoon intensity transitions to $(\xi_p^{Le})'$ in the left lane, given ξ_{tr} and ξ_p

Notation	Definition
$h_l^{Ri} \left((\xi_p^{Ri})' \xi_{tr}, \xi_p \right)$	Probability that the platoon intensity transitions to $(\xi_p^{Ri})'$ in the right lane, given ξ_{tr} and ξ_p
Actions	
$a_l \in \{Le, Ri\}$	Target lane
$a_p \in \{0,1\}$	Target platoon membership
$a_r \in S_l(l_i)$	Target route
$a = [a_l, a_p, a_r]$	The action taken by the subject vehicle
$A = \{a\}$	Action set

Let $s = (\mu, \phi) \in \mathcal{S}$ denote the state of the traffic dynamics process and $\mathcal{S}$ is the set of all possible states s . Vector $\mu = [l_i, \xi_{tr}, \xi_p]$ in this process denotes the location-dependent environment state, where $l_i \in L$ and L includes the location of the origin and destination of the trip (leg), denoted by l_o and l_d , respectively, and all other road pieces on all possible paths that connect the origin to the destination. The vector $\xi_{tr} = [\xi_{tr}^{Le}, \xi_{tr}^{Ri}]$ denotes the macroscopic state of traffic on the left and right lanes, respectively. More specifically, we consider three macroscopic traffic states of free-flow, onset-of-congestion, and congested. Vector $\xi_p = [\xi_p^{Le}, \xi_p^{Ri}]$ denotes the percentage of platoon-enabled vehicles on the left and right lanes, respectively.

Let $\phi = [\phi_l, \phi_p, d]$ denote the state of the subject vehicle. Here, $\phi_l \in \{Le, Ri\}$ denotes the lateral position of the subject vehicle, where ‘*Le*’ and ‘*Ri*’ refer to the left and right lanes, respectively. Furthermore, $\phi_p \in \{0,1\}$ is a binary indicator denoting the platoon membership status of the vehicle, where $\phi_p = 0$ indicates that the subject vehicle is not a platoon member and $\phi_p = 1$ indicates otherwise. Let d denote the number of road pieces to the scheduled splitting of the platoon the subject vehicle is a member of. We set $d = -1$ if the subject vehicle is not in a platoon. We assume that before merging, vehicles that will stay in a same platoon will negotiate and reach consensus on the scheduled splitting position $d = k$. Vehicles in a platoon moving to the next road piece will have their d decreased by 1. The platoon has to split/dissolve when $d = 0$. (We assume the subject vehicle can optimize its action periodically and thus actively split before the scheduled splitting position in our current model, but this can be easily modified by disabling the first, third and fourth expression in Equation (8).

Let $a = [a_l, a_p, a_r]$ denote the action taken by the subject vehicle in the beginning of each road piece, where $a_l \in \{Le, Ri\}$ denotes the target lane for the subject vehicle, $a_p \in \{0,1\}$ denotes the target platoon membership, where $a_p = 0$ indicates that the vehicle stays as a free agent and $a_p = 1$ indicates that the vehicle merges into a platoon, and $a_r \in S_l(l_i)$ denotes the path selected by the vehicle.

Let $c^f(s)$ and $c^t(s)$ denote the fuel cost and time cost of the subject vehicle at the state s , respectively. See (X. Liu et al. 2021, in press) for the computation of fuel cost, $c^f(s)$. The time cost of a trip (leg) can be computed as the length of the road piece len_i over the velocity in lane ϕ_l under traffic condition $\xi_{tr}, v(\xi_{tr}, \phi_l)$, i.e.,

$$c^t(s) = len_i / v(\xi_{tr}, \phi_l) \quad (1)$$

Let $c^{di}(s_1, s_2)$ denote the cost associated with passenger discomfort/safety risk for transitioning from state s_1 to s_2 . The passenger discomfort/safety cost is assumed to be realized when the vehicle is changing lanes, and increase linearly with the number of lane changes. Therefore,

$$c^{di}(s_1, s_2) = g(N_{lc}) \quad (2)$$

where $g(\cdot)$ is a linear function and N_{lc} is the number of lane changes in the current road piece. In one road piece, the subject vehicle is not expected to change lanes more than once, i.e., $N_{lc} \in \{0, 1\}$.

Let $C_{s_1}^{s_2}$ denote the sum of all three costs discussed above for a vehicle that starts a road piece in state s_1 and ends it in state s_2 . The exact transition position depends on the real-time traffic environment. For simplification, we assume the transition takes place in the middle of a road piece, and therefore $C_{s_1}^{s_2}$ can be formulated as:

$$C_{s_1}^{s_2} = \Lambda C(s_1, s_2) = \lambda_f (c^f(s_1) + c^f(s_2))/2 + \lambda_t (c^t(s_1) + c^t(s_2))/2 + \lambda_{di} c^{di}(s_1, s_2) \quad (3)$$

where the vector $\Lambda = [\lambda_f, \lambda_t, \lambda_{di}]$ contains the corresponding coefficients for each cost component, and $C(s_1, s_2) = \left[(c^f(s_1) + c^f(s_2))/2, (c^t(s_1) + c^t(s_2))/2, c^{di}(s_1, s_2) \right]^T$ is the cost vector for a vehicle transitioning from state s_1 to s_2 . Note that all costs are functions of our action, where the action is implied from the transition of the state from s_1 to s_2 . We assume that Λ can be different for each driver, since different cost terms are of different importance for each driver. The total travel cost $C_{s_1}^{s_2}$ describes the generalized cost of travel in a road piece. For example, the MDP cost for a vehicle that starts a road piece on the left lane as a free agent and ends the road piece on the right lane as a free agent can be denoted by $C_{Le,0,-1}^{Ri,0,-1}$.

An important part of the MDP model is the transition probability matrices that allow us to model the dynamics of the system. Let $p_l^{Le}((\xi_{tr}^{Le})' | \xi_{tr}, \xi_p)$ and $p_l^{Ri}((\xi_{tr}^{Ri})' | \xi_{tr}, \xi_p)$ denote the probability that given ξ_{tr} and ξ_p , the traffic state transitions to $(\xi_{tr}^{Le})'$ in the left lane and to $(\xi_{tr}^{Ri})'$ in the right lane in road piece l_i , respectively. Let $h_l^{Le}((\xi_p^{Le})' | \xi_{tr}, \xi_p)$ and $h_l^{Ri}((\xi_p^{Ri})' | \xi_{tr}, \xi_p)$ denote the probability that given ξ_{tr} and ξ_p , the platoon intensity transitions to $(\xi_p^{Le})'$ in the left lane and to $(\xi_p^{Ri})'$ in right lane, respectively. These transition probability matrices can be learnt from historical data.

Let $q_\phi^f(\mu)$ denote the probability that the subject vehicle fails to change lanes if such a decision has been made. Note that q_ϕ^f is a function of the traffic state in target lane. Let $g_l^1(\xi_p)$ and $g_l^0(\xi_p)$ denote the probability of successful platoon merging with and without lane changing, respectively. Note that g_l^1 is a function of the density of platoon-enabled vehicles in the target lane, and g_l^0 is a function of the availability of platoon-enabled vehicles in the immediate downstream of the subject vehicle in the original lane. Let l_i' and $\mu' = [l_i', \xi_{tr}', \xi_p']$ denote a candidate road piece directly connected to l_i and its corresponding environment state vector, respectively. Hence, the problem terminates when the vehicle reaches its destination, i.e., $l_i = l_d$. Finally, let $V([l_i, \xi_{tr}, \xi_p], [\phi_l, \phi_p, d])$ denote the minimum total expected discounted cost starting with the

vehicle state $[\phi_l, \phi_p, d]$ and the environment state $[l_i, \xi_{tr}, \xi_p]$. Hence, for $l_i = l_d$, the minimum total expected discounted cost is given by

$$V([l_d, \xi_{tr}, \xi_p], [\phi_l, \phi_p, d]) = \begin{cases} 0 & \text{if the vehicle is at the correct destination} \\ c_{fl} & \text{otherwise} \end{cases} \quad (4)$$

where c_{fl} is a cost incurred should the subject vehicle fail to reach its destination (e.g., the vehicle should be a single vehicle in the right lane at the target off-ramp piece).

For $l_i \neq l_d$, when $\phi_l = Le, \phi_p = 0, d = -1$, the minimum expected discounted cost is given by

$$V(\mu, Le, 0, -1) = \min_{a \in A} \begin{cases} \Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu', Le, 0, -1) & a_l = Le, a_p = 0 \\ g_l^0(\xi_p) \{ \Lambda C_{Le,0,-1}^{Le,1,k-1} + W(\mu', Le, 1, k-1) \} + \\ (1 - g_l^0(\xi_p)) \{ \Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu', Le, 0, -1) \} & a_l = Le, a_p = 1 \\ q_\phi^f(\mu) \{ \Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu', Le, 0, -1) \} + \\ (1 - q_\phi^f(\mu)) \{ \Lambda C_{Le,0,-1}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} & a_l = Ri, a_p = 0 \\ g_l^1(\xi_p) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Le,0,-1}^{Ri,1,k-1} + W(\mu', Ri, 1, k-1) \} + \\ \left((1 - g_l^1(\xi_p)) (1 - q_\phi^f(\mu)) \right) \{ \Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu', Le, 0, -1) \} & a_l = Ri, a_p = 1 \end{cases} \quad (5)$$

where U and W are described in Equations (6) and (7) as the minimum expected discounted cost of the remainder of the trip starting from the next road piece l'_i for a vehicle that intends to maintain its state and join a platoon, respectively.

$$U(\mu', \phi_l, \phi_p, d) = \alpha \sum_{\xi_{tr}' \xi_p'} p_l^{Le} ((\xi_{tr}^{Le})' | \xi_{tr}, \xi_p) p_l^{Ri} ((\xi_{tr}^{Ri})' | \xi_{tr}, \xi_p) \quad (6)$$

$$h_l^{Le} ((\xi_p^{Le})' | \xi_p, \xi_p) h_l^{Ri} ((\xi_p^{Ri})' | \xi_p, \xi_p) V(\mu', \phi_l, \phi_p, d)$$

$$W(\mu', \phi_l, 1, k-1) = \alpha \sum_{\xi_{tr}' \xi_p'} \sum_k w(k) p_l^{Le} ((\xi_{tr}^{Le})' | \xi_{tr}, \xi_p) p_l^{Ri} ((\xi_{tr}^{Ri})' | \xi_{tr}, \xi_p) \quad (7)$$

$$h_l^{Le} ((\xi_p^{Le})' | \xi_p, \xi_p) h_l^{Ri} ((\xi_p^{Ri})' | \xi_p, \xi_p) V(\mu', \phi_l, 1, k-1)$$

The four arguments of the min function in Equation (5) correspond to the costs of the lane changing and platoon merging actions. The expected discounted cost (with the initial values as specified) is then the minimum cost over the entire action set, which consists of lane changing, platoon merging, and route choice.

