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

Xiangguo Liu, Ph.D. Student
Department of Electrical and Computer Engineering
Northwestern University, Evanston 60201
xiangguoliu2023@u.northwestern.edu

Neda Masoud (corresponding author), Assistant Professor
Department of Civil and Environmental Engineering
University of Michigan, Ann Arbor 48109
nmasoud@umich.edu

Qi Zhu, Associate Professor
Department of Electrical and Computer Engineering
Northwestern University, Evanston 60201
qzhu@northwestern.edu

Anahita Khojandi, Assistant Professor
Department of Industrial and Systems Engineering
University of Tennessee, Knoxville 37996
khojandi@utk.edu

6,929 (words) + 250 (1 table) = 7,179

Aug. 1, 2020

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, have been able to 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 model that can capture network-level information. By optimizing for a combined short- and long-term fuel and time cost, our proposed framework can guarantee safety and minimize the generalized cost of an entire trip. To showcase the benefits of incorporating network-level data when devising trajectories, we conduct a comprehensive simulation study in two experimental settings, namely a straight highway with on- and off-ramps, and a small network with route choice. The simulation results indicate that further statistically significant efficiency can be obtained for the subject vehicle and its surrounding vehicles in different traffic states under all experimental settings.

Keywords: Connected and Automated Vehicles, Trajectory planning

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) (1) 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) (2, 3, 4), Connected Cruise Control (CCC) (5, 6) and Platooning (7, 8, 9, 10). 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 where local information is leveraged to make locally-optimal decisions (11). 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-wide information with granular local information to devise globally optimal cruising, routing, lane-changing, and platoon-merging decisions for a CAVs in a mixed traffic, as shown in Figure 1. As demonstrated in this figure, the proposed framework combines an optimal control (OC) trajectory planning model proposed in (12) with a Markov decision process (MDP) model developed in this paper to devise an efficient trajectory for an entire trip. This Markov decision process model can capture the progression of traffic as a stochastic process at an aggregate level, thereby complementing 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 (12), 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.

RELATED WORKS

Motion Planning

Motion planning for automated vehicles has been an active research topic (11, 13, 14, 15). 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. (16) reviewed the planning and control techniques in an urban environment, Claussmann et al. (17) reviewed motion planning techniques for highway driving.

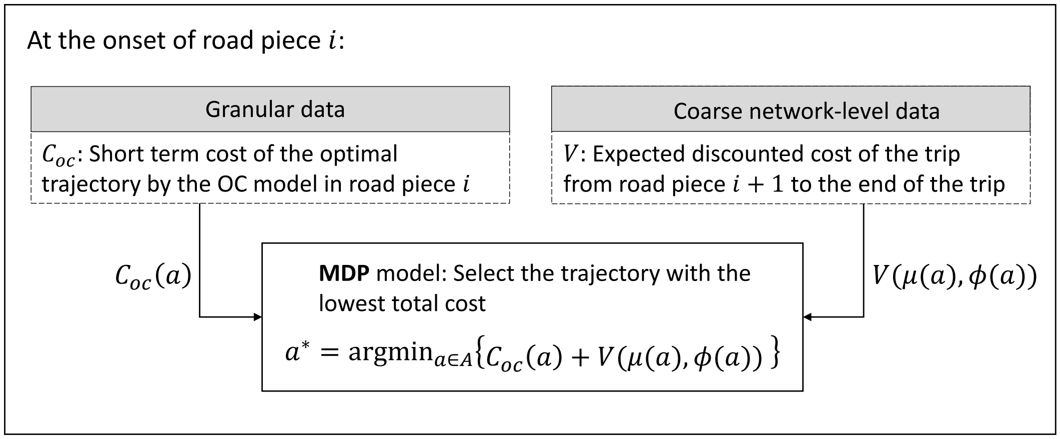

FIGURE 1 : Structure of the proposed MDP framework. The optimal control (OC) model plans a trajectory to determine the short-term cost (12) associated with a higher-layer action $a \in A$, which includes a combination of route choice, lane changing, and platoon merging. The MDP model assesses the long-term cost associated with the same action 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.

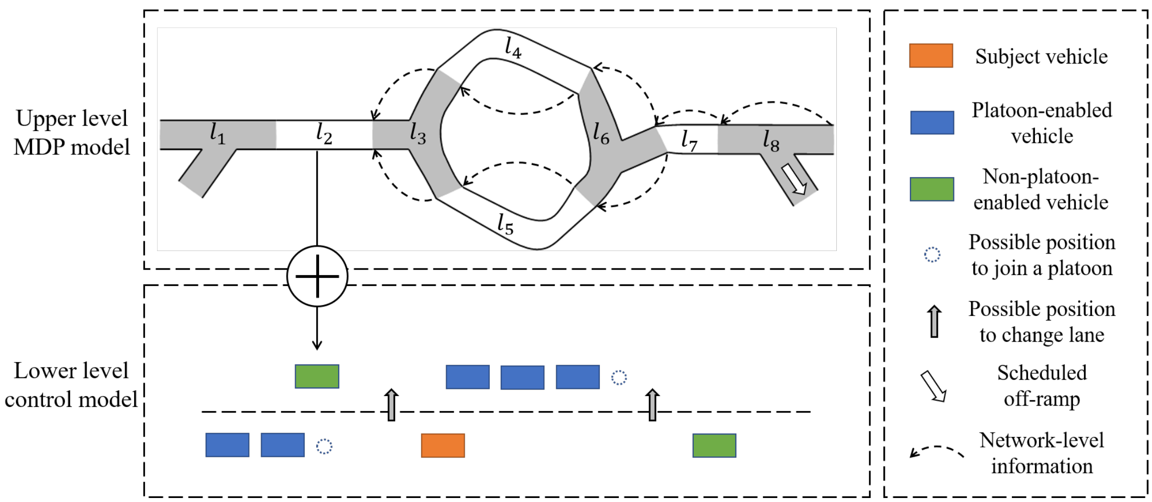

FIGURE 2 : The upper figure displays a freeway stretch segmented into merge (on-ramp), diverge (off-ramp), and regular road pieces, where the MDP model operates. The lower figure displays a zoomed down view of the current road piece, where the cost of each action (i.e., lane-changing and platoon-merging decisions) are determined based on local information.

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 (18, 19), some of which have considered the uncertainty of surrounding vehicles' motion (20, 21). Besides safety guarantee, efficiency manifested in the form of reducing travel time (22) and fuel consumption (22, 23, 24), or increasing traffic flow (25, 26) 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.

In the literature, hierarchical design is sometimes referred to as the combination of trajectory planning and tracking (18, 27, 28), and sometimes as the combination of long- and short-horizon planning. 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. More specifically, we consider route choice, lane changing, and platoon merging as higher-layer decisions, and cruising decisions, in which the acceleration profile of the vehicle is determined, as lower-layer decisions. There have been several attempts in the literature to conduct longer-term horizon planning using hierarchical design, under specific assumptions. 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 (29, 30, 31). 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 (32). 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 plan at either the lower- or higher-layer, or both, may also be outdated at the time of execution in a fast-changing traffic environment. To combat outdated decisions, Boriboonsomsin et al. (24) proposed to update the higher-layer plan, while Alia et al. (18) considered updating the lower-layer plan. Huang et al. (29) 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 network-level evolution of traffic. Additionally, our framework is more comprehensive as it includes decision making for routing, lane-changing, platooning, and cruising.

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 discounted cumulative reward in a stochastic environment (33, 34).

In the transportation field, MDP and its variant, partially observable Markov decision process (POMDP), have also been applied for vehicle behavior analysis and prediction (35, 36) and driving entity switching policy (37). Brechtel et al. (32) 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.

The studies above mostly employ MDPs to determine the velocity of the subject vehicle, leaving out higher-layer decisions. A recent work (38) developed a hierarchical framework in which an MDP model is 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 the trajectory by extending the MDP model to a more general motion planner, which includes routing, lane-changing, and platoon-merging. 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 demonstrate that our proposed method results in statistically significant reductions in the long-term generalized trip cost using simulations.

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 (12); 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 incorporating the 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 more importantly its surrounding vehicles, which may be CAVs or legacy vehicles, experience second-hand cost-reduction benefit. 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.

METHODOLOGY

Overview

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, and coarse decisions are made by an MDP model. For each coarse action (where a coarse action is a unique combination 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 obtains the long-term cost of completing the same action. Finally,the action that provides the the lowest total cost will be selected and pursued by the MDP framework. This framework is demonstrated in Figure 1.

An example network is displayed in Figure 2, 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 path 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-layer 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 and travel time. 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 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 2, for example, l_1 is an on-ramp or merge piece, while l_4 and l_5 are regular pieces. The trajectory planning model operates within a short prediction horizon and determines the cost of each action within the prediction horizon in real time. The local cost of the trip under a fixed set of coarse actions will be combined with long-term consequences of these actions with the MDP framework, giving rise to the final optimal solution.

The trajectory planning models are re-optimized dynamically, while 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. The optimal trajectory model guides the trajectory of the subject vehicle within a short prediction horizon. The MDP model uses the policy table to look up the long-term cost of following such a trajectory. Ultimately, the optimal trajectory to be followed is the one that minimizes sum of both short- and long-term costs. This process is repeated at period t_{upd}, which is set to 0.4 sec following (12). The MDP model is elaborated in the next section.

The MDP Framework

In this context, three traffic states, namely, free-flow, onset-of-congestion, and congested traffic, are considered. The traveling speed of the subject vehicle is determined based on the traffic state of the road piece the vehicle is traversing. Every time 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 $S_l(l_i) = \{l_{i1}^{'}, l_{i2}^{'}, l_{i3}^{'}, \ldots\}$ is the set of road pieces connected to l_i.

Let $s = (\mu, \phi) \in S$ denote the state of the traffic dynamics process. Vector $\mu = [l, \xi_{\text{tr}}, \xi_p]$ in this process denotes the location-dependent environment state, where $l \in L$ denotes the location of the subject vehicle, 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 possible paths that connect the origin to the destination. In addition, the vector $\xi_{\text{tr}} = [\xi_{\text{tr}}^{Le}, \xi_{\text{tr}}^{Ri}]$ 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.

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 (12) 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 ξ_{tr}, $v(\xi_{\text{tr}}, \phi_l)$, i.e.,

$$c^t(s) = \text{len}_i / v(\xi_{\text{tr}}, \phi_l) \tag{1}$$

Let $c^{di}(s)$ denote the cost associated with passenger discomfort/safety risk for lane changing. The passenger discomfort/safety cost is assumed to increase linearly with the number of lane changes. Therefore,

$$c^{di}(s) = N_{lc} \tag{2}$$

where $N_{lc} \in \{0, 1\}$ is the number of lane changes in the current road piece.