The first expression in Equation (5) corresponds to the action that results in no change in the state of the vehicle; that is, the subject vehicle stays on the left lane as a single agent. The cost of this action is equal to the cost of continuing with the initial state $(Le, 0, -1)$ on the current road piece, plus the min expected discounted cost of starting the next road piece under the same initial state.

The second expression in Equation (5) corresponds to the action of staying on the left lane, but joining a platoon. The first term here corresponds to the expected cost of the scenario where the vehicle successfully joins a platoon. Under this scenario, the vehicle incurs both the cost of this new trajectory on the current road piece and the expected discounted cost of the rest of the trip starting from its new state as a platoon member. In case the execution of this action fails (i.e., the vehicle cannot join a platoon), the vehicle will continue under the previous state on the current road piece, and incurs an expected discounted cost for the rest of the trip starting from the left lane as a single agent. This cost is demonstrated in the second term.

The third expression in Equation (5) corresponds to the action of changing to the right lane and remaining a free agent. Similar to the previous case, the first term captures the expected cost if the action can be completed, and the second term corresponds to the cost of the trajectory if the vehicle fails to complete the action.

Finally, the last expression in Equation (5) corresponds to the action of changing lanes and joining a platoon. In this case, the expected discounted cost is the summation of two terms, the first term corresponding to the entire action being completed, and the second term corresponding to the action failing.

For the case where the subject vehicle is a platoon member and the platoon splitting time has not been reached (i.e., $l_i \neq l_d$, when $\phi_l = Le$, $\phi_p = 1$, $d > 0$), the minimum expected discounted cost is given by

$$V(\mu, Le, 1, d) = \min_{a \in A} \begin{cases} \Lambda C_{Le,1,d}^{Le,0,-1} + U(\mu', Le, 0, -1) & a_l = Le, a_p = 0 \\ \Lambda C_{Le,1,d}^{Le,1,d-1} + U(\mu', Le, 1, d - 1) & a_l = Le, a_p = 1 \\ q_\phi^f(\mu) \{ \Lambda C_{Le,1,d}^{Le,1,d-1} + U(\mu', Le, 1, d - 1) \} + \\ (1 - q_\phi^f(\mu)) \{ \Lambda C_{Le,1,d}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} & a_l = Ri, a_p = 0 \\ (1 - g_l^1(\xi_p) (1 - q_\phi^f(\mu))) \{ \Lambda C_{Le,1,d}^{Le,1,d-1} + U(\mu', Le, 1, d - 1) \} \\ + g_l^1(\xi_p) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Le,1,d}^{Ri,1,k-1} + W(\mu', Ri, 1, k - 1) \} & a_l = Ri, a_p = 1 \end{cases} \quad (8)$$

The first expression in the min function in Equation (8) refers to the case that the subject vehicle splits from its platoon without changing lanes. Since this can always be achieved, the expected discounted cost of this action is the cost of the subject vehicle traveling on its current road piece as a free agent, plus its expected discounted cost of continuing to travel as a free agent starting from the next road piece.

The second expression in Equation (8) describes the scenario where the subject vehicle maintains its current state. Under this scenario, the subject vehicle traverses its current road piece while maintaining its state, and continues the rest of its trip with the platoon splitting time reduced by one unit.

The third expression in Equation (8) has the subject vehicle splitting from the platoon and changing lanes. When the subject vehicle decides to change lanes while in a platoon, it has to split from its platoon first. The first term here captures the scenario where the subject vehicle is not able to change lanes, under which case it will continue in its current platoon. Note that the OC model will inform the subject vehicle whether it can successfully change lanes. As such, if OC determines that changing lanes cannot take place safely, the subject vehicle will not split from its platoon. If the subject vehicle can change lanes, it will split from its platoon and continue the rest of the trip on the right lane as a free agent.

Finally, the fourth expression in Equation (8) has the subject vehicle changing lanes and traveling on the right lane in a platoon. For this action to take place, the subject vehicle should split from its current platoon, change lanes, and join a platoon on the right lane. Since we are assuming that the subject vehicle is always able to split from its current platoon, the probability of completing this action is the probability of successfully changing lanes and joining a platoon in the new lane. The first term here captures the cost of this action failing, in which case the subject vehicle would continue on the left lane in its current platoon. The second term captures the cost of the action being completed successfully.

For the case where the vehicle is a platoon member on the left lane and the platoon splitting time has arrived (i.e., $l_i \neq l_d$, when $\phi_l = Le$, $\phi_p = 1$, $d = 0$), the minimum expected discounted cost is given by

$$V(\mu, Le, 1, 0) = \min_{a \in A} \begin{cases} \Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1) & a_l = Le, a_p = 0 \\ g_l^0(\xi_p) \{ \Lambda C_{Le,1,0}^{Le,1,k-1} + W(\mu', Le, 1, k-1) \} + \\ (1 - g_l^0(\xi_p)) \{ \Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1) \} & a_l = Le, a_p = 1 \\ q_\phi^f(\mu) \{ \Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1) \} + \\ (1 - q_\phi^f(\mu)) \{ \Lambda C_{Le,1,0}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} & a_l = Ri, a_p = 0 \\ g_l^1(\xi_p) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Le,1,0}^{Ri,1,k-1} + W(\mu', Ri, 1, k-1) \} \\ + (1 - g_l^1(\xi_p)) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1) \} & a_l = Ri, a_p = 1 \end{cases} \quad (9)$$

In Equation (9), $d = 0$ indicates that the platoon is dissolving and the subject vehicle has to split from it in the current road piece. The first expression in Equation (9) captures the scenario where the subject vehicle continues to travel on the left lane as a free agent after splitting from its current platoon.

The second expression in Equation (9) captures the case where the subject vehicle decides to join another platoon in the left lane, which may fail due to the absence of platoon-enabled vehicles in the left lane (second term).

The third expression in Equation (9) indicates that the subject vehicle plans to change lanes and continue to travel as a free agent. This action may fail if the subject vehicle cannot change lanes (first term), in which case the subject vehicle continues to travel on the left lane as a free agent. Otherwise, the subject vehicle travels on the right lane as a free agent.

The fourth expression in Equation (9) captures the scenario where the subject vehicle switches to the right lane and joins a platoon. The first term is the cost of the case where this action can be completed successfully, and the second term captures the case where this action fails.

For other cases that the vehicle is on the right lane (i.e., $\phi_l = Ri$), the minimum expected discounted cost has similar formulas as above. Refer to 7 for details.

4.3. The Optimal Control (OC) Model

The MDP model creates a policy that advises the set of coarse actions the vehicle needs to take in order to complete its trip in the most cost-effective way. However, the MDP model cannot provide a full, implementable trajectory for the subject vehicle that includes its target acceleration profile. As such, the MDP framework utilizes an optimal control (OC) model to bridge this gap. The role of the OC model is two-fold: First, it devises an acceleration profile for the subject vehicle to complete coarse actions (or determines the infeasibility of completing the coarse actions) following a quintic trajectory function and subject to collision avoidance and bounds on the vehicle's speed, acceleration, and jerk (X. Liu et al. 2021, in press). The quintic trajectory function is selected due to its ability to provide a smooth trajectory. This function is demonstrated in Equation (10). In this equation, $x(t)$ and $y(t)$ indicate the longitudinal and lateral positions of the vehicle at time t , respectively, Coefficients a_0^i through a_5^i and b_0^i through b_5^i are decision variables that determine the optimal solution.

$$\begin{cases} x(t) &= a_5^i t^5 + a_4^i t^4 + a_3^i t^3 + a_2^i t^2 + a_1^i t + a_0^i \\ y(t) &= b_5^i t^5 + b_4^i t^4 + b_3^i t^3 + b_2^i t^2 + b_1^i t + b_0^i \end{cases} \quad (10)$$

Additionally, the OC model quantifies the short-term cost of completing the coarse actions based on the acceleration profile of the vehicle (X. Liu et al. 2021, in press). More specifically, given the action $\mathbf{a} = \{a_l, a_p, a_r\}$, the OC model plans a trajectory that minimizes a convex combination of fuel and time costs, subject to safety and comfort guarantees. The details on the OC model can be found in (X. Liu et al. 2021, in press).

For each action $\mathbf{a} \in A$, this short term cost C_{oc} is then combined with the expected long-term cost $V(\mu, \phi)$ in the MDP framework. The MDP framework enumerates all coarse actions $\mathbf{a} \in A$, and selects the action that minimizes the total cost by the OC and MDP models.

5. Experiments and Analysis

In this section, we will conduct simulations in three experimental settings, namely a circular track, a straight highway, and a small network with route choice. We compare the performance of the local

OC model and the MDP framework, in which the OC and MDP models are combined, under different traffic states in all three experimental settings. Our simulations are based on a previously built simulation platform in (X. Liu et al. 2021, in press), in which surrounding vehicles follow the Intelligent Driver Model (Jin and Orosz 2016). We consider aerodynamic, rolling, grade, and inertial resistance forces for fuel cost computation (Gillespie 1992), and set the value of time (VoT) to 10 dollars per hour.

5.1. Model Calibration

In a future connected and automated vehicle system, parameters of the MDP framework can be calibrated using historical data. Note that even when abundant CAV data becomes available, it could still be a difficult task to fully and precisely represent every single driving scenario due to the complexity of human behavior, non-linearity of interactions between vehicles, and the dynamic nature of the transportation network. Therefore, a more practical approach would be to use historical data to partition ξ_{tr}^{Le} , ξ_{tr}^{Ri} , ξ_p^{Le} and ξ_p^{Ri} into different clusters, representing different traffic states and platoon intensities in the left and right lanes, respectively. The transition probabilities can then be estimated using the maximum likelihood principle, based on the occurrence percentages of the corresponding state transitions in historical records. Furthermore, once data is available, we can use it in a maximum likelihood estimation framework to calibrate functions $q_\phi^f(\mu)$, $g_l^1(\xi_p)$ and $g_l^0(\xi_p)$, and $w(k)$.