Let C_s denote the sum of all three costs discussed above, then C_s can be formulated as:

$$C_s = \Lambda C(s) = \lambda_f c^f(s) + \lambda_t c^t(s) + \lambda_{di} c^{di}(s) \tag{3}$$

where the vector $\Lambda = [\lambda_f, \lambda_t, \lambda_{di}]$ contains the corresponding coefficients for each cost component, and $C(s) = [c^f(s), c^t(s), c^{di}(s)]^\top$ is the cost vector at state s. We assume that Λ can be different

TABLE 1 : Table of notation

Notation	Definition
Le, Re	Left and right lanes, respectively
$l_i \in L$	Road piece i, which is a member of the set of road pieces L
$l_{ij}^{'}$	The jth road piece directly connected to road piece l_i
$S_l(l_i)$	Set of road pieces directly connected to road piece l_i
l_o, l_d	The road piece at the origin and destination of the trip, respectively
$\xi_{\text{tr}}^{Le}, \xi_{\text{tr}}^{Ri}$	Macroscopic state of traffic in the left and right lanes, respectively
$\xi_{\text{tr}} = [\xi_{\text{tr}}^{Le}, \xi_{\text{tr}}^{Ri}]$	Vector specifying the macroscopic state of traffic
ξ_p^{Le}, ξ_p^{Ri}	Percentage of platoon-enabled vehicles in the left and right lanes, respectively
$\xi_p = [\xi_p^{Le}, \xi_p^{Ri}]$	Vector specifying the percentage of platoon-enabled vehicles
$\mu = [l, \xi_{\text{tr}}, \xi_p]$	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	Number of road pieces to the platoon scheduled splitting ($d = -1$ for a free agent)
$s = (\mu, \phi) \in S$	State of the traffic dynamics process
$c^f(s)$, $c^{di}(s)$, $c^t(s)$	The fuel, comfort, and time costs of the subject vehicle at state s, respectively
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
$C(s)$	Cost vector at the state s, i.e., $[c^f(s), c^t(s), c^{di}(s)]^\top$
$C_s = \Lambda C(s)$	Sum of fuel, time, and comfort/safety costs
$V([l, \xi_{\text{tr}}, \xi_p], [\phi_l, \phi_p, d])$	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
$g_l^1(\xi_p)$	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, to travel within a platoon
Transition matrices	
$p_l^{Le}((\xi_{\text{tr}}^{Le})^{'} \mid \xi_{\text{tr}})$	Probability that the traffic state transitions from ξ_{tr}^{Le} to $(\xi_{\text{tr}}^{Le})^{'}$ in the left lane in l
$p_l^{Ri}((\xi_{\text{tr}}^{Ri})^{'} \mid \xi_{\text{tr}})$	Probability that the traffic state transitions from ξ_{tr}^{Ri} to $(\xi_{\text{tr}}^{Ri})^{'}$ in the right lane in l
$h_l^{Le}((\xi_p^{Le})^{'} \mid \xi_p)$	Probability that the platoon intensity transitions from ξ_p^{Le} to $(\xi_p^{Le})^{'}$ in the left lane
$h_l^{Ri}((\xi_p^{Ri})^{'} \mid \xi_p)$	Probability that the platoon intensity transitions from ξ_p^{Ri} to $(\xi_p^{Ri})^{'}$ in the right lnae
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
$a = [a_l, a_p, a_r] \in A$	The action, within the action set A, taken by the subject vehicle

for each driver, since different cost terms are of different importance for each driver. The total travel cost C_s describes the generalized cost of travel in a road piece. We assume that the cost c for a road piece can be calculated as a function of the vehicle state in the beginning and end points of a road piece, i.e., C_e^s, where the superscript s denotes the state of the vehicle in the beginning of the road piece and the subscript e denotes its state at the end of the 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 captured by the cost of changing lanes. Notation-wise, we use $\Lambda C_{Le,0,-1}^{Ri,0,-1}$ to denote the cost of the subject vehicle transitioning from state ($\phi_l = Le$, $\phi_p = 0$, $d = -1$) to state ($\phi_l = Ri$, $\phi_p = 0$, $d = -1$).

An important part of the MDP model is the transition probability matrices that allow us to model the dynamics of the system. Let $p_l^{Le}((\xi_{\text{tr}}^{Le})'|\xi_{\text{tr}})$ and $p_l^{Ri}((\xi_{\text{tr}}^{Ri})'|\xi_{\text{tr}})$ denote the probability that the traffic state transitions from ξ_{tr}^{Le} to $(\xi_{\text{tr}}^{Le})'$ in the left lane and from ξ_{tr}^{Ri} to $(\xi_{\text{tr}}^{Ri})'$ in the right lane in road piece l, respectively. Let $h_l^{Le}((\xi_p^{Le})'|\xi_p)$ and $h_l^{Ri}((\xi_p^{Ri})'|\xi_p)$ denote the probability that the platoon intensity transitions from ξ_p^{Le} to $(\xi_p^{Le})'$ in the left lane and from ξ_p^{Ri} 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, which 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 in the original lane. Let $w(k)$ denote the probability distribution for the number of road pieces, k, for which the subject vehicle can stay with a platoon it has met. Note that k is the number of road pieces before the platoon has to split/dissolve. If it merges into an existing platoon, then it will have the same scheduled splitting position as other vehicles in the platoon. If it merges with another vehicle to form a new platoon, the scheduled splitting position will be decided at this time by $w(k)$. We assume that $w(k)$ follows a normal distribution, $w(k) = N(\mu_{Ri}, \sigma_{Ri})$ for the right lane, and $w(k) = N(\mu_{Le}, \sigma_{Le})$ for the left lane. Let l' and $\mu' = [l', \xi_{\text{tr}}', \xi_p']$ denote a candidate road piece directly connected to l and its corresponding environment state vector, respectively. Hence, the problem terminates when the vehicle reaches its destination, i.e., $l = l_d$. Finally, let $V([l, \xi_{\text{tr}}, \xi_p], [\phi_l, \phi_p, d])$ denote the minimum total expected discounted cost starting with the vehicle state ϕ_l, ϕ_p under the environment state $[l, \xi_{\text{tr}}, \xi_p]$, when the next splitting of the platoon of which the subject vehicle is a member is scheduled after d road pieces. Hence, for $l = l_d$, the minimum total expected discounted cost is given by

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

where c_{fl} is a cost incurred should the subject vehicle not be at its destination (e.g., the vehicle should be a single vehicle in the right lane at the target off-ramp piece, or it will miss its exit).

To simplify following formulas, we introduce the following notations:

$$\begin{aligned} U(\mu', \phi_l, 0, -1) = \min_{l' \in S_l(l)} \alpha \sum_{\xi_{\text{tr}}' \xi_p'} & p_l^{Le}((\xi_{\text{tr}}^{Le})'|\xi_{\text{tr}}^{Le}) p_l^{Ri}((\xi_{\text{tr}}^{Ri})'|\xi_{\text{tr}}^{Ri}) \\ & h_l^{Le}((\xi_p^{Le})'|\xi_p^{Le}) h_l^{Ri}((\xi_p^{Ri})'|\xi_p^{Ri}) V(\mu', \phi_l, 0, -1) \end{aligned} \tag{5}$$

$$W(\mu^{'}, \phi_l, 1, k-1) = \min_{l^{'} \in S_l(l)} \alpha \sum_{\xi_{\text{tr}}^{'} \xi_p^{'}} \sum_k w(k) p_l^{Le}((\xi_{\text{tr}}^{Le})^{'} | \xi_{\text{tr}}^{Le}) p_l^{Ri}((\xi_{\text{tr}}^{Ri})^{'} | \xi_{\text{tr}}^{Ri}) \\ h_l^{Le}((\xi_p^{Le})^{'} | \xi_p^{Le}) h_l^{Ri}((\xi_p^{Ri})^{'} | \xi_p^{Ri}) V(\mu^{'}, \phi_l, 1, k-1) \tag{6}$$

where Equations (5) and (6) are the minimum expected discounted cost of the remainder of the trip starting from the next road piece with the vehicle as a free agent and a platoon member, respectively.

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

$$V(\mu, Le, 0, -1) = \\ \min \begin{cases} \Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu^{'}, Le, 0, -1) & a_l = Le, a_p = 0 \\ \\ g_l^0(\xi_p^{Le})\{\Lambda C_{Le,0,-1}^{Le,2,k-1} + W(\mu^{'}, Le, 1, k-1)\} + \\ (1 - g_l^0(\xi_p^{Le}))\{\Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu^{'}, Le, 0, -1)\} & a_l = Le, a_p = 1 \\ \\ q_\phi^f(\mu)\{\Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu^{'}, Le, 0, -1)\} + \\ (1 - q_\phi^f(\mu))\{\Lambda C_{Le,0,-1}^{Ri,0,-1} + U(\mu^{'}, Ri, 0, -1)\} & a_l = Ri, a_p = 0 \\ \\ g_l^1(\xi_p^{Ri})(1 - q_\phi^f(\mu))\{\Lambda C_{Le,0,-1}^{Ri,2,k-1} + W(\mu^{'}, Ri, 1, k-1)\} \\ +(1 - g_l^1(\xi_p^{Ri})(1 - q_\phi^f(\mu)))\{\Lambda C_{Le,0,-1}^{Le,0,-1} + U(\mu^{'}, Le, 0, -1)\} & a_l = Ri, a_p = 1 \end{cases} \tag{7}$$

The four arguments of the min function in Equation (7) correspond to the costs of these four actions. The expected discounted cost (with the initial values as specified) is then the minimum cost of all four actions. The first expression indicates 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 argument of the min function in Equation (7) 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 travel starting from its new state as a platoon member. In case this action fails, the vehicle continues 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 argument of the min function provides the cost under the action of changing to the right lane and remaining a free agent. Similar to the previous case, the first expression 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 argument of the min function 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., $l \neq l_d$, when $\phi_l = Le$, $\phi_p = 1$, $d > 0$), the minimum expected discounted cost is given by

$$V(\mu, Le, 1, d) = \min \begin{cases} \Lambda C_{Le,1,d}^{Le,0,-1} + U(\mu', Le, 0, -1) & a_l = Le, a_p = 0 \\ \\ \Lambda C_{Le,1,d}^{Le,2,d-1} + U(\mu', Le, 2, d-1) & a_l = Le, a_p = 1 \\ \\ q_\phi^f(\mu)\{\Lambda C_{Le,1,d}^{Le,2,d-1} + U(\mu', Le, 2, d-1)\} + \\ (1 - q_\phi^f(\mu))\{\Lambda C_{Le,1,d}^{Ri,0,-1} + U(\mu', Ri, 0, -1)\} & a_l = Ri, a_p = 0 \\ \\ \big(1 - g_l^1(\xi_p^{Ri})(1 - q_\phi^f(\mu))\big)\{\Lambda C_{Le,1,d}^{Le,2,d-1} + U(\mu', Le, 2, d-1)\} \\ + g_l^1(\xi_p^{Ri})(1 - q_\phi^f(\mu))\{\Lambda C_{Le,1,d}^{Ri,2,k-1} + W(\mu', Ri, 1, k-1)\} & a_l = Ri, a_p = 1 \end{cases} \tag{8}$$