For the current study, since historical data does not exist, we use simulations to create CAV driving scenarios, and treat observations within simulations as historical data. We conduct simulations using the OC model proposed in (X. Liu et al. 2021, in press), in which a mixed traffic of CAVs and legacy vehicle can travel in a traffic stream. The parameter values used in these simulations are specified in 8. After a warm-up period of about 20 minutes, we estimate the required parameters for this study using the maximum likelihood principle.

In this work, we assume that only the subject vehicle is adopting the MDP framework, and thus the actions taken by a single vehicle do not change the macroscopic traffic state of the system. If the penetration rate of vehicles that adopt the MDP framework is high, actions taken by these vehicles could change the state. In this case, model parameters and the optimal MDP policy can be updated periodically to capture such changes.

5.2. A Circular Track Scenario

Circular track is a great experimental setting as it can demonstrate the impact of the proposed methodology not only on the generalized cost of a trip, but also on the properties of traffic wave propagation (Sugiyama et al. 2008; Tadaki et al. 2013). Stern et al. (2018) demonstrate that a low penetration of autonomous vehicles can effectively dampen the stop-and-go wave using a circular track. Here, we conduct our simulations in a circular track, where the surrounding vehicles can merge into platoons, but cannot change lanes, enter through on-ramps, or exit from off-ramps. In these simulations, the subject vehicle will have a trip of 10.8 kilometers in length, and different traffic states (e.g., free-flow, onset-of-congestion and congested) are generated similar to (X. Liu et al. 2021, in press), by utilizing a fundamental diagram of traffic flow.

In figures presented in this paper, *OC* and *MDP* refer to the local optimal controller and the MDP framework (also referred to as the MDP controller), respectively. The suffix $_xK$ indicates that the circle length is x kilometers. The suffix $_low$, $_medium$ and $_high$ represent the penetration of

platoon-enabled vehicles. Specifically, *_low* indicates that all surrounding vehicles are non-platoon-enabled, *_medium* indicates a subset (about 30%) of surrounding vehicles are platoon-enabled, and *_high* indicates that all surrounding vehicles are platoon-enabled.

Figure (3) shows the generalized cost incurred by the subject vehicle under the OC and MDP controllers when the circular track is 2, 5 and 10 km in perimeter, respectively. This figure indicates that the circle perimeter does not significantly affect the subject vehicle’s generalized cost. In the free-flow and onset-of-congestion states, the MDP controller provides statistically significant (at the 5% significance level) lower costs. In the congested traffic state, no statistically significant difference in cost is observed between the MDP and OC controllers, although the variance of cost is lower under the MDP controller.

Figure (4) shows the generalized cost for the subject vehicle under different controllers and a track perimeter of 5km, as the penetration rate of platoon-enabled vehicles in the surrounding traffic changes. In the free-flow state, it is only under a high penetration rate that the MDP controller results in a significantly smaller cost compared with the OC controller, and there is no significant difference when penetration rate is low or medium. In onset-of-congestion traffic state, the MDP controller has significantly smaller costs than the OC controller at all penetration rates. In congested traffic, the MDP and OC controllers are not different in a statistically significant manner, although the generalized cost is much lower under a high penetration rate of platoon-enabled vehicles. Generally, higher intensity of platoon-enabled vehicles gives rise to more opportunities for the subject vehicle to join a platoon, thereby resulting in less cost.

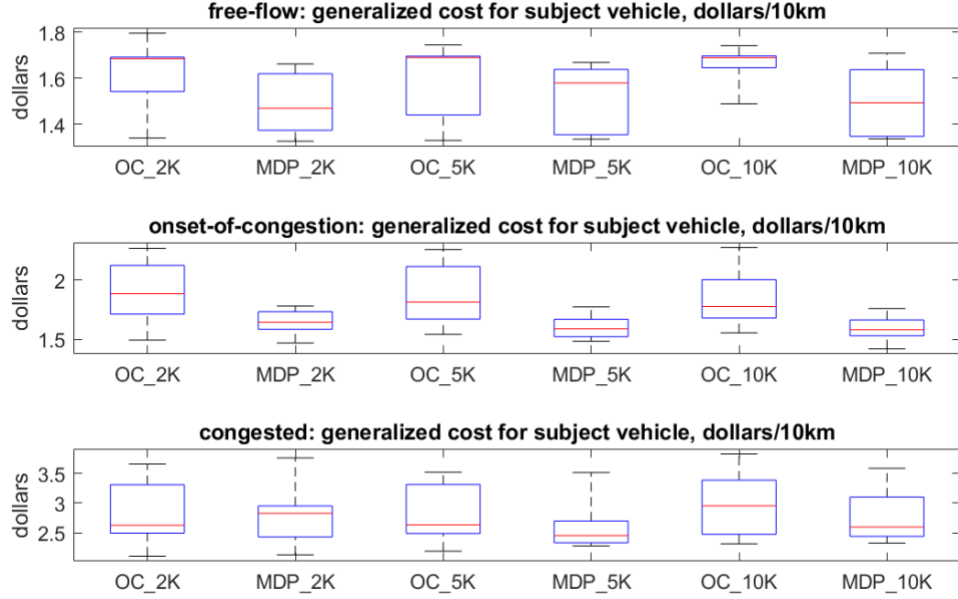


Figure 3: The simulation environment is a circular track. The top, middle and bottom sub-figures represent the free-flow, onset-of-congestion, and congested traffic states, respectively. The vertical axes show the generalized costs with VoT set to 10 dollars per hour. Along the horizontal axes, the generalized costs of the subject vehicle under different controllers in circular tracks of different lengths are compared. Here ‘OC’ and ‘MDP’ denote local optimal and the MDP controllers, respectively. The suffix ‘xK’ indicates that the length of the circular track is x kilometers.

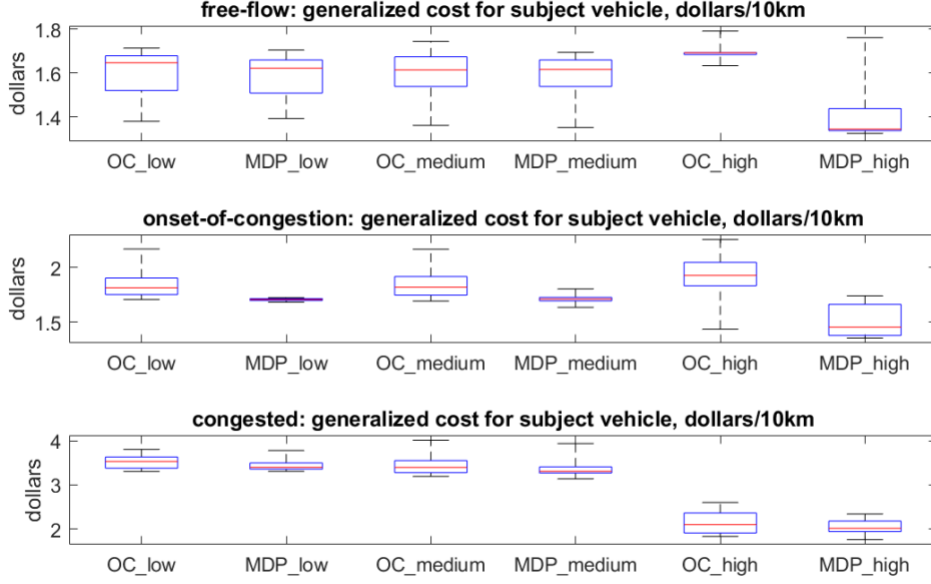


Figure 4: The suffixes ‘low’, ‘medium’ and ‘high’ represent different levels of intensities of platoon-enabled vehicles in the environment. Specifically, ‘low’ indicates that all surrounding vehicles are non-platoon-enabled, ‘medium’ indicates that a proportion (about 30%) of the surrounding vehicles are platoon-enabled, and ‘high’ indicates that all surrounding vehicles are platoon-enabled. Other settings are the same as Figure 3.

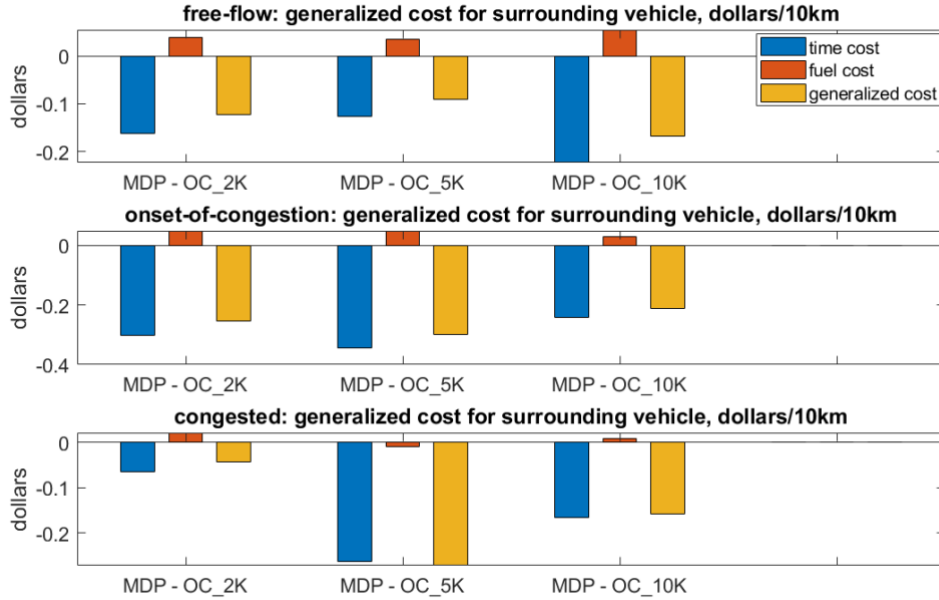


Figure 5: Differences in average time, fuel, and generalized costs of the vehicles upstream to the subject vehicle for different track lengths. A positive value indicates that MDP results in a higher cost than OC, while a negative value indicates that MDP brings more cost savings than OC. The simulation settings are similar to those in Figure 3.

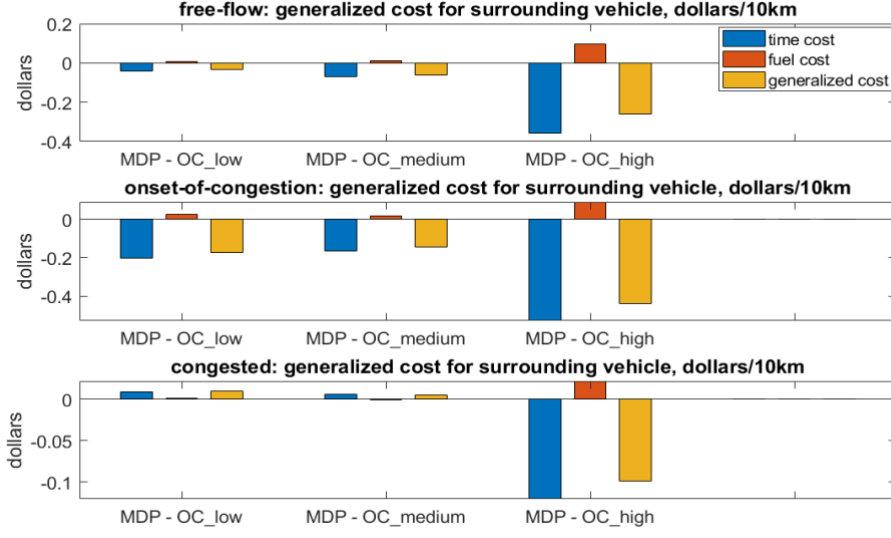


Figure 6: Differences in average time, fuel, and generalized costs of the vehicles upstream to the subject vehicle under different penetration rates of platoon-enabled vehicles. A positive value indicates that MDP results in a higher cost than OC, while a negative value indicates that MDP brings more cost savings than OC. The simulation settings are similar to those in Figure 4.

5.3. A Two-lane Highway Scenario

In this highway scenario, we adopt the same surrounding environment setting as in (X. Liu et al. 2021, in press). Surrounding vehicles can change lanes, merge/exit from the highway, and join into/split from a platoon. Figure (7) demonstrates the generalized costs of different controllers, where the number in the controller name is the value of α , i.e., the discount factor used in the MDP model. This figure shows that in all traffic states, the larger the discount factor (i.e., the more weight on the expectation of the long term cost), the smaller the cost for the subject vehicle along the entire trip, which highlights the importance of accounting for the long-term trip cost. Figure (8) shows the generalized cost of the surrounding vehicles. In the free-flow traffic state, the MDP controller results in significantly smaller cost for the surrounding vehicles, and these savings grow as the MDP discount factor increases. However, under the onset-of-congestion and congested traffic states, the OC and MDP controllers do not show significant differences in cost.

Figure (9) shows the generalized costs incurred by the subject vehicle and its immediate downstream vehicles for an example trip in the onset-of-congestion traffic state, as well as the lateral position and platoon membership status of the subject vehicle. The top plot in this figure pertains to the trajectory formed by the OC model, and the bottom plot demonstrates the trajectory devised by the MDP controller. In the top plot, the subject vehicle makes decisions based solely on local information; as such, its trajectory tends to closely follow the trajectory of its downstream vehicle. This figure shows that under the OC controller, the subject vehicle changes to the left lane at about 2950 time steps, and then returns to its original lane at about 3750 time steps, an indicator of short-sighted decisions. The subject vehicle’s platoon membership status also changes frequently starting at about 4600 time steps. These actions disturb the traffic stream and increase the generalized cost of the subject vehicle and its surrounding vehicles. In the bottom plot, the subject vehicle changes to the left lane at an early time, in which it travels for the rest of its trip. The subject vehicle also joins

platoons twice during its trips, but for longer periods of time. In general, the cost of the subject vehicle under the OC controller is much higher than that of the MDP controller.

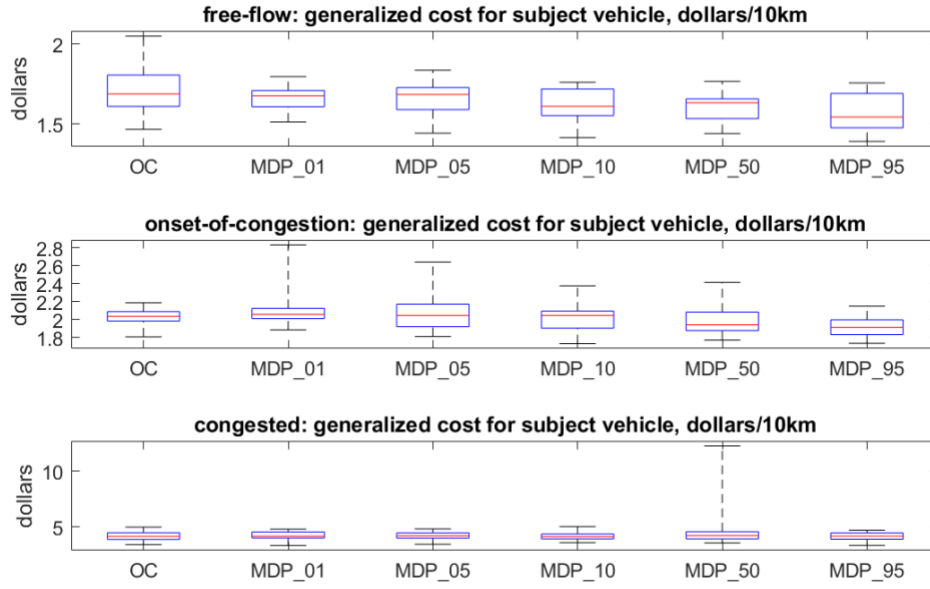


Figure 7: The simulation environment is a highway with on- and off-ramps. The value following ‘MDP ’ in the name of the controller specifies the discount factor, α in the MDP model. Other settings are similar to those in Figure 3.

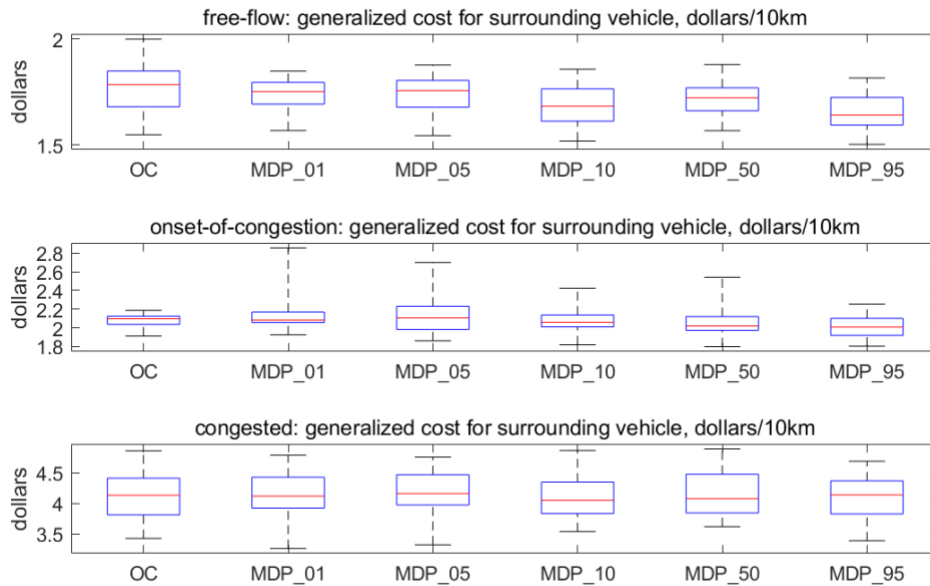


Figure 8: The average generalized cost of the surrounding vehicles. The value following MDP in the name of the controller specifies the discount factor, α in the MDP model. Other settings are similar to those in Figure 7.

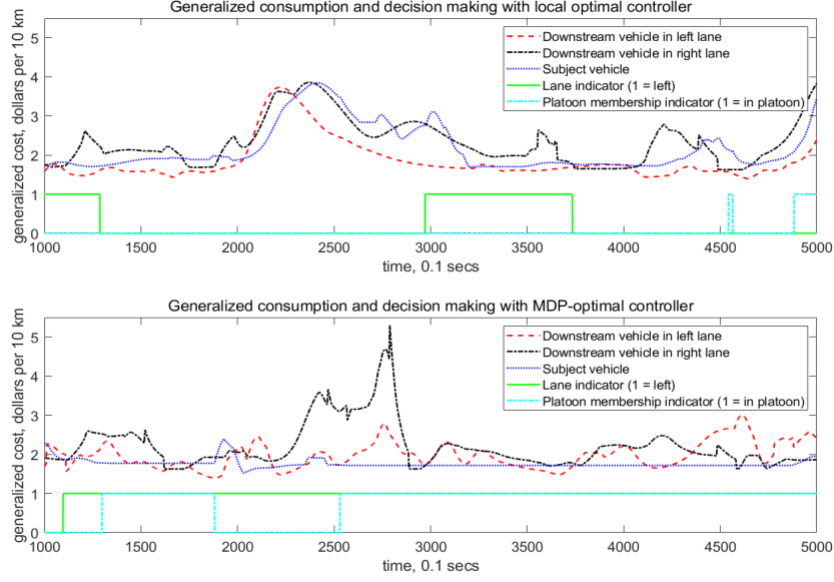


Figure 9: The vertical axis shows the generalized cost, with the unit of dollars per 10 km. The horizontal axis is time, with the unit of 0.1 second. Generalized cost of the subject vehicle and its immediate upstream vehicles, as well as its lane position and platoon membership status are shown. In the top plot, the subject vehicle is traveling under the OC controller, while in the bottom plot, the subject vehicle is traveling under the MDP controller.

5.4. A Network-level Scenario with Route Choice

In these experiments, we show the extensibility of the MDP framework in a joint decision making scenario, in which the framework makes routing, lane-changing, and platoon-merging decisions. In the scenario shown in Figure 10, the subject vehicle has two possible routes to the destination, namely ‘Route1’ and ‘Route2’. Figure 11 shows the results under three scenarios. Under Route1 and Route2, the traveling route is fixed, and the OC model determines the lane changing and platoon merging decisions. Under MDP, the MDP framework makes all three sets of decisions. This figure demonstrates that under all traffic states, the MDP model results in statistically significant savings in the generalized cost compared to the OC model with a fixed route.