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 term in the min function 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 term in the min function has the subject vehicle splitting from the platoon and changing lane. When the subject vehicle decides to change lanes while in a platoon, it has to split from its platoon first. The first expression here captures the scenario where the subject vehicle is not able to change lanes, under which 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 term in the min function has that the subject vehicle changing lanes and traveling on the right lane in a platoon. For the action to completely take place, the subject vehicle should dissolve 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 expression 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 expression 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., $l \neq l_d$, when $\phi_l = Le$, $\phi_p = 1$, $d = 0$), the minimum expected discounted cost is given by

$$V(\mu, Le, 1, 0) = \min \begin{cases} \Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1) & a_l = Le, a_p = 0 \\ \\ g_l^0(\xi_p^{Le})\{\Lambda C_{Le,1,0}^{Le,2,k-1} + W(\mu', Le, 1, k-1)\} + \\ (1 - g_l^0(\xi_p^{Le}))\{\Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1)\} & a_l = Le, a_p = 1 \\ \\ q_\phi^f(\mu)\{\Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1)\} + \\ (1 - q_\phi^f(\mu))\{\Lambda C_{Le,1,0}^{Ri,0,-1} + U(\mu', Ri, 0, -1)\} & a_l = Ri, a_p = 0 \\ \\ g_l^1(\xi_p^{Ri})(1 - q_\phi^f(\mu))\{\Lambda C_{Le,1,0}^{Ri,2,k-1} + W(\mu', Ri, 1, k-1)\} \\ +\big(1 - g_l^1(\xi_p^{Ri})(1 - q_\phi^f(\mu))\big)\{\Lambda C_{Le,1,0}^{Le,0,-1} + U(\mu', Le, 0, -1)\} & a_l = Ri, a_p = 1 \end{cases} \tag{9}$$

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 element of the min function 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 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 line). The third expression 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 line), 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 argument on the min function captures the scenario where the subject vehicle switches to the right lane and joins a platoon. The first line of this argument 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.

EXPERIMENTS AND ANALYSIS

In this section, we will conduct simulations in two experimental settings, namely a straight highway, and a small network with route choice. We compare the performance of the local OC model and the MDP framework under different traffic states in all three experimental settings. Our simulations are based on previously built simulation platform in (12), in which surrounding vehicles follow the Intelligent Driver Model (39). We consider aerodynamic resistance force, rolling resistance force, grade resistance force and inertia resistance force for fuel cost computation (40), and set value of time (VoT) to be 10 dollars per hour.

Parameters Calibration

In a future connected and automated vehicle system, parameters of the MDP model can be calibrated using historical data. According to the historical data, we partition ξ_{tr}^{Le}, ξ_{tr}^{Ri}, ξ_{p}^{Le} and ξ_{p}^{Ri} into different clusters, representing different traffic states in the left and right lanes, and different 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. We use piece-wise constants to fit the functions $q_{\phi}^{f}(\mu)$, $g_{l}^{1}(\xi_p)$ and $g_{l}^{0}(\xi_p)$, and the normal distribution to fit the function $w(k)$. For the current study, since historical data does not exist, we use simulations to set the values of these parameters.

Two-lane Highway

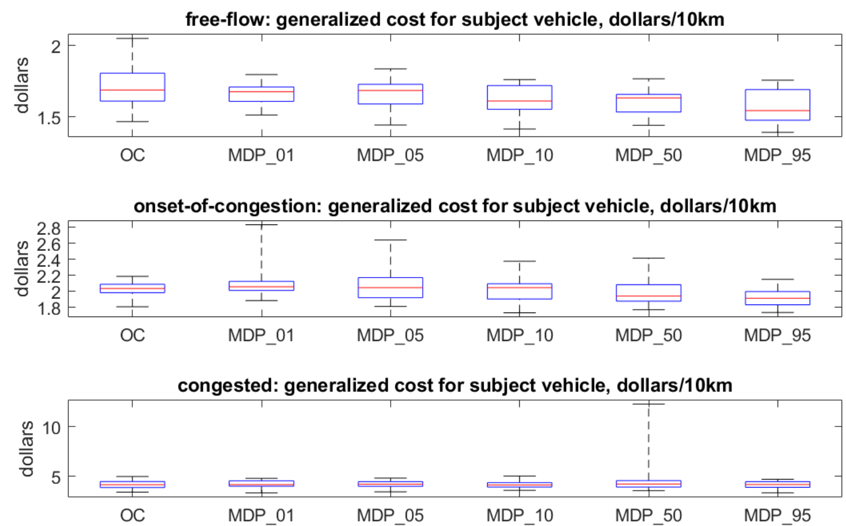

FIGURE 3 : The simulation environment is a straight highway with on- and off-ramps. The top, middle and bottom sub-figures represent the free-flow, onset-of-congestion, and congested traffic states, respectively. The vertical axes show the generalized costs with VoT set to 10 dollars per hour. Along the horizontal axes, the generalized costs of the subject vehicle under different controllers are compared. Here 'OC' and 'MDP' denote local optimal and the MDP controllers, respectively. The value following 'MDP_' in the name of the controller specifies the discount factor, α in the MDP model.