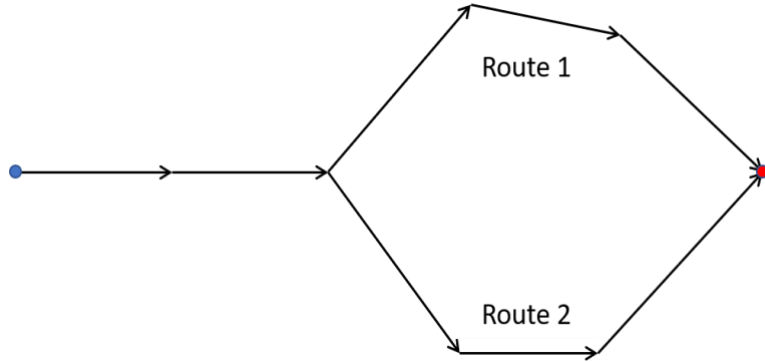


Figure 10: The subject vehicle has two available routes from the origin (blue point) to the destination (red point). Route 1 has a slightly shorter distance, but it is more congested compared with route 2.

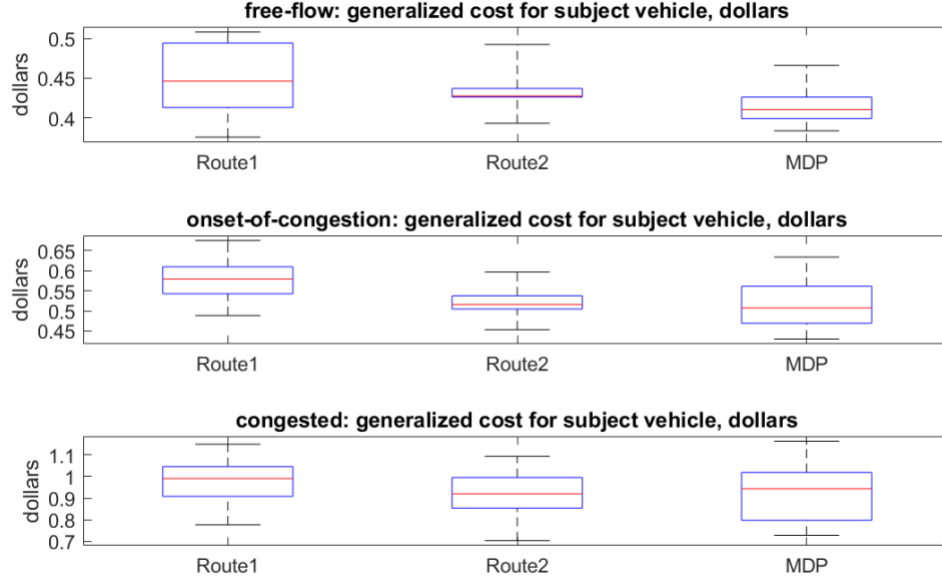


Figure 11: ‘Route1’ and ‘Route2’ refer to scenarios where the subject vehicle will take routes 1 and 2, respectively. In these two scenarios, the OC controller is applied. The ‘MDP’ label refers to the case where the MDP framework selects the adopted route. Other settings are the same as Figure 3.

6. Conclusion

In this paper we proposed a motion planning framework for a CAV in a mixed traffic environment. The framework design leverages an optimal control model to quantify the short-term cost of a trip and an MDP model to capture its long-term cost. This general framework outputs the target acceleration profile of the vehicle as well as routing, platooning and lane changing decisions in a dynamic traffic environment. We implemented this motion planning framework in three experimental scenarios including a highway section with multiple on- and off-ramps, a circular track, and an urban network with route choice, and conducted a comprehensive set of simulations to quantify the long-term benefits the subject vehicle and its surrounding vehicles may experience as a result of incorporating network-level information into the decision-making process. Our experiments indicate that, generally speaking, the MDP framework outperforms a local OC controller in reducing the generalized trip cost. With higher intensity of platoon-enabled vehicles or higher weight on long-term cost (larger discounting factor), the reduction in generalized cost for both the subject vehicle and its upstream vehicles is statistically significant. This significant cost saving, which originates from accounting for network-level conditions, exists in all simulated environments, under various traffic states.

ACKNOWLEDGMENT

The work described in this paper is supported by research grants from the National Science Foundation (CPS-1837245, CPS-1839511, IIS-1724341).

Appendix A. Expected Discounted Cost in Right Lane

For $l_i \neq l_d$, when $\phi_l = Ri$, $\phi_p = 0$, and $d = -1$, the minimum expected discounted cost is given by

$$\begin{aligned}
 V(\mu, Ri, 0, -1) = & \\
 \min_{a \in A} & \begin{cases}
 \Lambda C_{Ri,0,-1}^{Ri,0,-1} + U(\mu', Ri, 0, -1) & a_l = Ri, a_p = 0 \\
 g_l^0(\xi_p) \{ \Lambda C_{Ri,0,-1}^{Ri,1,k-1} + W(\mu', Ri, 1, k-1) \} + \\
 (1 - g_l^0(\xi_p)) \{ \Lambda C_{Ri,0,-1}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} & a_l = Ri, a_p = 1 \\
 q_\phi^f(\mu) \{ \Lambda C_{Ri,0,-1}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} + \\
 (1 - q_\phi^f(\mu)) \{ \Lambda C_{Ri,0,-1}^{Le,0,-1} + U(\mu', Le, 0, -1) \} & a_l = Le, a_p = 0 \\
 g_l^1(\xi_p) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Ri,0,-1}^{Le,1,k-1} + W(\mu', Le, 1, k-1) \} \\
 + (1 - g_l^1(\xi_p)) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Ri,0,-1}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} & a_l = Le, a_p = 1
 \end{cases} \tag{A. 1}
 \end{aligned}$$

The explanation for the case that the subject vehicle is a free agent in the right lane is similar to that in the left lane.

For $l_i \neq l_d$, when $\phi_l = Ri$, $\phi_p = 1$, $d > 0$, the minimum expected discounted cost is given by

$$\begin{aligned}
 V(\mu, Ri, 1, d) = & \\
 \min_{a \in A} & \begin{cases}
 \Lambda C_{Ri,1,d}^{Ri,0,-1} + U(\mu', Ri, 0, -1) & a_l = Ri, a_p = 0 \\
 \Lambda C_{Ri,1,d}^{Ri,1,d-1} + U(\mu', Ri, 1, d-1) & a_l = Ri, a_p = 1 \\
 q_\phi^f(\mu) \{ \Lambda C_{Ri,1,d}^{Ri,1,d-1} + U(\mu', Ri, 1, d-1) \} + \\
 (1 - q_\phi^f(\mu)) \{ \Lambda C_{Ri,1,d}^{Le,0,-1} + U(\mu', Le, 0, -1) \} & a_l = Le, a_p = 0 \\
 (1 - g_l^1(\xi_p)) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Ri,1,d}^{Ri,1,d-1} + U(\mu', Ri, 1, d-1) \} \\
 + g_l^1(\xi_p) (1 - q_\phi^f(\mu)) \{ \Lambda C_{Ri,1,d}^{Le,1,k-1} + W(\mu', Le, 1, k-1) \} & a_l = Le, a_p = 1
 \end{cases} \tag{A. 2}
 \end{aligned}$$

The explanation for the case that the subject vehicle is a member of platoon in the right lane is similar to the case that in the left lane.

For $l_i \neq l_d$, when $\phi_l = Ri$, $\phi_p = 1$, $d = 0$, the minimum expected discounted cost is given by

$$\begin{aligned}
V(\mu, Ri, 1, 0) = & \\
\min_{a \in A} & \left\{ \begin{aligned}
& \Lambda C_{Ri,1,0}^{Ri,0,-1} + U(\mu', Ri, 0, -1) & a_l = Ri, a_p = 0 \\
& g_l^0(\xi_p) \{ \Lambda C_{Ri,1,0}^{Ri,1,k-1} + W(\mu', Ri, 1, k-1) \} + \\
& \left(1 - g_l^0(\xi_p) \right) \{ \Lambda C_{Ri,1,0}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} & a_l = Ri, a_p = 1 \\
& q_\phi^f(\mu) \{ \Lambda C_{Ri,1,0}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} + \\
& \left(1 - q_\phi^f(\mu) \right) \{ \Lambda C_{Ri,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1) \} & a_l = Le, a_p = 0 \\
& g_l^1(\xi_p) \left(1 - q_\phi^f(\mu) \right) \{ \Lambda C_{Ri,1,0}^{Le,1,k-1} + W(\mu', Le, 1, k-1) \} \\
& + \left(1 - g_l^1(\xi_p) \right) \left(1 - q_\phi^f(\mu) \right) \{ \Lambda C_{Ri,1,0}^{Ri,0,-1} + U(\mu', Ri, 0, -1) \} & a_l = Le, a_p = 1
\end{aligned} \right. \quad (A. 3)
\end{aligned}$$

The explanation for the case that the subject vehicle is in the right lane is similar to the last case that it is in the left lane.

Appendix B. Parameters for Generating Simulations

Table B.1: Summary of parameters

Parameter	Value	Definition
t_{upd}	0.4 secs	the updating period of the trajectory of the subject vehicle
p_{on}	0.6	the probability that a vehicle is interested in joining the freeway from an on-ramp
p_{off}	0.6	the probability that a vehicle is interested in taking an off-ramp
p_{npe}	0.5	the probability that the vehicle is not platoon-enabled
p_{merge}	0.6	the probability of that a vehicle intends to merge
p_{change}	0.1	the probability of that the vehicle intends to change lane
t_p	3.5 secs	the time gap between two successive vehicles not in a platoon
t_g	0.55 secs	the time gap between two successive vehicles in a platoon
t_{lcp}	3.6 secs	the period of time within which the surrounding vehicles complete changing lanes
t_{lc}	5 secs	the minimum time interval between two successive lane changes by two successive vehicles in the same lane
τ_s	0.4 secs	the reaction time delay in the car-following model
$\tau_{N_{\text{act}}}$	10 secs	the prediction horizon in the optimal control model
v_m^{le}	20 m/s	the velocity in the left lane at the maximum flow rate
v_m^{ri}	14 m/s	the velocity in the right lane at the maximum flow rate
$v_{\text{max}}^{\text{le}}$	30 m/s	maximum velocity in left lane

Parameter	Value	Definition
$a_{\max}$	2 m/s ²	maximum acceleration for the subject vehicle
$j_{\max}$	3.5 m/s ³	maximum jerk for the subject vehicle
d_{cg}	50 m	critical gap to decide whether it is feasible to change lanes
l_{car}	5 m	length of a vehicle
h_{st}	5 m	vehicle would stop at headway of this value
a	2 m/s ²	the maximum desired acceleration
b	3 m/s ²	the comfortable deceleration
γ_{AR}	0.3987	coefficient for air resistance force
γ_{RR}	281.547	coefficient for rolling resistance force
γ_{GR}	0	coefficient for grade resistance force
γ_{IR}	1750	coefficient for inertial resistance force
η_f	5.98×10^{-8}	fuel cost for a unit energy consumed by the vehicle (dollars/Joule)
P_{sch}	{2, 10, 50}	the scheduled splitting position can be in 2, 10 or 50 road pieces
$N(\mu_{\text{sch}}, \sigma_{\text{sch}})$	$N(2, 5)$, left, $N(-1, 5)$, right	the normal distribution of the scheduled splitting position in the left and right lanes

Appendix C. Sensitivity Analysis over Parameters in Traffic Environment

To demonstrate the performance of our method under various settings, we conduct sensitivity analysis over parameters p_{on} , p_{off} , p_{npe} , p_{merge} and p_{change} in the two-lane highway scenario.