In this highway scenarios, we adopt the same surrounding environment setting as in (12). Surrounding vehicles can change lanes, merge/exit from the highway, and join into/split from a platoon. 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 (12), by utilizing a fundamental diagram of traffic flow. In figures presented in this paper, *OC* and *MDP* refer to the local optimal and MDP controllers, respectively. Figure 3 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

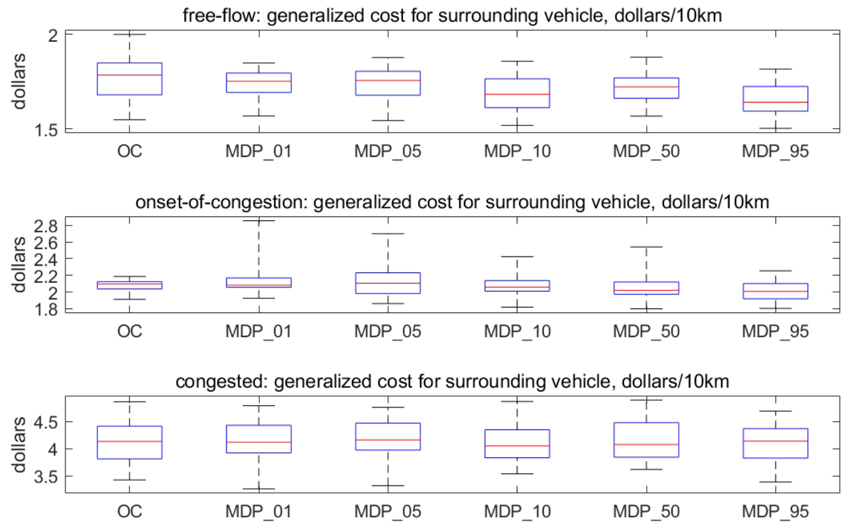

FIGURE 4 : The average generalized cost of the surrounding vehicles. The value following MDP in the name of the controller specifies the discount factor, α in the MDP model. Other settings are the same as Figure 3.

larger 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. Figure 4 shows the generalized cost of the surrounding vehicles. In the free-flow traffic state, the MDP controller results in significantly less 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 5 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 trajectories 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: 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 goes back to its original lane at about 3750 time steps, another 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.

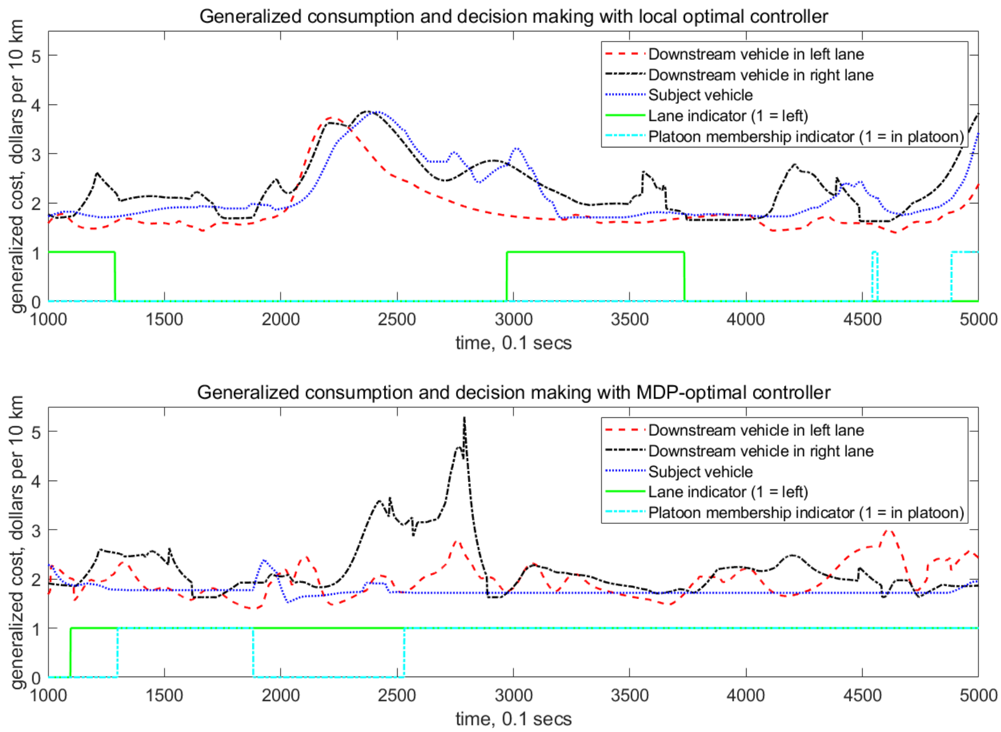

FIGURE 5 : The vertical axis shows the generalized cost, with the unit of dollars per 10 km. The horizontal axis is time, with the unit of 0.1 second. Generalized cost of the subject vehicle and its immediate downstream vehicles in both lanes, as well as its lane position and platoon membership status 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.

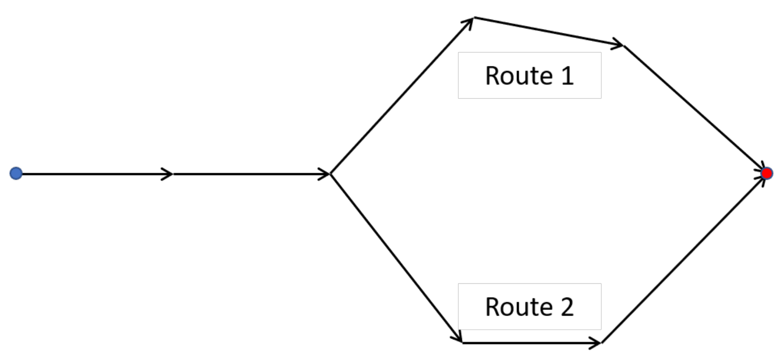

FIGURE 6 : The subject vehicle has two available routes from the origin (blue point) to the destination (red point). Route 1 has a slightly shorter distance, but it is usually more congested compared with route 2.

Routing

We show the extensibility of our method in a joint decision making framework, in which the model makes routing, lane-changing, and platoon-merging decisions. In a scenario as shown in

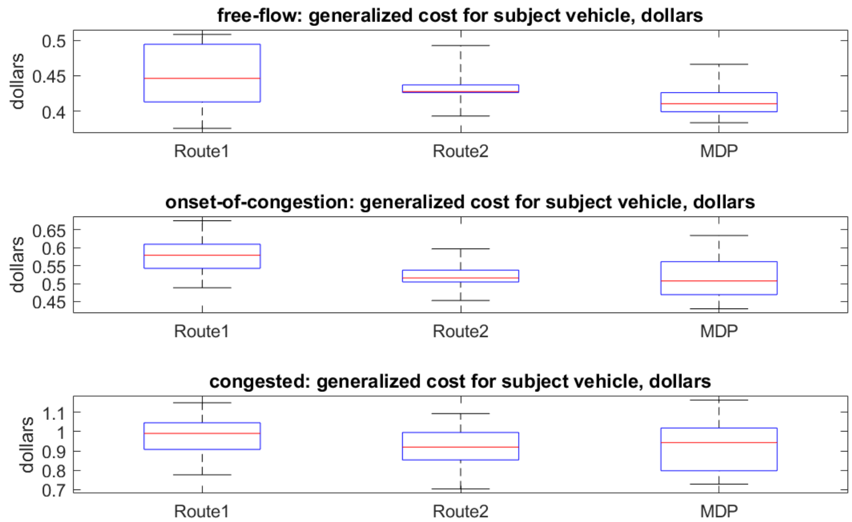

FIGURE 7 : Here 'Route1' and 'Route2' means that the subject vehicle is fixed to take Route1 and Route2 as shown in Figure 6, respectively. In these two modes, OC controller is applied. Other settings are the same as Figure 3.

Figure 6, the subject vehicle has two possible routes to the destination, namely 'Route1' and 'Route2'. Figure 7 shows the results of three decision-making 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 model 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.

CONCLUSION

In this paper we proposed an hierarchical motion planning framework for a CAV in a mixed traffic environment. The hierarchical 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 two experimental scenarios including a highway section with multiple on- and off-ramps, and a small network, and conducted a comprehensive set of simulations to quantify the long-term benefits the subject vehicle and its surrounding vehicles can experience as a result of incorporating network-level information into the decision-making process. Our experiments indicate that, generally speaking, the MDP framwork outperforms a local OC controller in reducing the generalized trip cost. With 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 originating from considering network-level information exists in all simulated environments, under various traffic states.

ACKNOWLEDGEMENT

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

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: study conception and design: X. Liu, N. Masoud, Q. Zhu, A. Khojandi. Analysis and interpretation of results: X. Liu, N. Masoud, Q. Zhu, A. Khojandi. Draft manuscript preparation: X. Liu, N. Masoud, Q. Zhu, A. Khojandi. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

[1] Kenney, J. B., Dedicated short-range communications (DSRC) standards in the United States. *Proceedings of the IEEE*, Vol. 99, No. 7, 2011, pp. 1162–1182.

[2] Shladover, S. E., C. Nowakowski, X. Y. Lu, and R. Ferlis, COOPERATIVE ADAPTIVE CRUISE CONTROL (CACC) DEFINITIONS AND OPERATING CONCEPTS. In *Trb Conference*, 2015.

[3] Wang, Z., G. Wu, and M. J. Barth, A Review on Cooperative Adaptive Cruise Control (CACC) Systems: Architectures, Controls, and Applications, 2018.

[4] Milanés, V. and S. E. Shladover, Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Research Part C: Emerging Technologies*, Vol. 48, 2014, pp. 285–300.

[5] Zhang, L., J. Sun, and G. Orosz, Hierarchical design of connected cruise control in the presence of information delays and uncertain vehicle dynamics. *IEEE Transactions on Control Systems Technology*, Vol. 26, No. 1, 2017, pp. 139–150.

[6] Orosz, G., Connected cruise control: modelling, delay effects, and nonlinear behaviour. *Vehicle System Dynamics*, Vol. 54, No. 8, 2016, pp. 1147–1176.

[7] Lioris, J., R. Pedarsani, F. Y. Tascikaraoglu, and P. Varaiya, Platoons of connected vehicles can double throughput in urban roads. *Transportation Research Part C: Emerging Technologies*, Vol. 77, 2017, pp. 292–305.

[8] Maiti, S., S. Winter, and L. Kulik, A conceptualization of vehicle platoons and platoon operations. *Transportation Research Part C Emerging Technologies*, Vol. 80, 2017, pp. 1–19.

[9] Huang, Z., D. Chu, C. Wu, and Y. He, Path planning and cooperative control for automated vehicle platoon using hybrid automata. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, No. 3, 2018, pp. 959–974.

[10] Bhoopalam, A. K., N. Agatz, and R. Zuidwijk, Planning of truck platoons: A literature review and directions for future research. *Transportation research part B: methodological*, Vol. 107, 2018, pp. 212–228.

[11] González, D., J. Pérez, V. Milanés, and F. Nashashibi, A review of motion planning techniques for automated vehicles. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 17, No. 4, 2015, pp. 1135–1145.

[12] Liu, X., G. Zhao, N. Masoud, and Q. Zhu, Trajectory Planning for Connected and Automated Vehicles: Cruising, Lane Changing, and Platooning. *arXiv e-prints*, 2020, p. arXiv:2001.08620.

[13] Zheng, Z., Recent developments and research needs in modeling lane changing.