Under univariate analysis, we adjust the value of one parameter at a time while keeping the values of other parameters unchanged. To maintain a relatively steady traffic environment, i.e., to avoid changes in traffic state, we use the same value for p_{on} and p_{off} to balance the number of vehicles entering and exiting the highway.

Figures C.1 and C.2 display the generalized cost when $p_{\text{on}} = p_{\text{off}} = 0.4$ and $p_{\text{on}} = p_{\text{off}} = 0.8$ for the subject vehicle and surrounding vehicles, respectively. Figures C.3 and C.4 demonstrate the generalized costs when $p_{\text{npe}} = 0.1$ and $p_{\text{npe}} = 0.9$, respectively. Figures C.5 and C.6 correspond to the cases where $p_{\text{merge}} = 0.4$ and $p_{\text{merge}} = 0.8$. Figures C.7 and C.8 show the cost when $p_{\text{change}} = 0.05$ and $p_{\text{change}} = 0.3$.

Under all these settings, our MDP framework generally results in statistically significant cost savings for subject vehicle and its surrounding vehicles in free-flow and onset-of-congestion states, and there is no significant difference in the congested state.

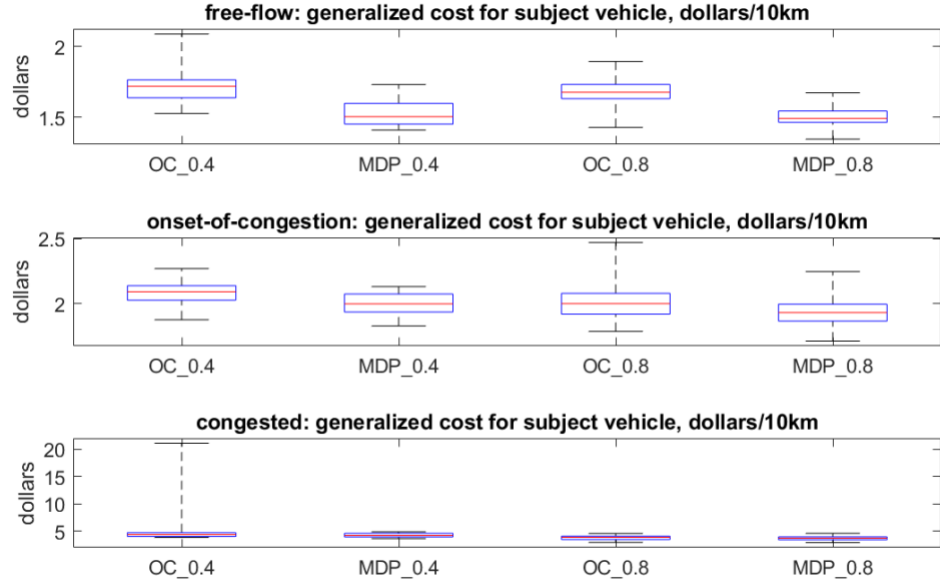


Figure C. 1: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{on} and p_{off} . Other settings are similar to those in Figure 7.

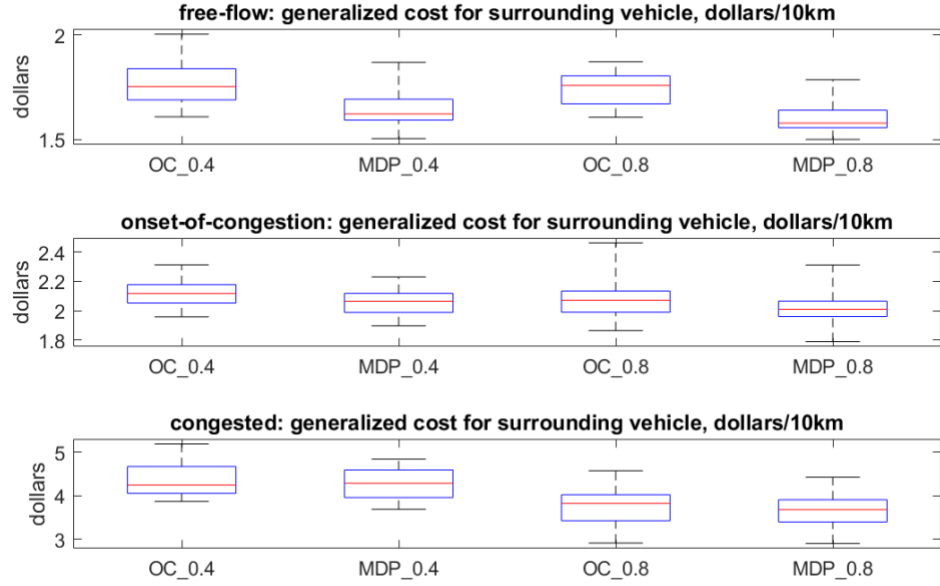


Figure C. 2: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{on} and p_{off} . Other settings are similar to those in Figure 8.

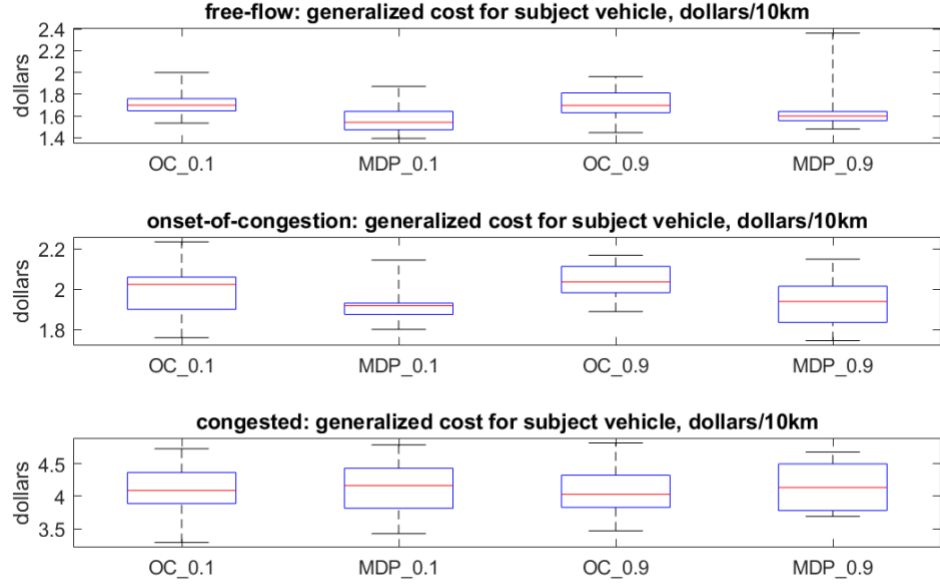


Figure C. 3: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{npe} . Other settings are similar to those in Figure 7.

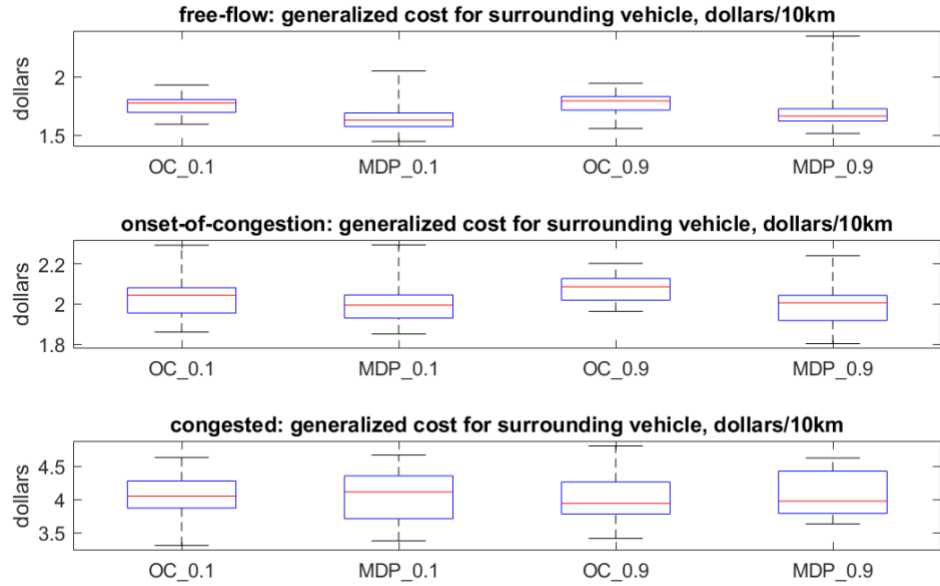


Figure C. 4: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{npe} . Other settings are similar to those in Figure 8.

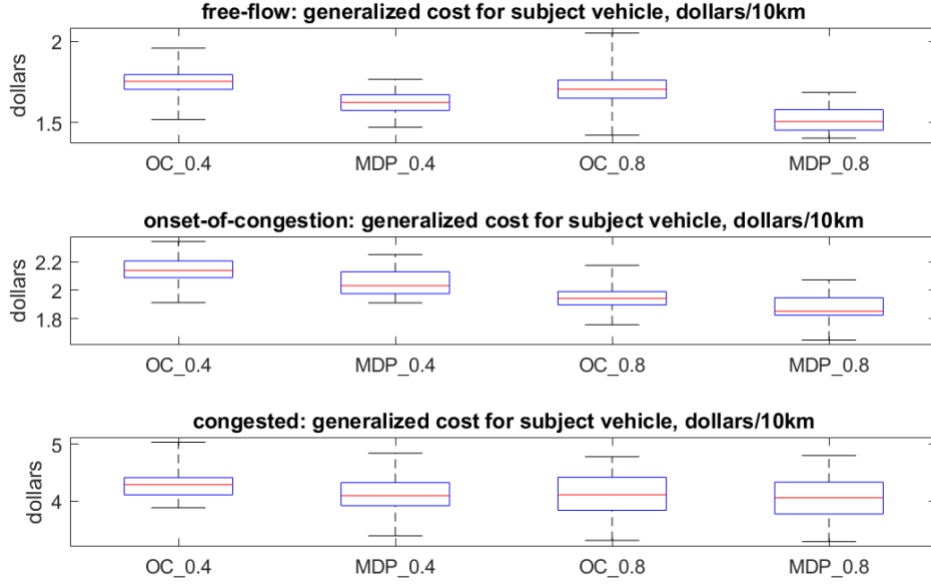


Figure C. 5: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{merge} . Other settings are similar to those in Figure 7.

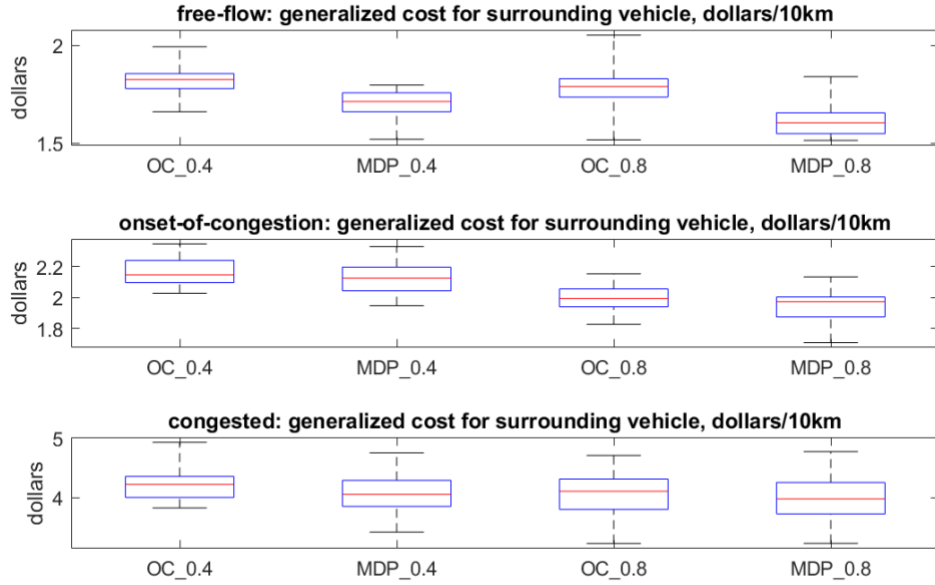


Figure C. 6: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{merge} . Other settings are similar to those in Figure 8.

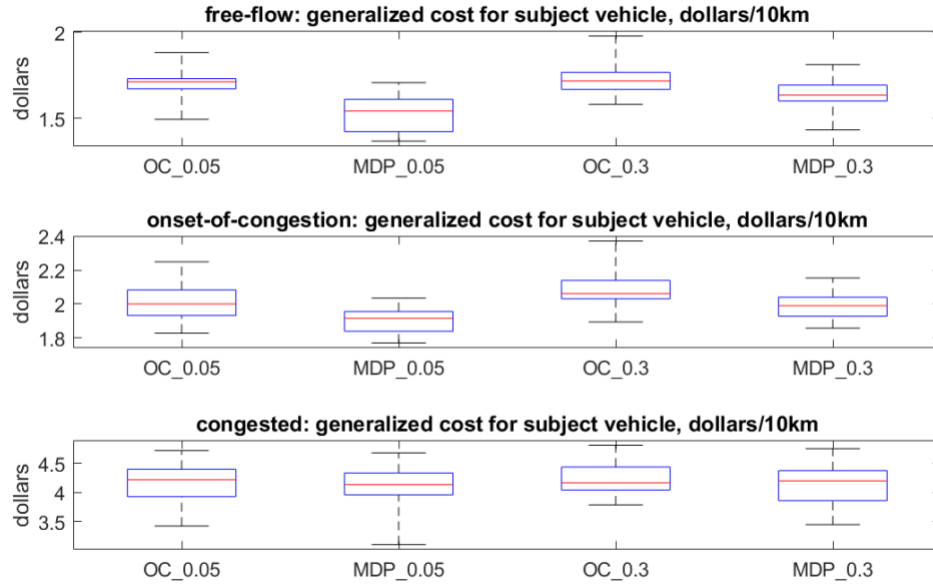


Figure C. 7: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{change} . Other settings are similar to those in Figure 7.

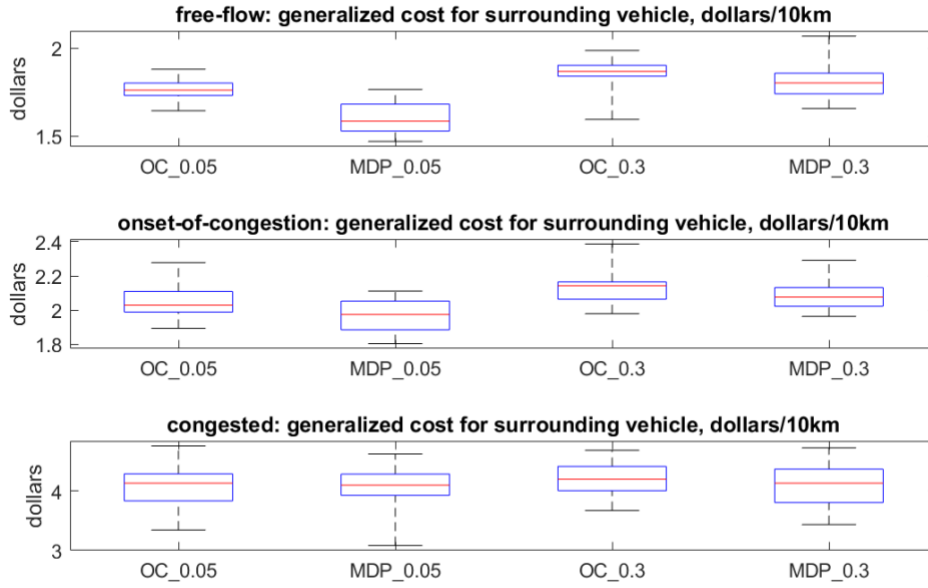


Figure C. 8: The value following ‘OC_’ or ‘MDP_’ in the name of the controller specifies the value of p_{change} .

References

- Ahn, Kyoungcho, and Hesham A Rakha. 2013. “Network-Wide Impacts of Eco-Routing Strategies: A Large-Scale Case Study.” *Transportation Research Part D: Transport and Environment* 25: 119–30.
- Alia, Chebly, Tagne Gilles, Talj Reine, and Charara Ali. 2015. “Local Trajectory Planning and Tracking of Autonomous Vehicles, Using Clothoid Tentacles Method.” In *2015 IEEE Intelligent Vehicles Symposium (IV)*, 674–79. IEEE.
- Bandyopadhyay, Tirthankar, Kok Sung Won, Emilio Frazzoli, David Hsu, Wee Sun Lee, and Daniela Rus. 2013. “Intention-Aware Motion Planning.” In *Algorithmic Foundations of Robotics x*, 475–91. Springer.
- Bellman, Richard. 1957. “A Markovian Decision Process.” *Journal of Mathematics and Mechanics*, 679–84.
- Bhoopalam, Anirudh Kishore, Niels Agatz, and Rob Zuidwijk. 2018. “Planning of Truck Platoons: A Literature Review and Directions for Future Research.” *Transportation Research Part B: Methodological* 107: 212–28.
- Boriboonsomsin, Kanok, Matthew J Barth, Weihua Zhu, and Alexander Vu. 2012. “Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information.” *IEEE Transactions on Intelligent Transportation Systems* 13 (4): 1694–704.
- Brehtel, Sebastian, Tobias Gindele, and Rüdiger Dillmann. 2011. “Probabilistic MDP-Behavior Planning for Cars.” In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 1537–42. IEEE.
- . 2014. “Probabilistic Decision-Making Under Uncertainty for Autonomous Driving Using Continuous POMDPs.” In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 392–99. IEEE.
- Cheng, JiuJun, JunLu Cheng, MengChu Zhou, FuQiang Liu, ShangCe Gao, and Cong Liu. 2015. “Routing in Internet of Vehicles: A Review.” *IEEE Transactions on Intelligent Transportation Systems* 16 (5): 2339–52.
- Claussmann, Laurène, Marc Revilloud, Dominique Gruyer, and Sébastien Glaser. 2019. “A Review of Motion Planning for Highway Autonomous Driving.” *IEEE Transactions on Intelligent Transportation Systems*.
- De Nunzio, Giovanni, Carlos Canudas De Wit, Philippe Moulin, and Domenico Di Domenico. 2016. “Eco-Driving in Urban Traffic Networks Using Traffic Signals Information.” *International Journal of Robust and Nonlinear Control* 26 (6): 1307–24.
- Duret, Aurelien, Meng Wang, and Andres Ladino. 2019. “A Hierarchical Approach for Splitting Truck Platoons Near Network Discontinuities.” *Transportation Research Part B: Methodological*.
- Galceran, Enric, Alexander G Cunningham, Ryan M Eustice, and Edwin Olson. 2015. “Multipolicy Decision-Making for Autonomous Driving via Changepoint-Based Behavior Prediction.” In *Robotics: Science and Systems*. Vol. 1. 2.
- Gillespie, Thomas D. 1992. “Fundamentals of Vehicle Dynamics.” SAE Technical Paper.