Transportation research part B: methodological, Vol. 60, 2014, pp. 16–32.

[14] Larsson, E., G. Sennton, and J. Larson, The vehicle platooning problem: Computational complexity and heuristics. *Transportation Research Part C: Emerging Technologies*, Vol. 60, 2015, pp. 258–277.

[15] Cheng, J., J. Cheng, M. Zhou, F. Liu, S. Gao, and C. Liu, Routing in internet of vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 16, No. 5, 2015, pp. 2339–2352.

[16] Paden, B., M. Čáp, S. Z. Yong, D. Yershov, and E. Frazzoli, A survey of motion planning and control techniques for self-driving urban vehicles. *IEEE Transactions on intelligent vehicles*, Vol. 1, No. 1, 2016, pp. 33–55.

[17] Claussmann, L., M. Revilloud, D. Gruyer, and S. Glaser, A review of motion planning for highway autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[18] Alia, C., T. Gilles, T. Reine, and C. Ali, Local trajectory planning and tracking of autonomous vehicles, using clothoid tentacles method. In *2015 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2015, pp. 674–679.

[19] Zhou, J., H. Zheng, J. Wang, Y. Wang, B. Zhang, and Q. Shao, Multi-objective optimization of lane-changing strategy for intelligent vehicles in complex driving environments. *IEEE Transactions on Vehicular Technology*, 2019.

[20] Hardy, J. and M. Campbell, Contingency planning over probabilistic obstacle predictions for autonomous road vehicles. *IEEE Transactions on Robotics*, Vol. 29, No. 4, 2013, pp. 913–929.

[21] Bandyopadhyay, T., K. S. Won, E. Frazzoli, D. Hsu, W. S. Lee, and D. Rus, Intention-aware motion planning. In *Algorithmic foundations of robotics X*, Springer, 2013, pp. 475–491.

[22] Rakha, H. A., K. Ahn, and K. Moran, Integration framework for modeling eco-routing strategies: Logic and preliminary results. *International Journal of Transportation Science and Technology*, Vol. 1, No. 3, 2012, pp. 259–274.

[23] Ahn, K. and H. A. Rakha, Network-wide impacts of eco-routing strategies: a large-scale case study. *Transportation Research Part D: Transport and Environment*, Vol. 25, 2013, pp. 119–130.

[24] Boriboonsomsin, K., M. J. Barth, W. Zhu, and A. Vu, Eco-routing navigation system based on multisource historical and real-time traffic information. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 13, No. 4, 2012, pp. 1694–1704.

[25] Duret, A., M. Wang, and A. Ladino, A hierarchical approach for splitting truck platoons near network discontinuities. *Transportation Research Part B: Methodological*, 2019.

[26] Wei, Y., C. Avcı, J. Liu, B. Belezamo, N. Aydın, P. T. Li, and X. Zhou, Dynamic programming-based multi-vehicle longitudinal trajectory optimization with simplified car following models. *Transportation research part B: methodological*, Vol. 106, 2017, pp. 102–129.

[27] Qian, X., A. De La Fortelle, and F. Moutarde, A hierarchical model predictive control framework for on-road formation control of autonomous vehicles. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, 2016, pp. 376–381.

[28] Neunert, M., C. De Crousaz, F. Furrer, M. Kamel, F. Farshidian, R. Siegwart, and J. Buchli, Fast nonlinear model predictive control for unified trajectory optimization and tracking. In *2016 IEEE international conference on robotics and automation (ICRA)*, IEEE, 2016, pp.

1398–1404.
[29] Huang, K., X. Yang, Y. Lu, C. C. Mi, and P. Kondlapudi, Ecological driving system for connected/automated vehicles using a two-stage control hierarchy. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 19, No. 7, 2018, pp. 2373–2384.
[30] Guo, G. and Q. Wang, Fuel-efficient en route speed planning and tracking control of truck platoons. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 20, No. 8, 2018, pp. 3091–3103.
[31] Zeng, X. and J. Wang, Globally energy-optimal speed planning for road vehicles on a given route. *Transportation Research Part C: Emerging Technologies*, Vol. 93, 2018, pp. 148–160.
[32] Brechtel, S., T. Gindele, and R. Dillmann, Probabilistic MDP-behavior planning for cars. In *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2011, pp. 1537–1542.
[33] Bellman, R., A Markovian decision process. *Journal of mathematics and mechanics*, 1957, pp. 679–684.
[34] Sutton, R. S. and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
[35] Kamrani, M., A. R. Srinivasan, S. Chakraborty, and A. J. Khattak, Applying Markov decision process to understand driving decisions using basic safety messages data. *Transportation Research Part C: Emerging Technologies*, Vol. 115, 2020, p. 102642.
[36] Brechtel, S., T. Gindele, and R. Dillmann, Probabilistic decision-making under uncertainty for autonomous driving using continuous POMDPs. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, IEEE, 2014, pp. 392–399.
[37] van Wyk, F., A. Khojandi, and N. Masoud, Optimal switching policy between driving entities in semi-autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, Vol. 114, 2020, pp. 517–531.
[38] Zhou, S., Y. Wang, M. Zheng, and M. Tomizuka, A hierarchical planning and control framework for structured highway driving. *IFAC-PapersOnLine*, Vol. 50, No. 1, 2017, pp. 9101–9107.
[39] Jin, I. G. and G. Orosz, Optimal control of connected vehicle systems with communication delay and driver reaction time. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 18, No. 8, 2016, pp. 2056–2070.
[40] Gillespie, T. D., *Fundamentals of vehicle dynamics*. SAE Technical Paper, 1992.