- González, David, Joshué Pérez, Vicente Milanés, and Fawzi Nashashibi. 2015. "A Review of Motion Planning Techniques for Automated Vehicles." *IEEE Transactions on Intelligent Transportation Systems* 17 (4): 1135–45.
- Gritschneider, Franz, Knut Graichen, and Klaus Dietmayer. 2018. "Fast Trajectory Planning for Automated Vehicles Using Gradient-Based Nonlinear Model Predictive Control." In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 7369–74. IEEE.
- Guo, Ge, and Qiong Wang. 2018. "Fuel-Efficient En Route Speed Planning and Tracking Control of Truck Platoons." *IEEE Transactions on Intelligent Transportation Systems* 20 (8): 3091–3103.
- Guo, Hongyan, Chen Shen, Hui Zhang, Hong Chen, and Rui Jia. 2018. "Simultaneous Trajectory Planning and Tracking Using an MPC Method for Cyber-Physical Systems: A Case Study of Obstacle Avoidance for an Intelligent Vehicle." *IEEE Transactions on Industrial Informatics* 14 (9): 4273–83.
- Hardy, Jason, and Mark Campbell. 2013. "Contingency Planning over Probabilistic Obstacle Predictions for Autonomous Road Vehicles." *IEEE Transactions on Robotics* 29 (4): 913–29.
- Huang, Chao, Fazel Naghdy, and Haiping Du. 2016. "Model Predictive Control-Based Lane Change Control System for an Autonomous Vehicle." In *2016 IEEE Region 10 Conference (TENCON)*, 3349–54. IEEE.
- Huang, Ke, Xianfeng Yang, Yang Lu, Chunting Chris Mi, and Prathyusha Kondlapudi. 2018. "Ecological Driving System for Connected/Automated Vehicles Using a Two-Stage Control Hierarchy." *IEEE Transactions on Intelligent Transportation Systems* 19 (7): 2373–84.
- Huang, Zichao, Duanfeng Chu, Chaozhong Wu, and Yi He. 2018. "Path Planning and Cooperative Control for Automated Vehicle Platoon Using Hybrid Automata." *IEEE Transactions on Intelligent Transportation Systems* 20 (3): 959–74.
- Jin, I Ge, and Gábor Orosz. 2016. "Optimal Control of Connected Vehicle Systems with Communication Delay and Driver Reaction Time." *IEEE Transactions on Intelligent Transportation Systems* 18 (8): 2056–70.
- Kamrani, Mohsen, Aravinda Ramakrishnan Srinivasan, Subhadeep Chakraborty, and Asad J Khattak. 2020. "Applying Markov Decision Process to Understand Driving Decisions Using Basic Safety Messages Data." *Transportation Research Part C: Emerging Technologies* 115: 102642.
- Katrakazas, Christos, Mohammed Quddus, Wen-Hua Chen, and Lipika Deka. 2015. "Real-Time Motion Planning Methods for Autonomous on-Road Driving: State-of-the-Art and Future Research Directions." *Transportation Research Part C: Emerging Technologies* 60: 416–42.
- Kenney, John B. 2011. "Dedicated Short-Range Communications (DSRC) Standards in the United States." *Proceedings of the IEEE* 99 (7): 1162–82.
- Larsson, Erik, Gustav Sennton, and Jeffrey Larson. 2015. "The Vehicle Platooning Problem: Computational Complexity and Heuristics." *Transportation Research Part C: Emerging Technologies* 60: 258–77.
- Li, Bai, Youmin Zhang, Yiheng Feng, Yue Zhang, Yuming Ge, and Zhijiang Shao. 2018. "Balancing Computation Speed and Quality: A Decentralized Motion Planning Method for Cooperative Lane Changes of Connected and Automated Vehicles." *IEEE Transactions on Intelligent Vehicles* 3 (3): 340–50.

- Li, Xiaohui, Zhenping Sun, Dongpu Cao, Daxue Liu, and Hangen He. 2017. "Development of a New Integrated Local Trajectory Planning and Tracking Control Framework for Autonomous Ground Vehicles." *Mechanical Systems and Signal Processing* 87: 118–37.
- Lioris, Jennie, Ramtin Pedarsani, Fatma Yildiz Tascikaraoglu, and Pravin Varaiya. 2017. "Platoons of Connected Vehicles Can Double Throughput in Urban Roads." *Transportation Research Part C: Emerging Technologies* 77: 292–305.
- Liu, Changliu, Chung-Yen Lin, and Masayoshi Tomizuka. 2018. "The Convex Feasible Set Algorithm for Real Time Optimization in Motion Planning." *SIAM Journal on Control and Optimization* 56 (4): 2712–33.
- Liu, Xiangguo, Guangchen Zhao, Neda Masoud, and Qi Zhu. 2021, in press. "Trajectory Planning for Connected and Automated Vehicles: Cruising, Lane Changing, and Platooning." *SAE International Journal of Connected and Automated Vehicles*, 2021, in press.
- Maiti, Santa, Stephan Winter, and Lars Kulik. 2017. "A Conceptualization of Vehicle Platoons and Platoon Operations." *Transportation Research Part C Emerging Technologies* 80: 1–19.
- Milanés, Vicente, and Steven E Shladover. 2014. "Modeling Cooperative and Autonomous Adaptive Cruise Control Dynamic Responses Using Experimental Data." *Transportation Research Part C: Emerging Technologies* 48: 285–300.
- Mouhagir, Hafida, Reine Talj, Véronique Cherfaoui, François Aioun, and Franck Guillemard. 2016. "Integrating Safety Distances with Trajectory Planning by Modifying the Occupancy Grid for Autonomous Vehicle Navigation." In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 1114–19. IEEE.
- Neunert, Michael, Cedric De Crousaz, Fadri Furrer, Mina Kamel, Farbod Farshidian, Roland Siegwart, and Jonas Buchli. 2016. "Fast Nonlinear Model Predictive Control for Unified Trajectory Optimization and Tracking." In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 1398–1404. IEEE.
- Orosz, Gábor. 2016. "Connected Cruise Control: Modelling, Delay Effects, and Nonlinear Behaviour." *Vehicle System Dynamics* 54 (8): 1147–76.
- Paden, Brian, Michal Čáp, Sze Zheng Yong, Dmitry Yershov, and Emilio Frazzoli. 2016. "A Survey of Motion Planning and Control Techniques for Self-Driving Urban Vehicles." *IEEE Transactions on Intelligent Vehicles* 1 (1): 33–55.
- Paikari, Elahe, Lina Kattan, Shahram Tahmasseby, and Behrouz H Far. 2013. "Modeling and Simulation of Advisory Speed and Re-Routing Strategies in Connected Vehicles Systems for Crash Risk and Travel Time Reduction." In *2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 1–4. IEEE.
- Qian, Xiangjun, Florent Althé, Philipp Bender, Christoph Stiller, and Arnaud de La Fortelle. 2016. "Optimal Trajectory Planning for Autonomous Driving Integrating Logical Constraints: An MIQP Perspective." In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 205–10. IEEE.
- Qian, Xiangjun, Arnaud De La Fortelle, and Fabien Moutarde. 2016. "A Hierarchical Model Predictive Control Framework for on-Road Formation Control of Autonomous Vehicles." In *2016 IEEE Intelligent Vehicles Symposium (IV)*, 376–81. IEEE.

- Rakha, Hesham A, Kyoungcho Ahn, and Kevin Moran. 2012. "Integration Framework for Modeling Eco-Routing Strategies: Logic and Preliminary Results." *International Journal of Transportation Science and Technology* 1 (3): 259–74.
- Rakha, Hesham, and Raj Kishore Kamalanathsharma. 2011. "Eco-Driving at Signalized Intersections Using V2i Communication." In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 341–46. IEEE.
- Shladover, Steven E, Christopher Nowakowski, Xiao Yun Lu, and Robert Ferlis. 2015. "COOPERATIVE ADAPTIVE CRUISE CONTROL (CACC) DEFINITIONS AND OPERATING CONCEPTS." In *Trb Conference*.
- Stern, Raphael E, Shumo Cui, Maria Laura Delle Monache, Rahul Bhadani, Matt Bunting, Miles Churchill, Nathaniel Hamilton, et al. 2018. "Dissipation of Stop-and-Go Waves via Control of Autonomous Vehicles: Field Experiments." *Transportation Research Part C: Emerging Technologies* 89: 205–21.
- Sugiyama, Yuki, Minoru Fukui, Macoto Kikuchi, Katsuya Hasebe, Akihiro Nakayama, Katsuhiro Nishinari, Shin-ichi Tadaki, and Satoshi Yukawa. 2008. "Traffic Jams Without Bottlenecks—Experimental Evidence for the Physical Mechanism of the Formation of a Jam." *New Journal of Physics* 10 (3): 033001.
- Sutton, Richard S, and Andrew G Barto. 2018. *Reinforcement Learning: An Introduction*. MIT press.
- Tadaki, Shin-ichi, Macoto Kikuchi, Minoru Fukui, Akihiro Nakayama, Katsuhiro Nishinari, Akihiro Shibata, Yuki Sugiyama, Taturu Yosida, and Satoshi Yukawa. 2013. "Phase Transition in Traffic Jam Experiment on a Circuit." *New Journal of Physics* 15 (10): 103034.
- Wang, Ziran, Guoyuan Wu, and Matthew J. Barth. 2018. "A Review on Cooperative Adaptive Cruise Control (CACC) Systems: Architectures, Controls, and Applications."
- Wei, Yuguang, Cafer Avci, Jiangtao Liu, Baloka Belezamo, Nizamettin Aydın, Pengfei Taylor Li, and Xuesong Zhou. 2017. "Dynamic Programming-Based Multi-Vehicle Longitudinal Trajectory Optimization with Simplified Car Following Models." *Transportation Research Part B: Methodological* 106: 102–29.
- Wyk, Franco van, Anahita Khojandi, and Neda Masoud. 2020. "Optimal Switching Policy Between Driving Entities in Semi-Autonomous Vehicles." *Transportation Research Part C: Emerging Technologies* 114: 517–31.
- You, Changxi, Jianbo Lu, Dimitar Filev, and Panagiotis Tsiotras. 2018. "Highway Traffic Modeling and Decision Making for Autonomous Vehicle Using Reinforcement Learning." In *2018 IEEE Intelligent Vehicles Symposium (IV)*, 1227–32. IEEE.
- Zeng, Xiangrui, and Junmin Wang. 2018. "Globally Energy-Optimal Speed Planning for Road Vehicles on a Given Route." *Transportation Research Part C: Emerging Technologies* 93: 148–60.
- Zhang, Linjun, Jing Sun, and Gábor Orosz. 2017. "Hierarchical Design of Connected Cruise Control in the Presence of Information Delays and Uncertain Vehicle Dynamics." *IEEE Transactions on Control Systems Technology* 26 (1): 139–50.
- Zheng, Zuduo. 2014. "Recent Developments and Research Needs in Modeling Lane Changing." *Transportation Research Part B: Methodological* 60: 16–32.

Zhou, Jian, Hongyu Zheng, Junmin Wang, Yulei Wang, Bing Zhang, and Qian Shao. 2019. “Multi-Objective Optimization of Lane-Changing Strategy for Intelligent Vehicles in Complex Driving Environments.” *IEEE Transactions on Vehicular Technology*.

Zhou, Shiyong, Yizhou Wang, Minghui Zheng, and Masayoshi Tomizuka. 2017. “A Hierarchical Planning and Control Framework for Structured Highway Driving.” *IFAC-PapersOnLine* 50 (1): 9101–7.