Mixed Traffic Flow of Human Driven Vehicles and Automated Vehicles on Dynamic Transportation Networks

Qiangqiang Guo[a], Xuegang (Jeff) Ban[b], HM Abdul Aziz[c]

[a,b] Department of Civil and Environmental Engineering, University of Washington, Seattle 98195 WA, United States

[c] Civil Engineering Department, Kansas State University, Manhattan, KS 66506, United States

Abstract: Improving the system performance of a traffic network by dynamically controlling the routes of connected and automated vehicles (CAVs) is an appealing profit that CAVs can bring to our society. Considering that there may be a long way to achieve 100% CAV penetration, we discuss in this paper the mixed traffic flow of human driven vehicles (HDVs) and CAVs on a transportation network. We first propose a double queue (DQ) based mixed traffic flow model to describe the link dynamics as well as the flow transitions at junctions. Based on this mixed flow model, we develop a dynamic bi-level framework to capture the behavior and interaction of HDVs and CAVs. This results in an optimal control problem with equilibrium constraints (OCPEC), where HDVs' route choice behavior is modeled at the lower level by the instantaneous dynamic user equilibrium (IDUE) principle and the CAVs' route choice is modelled by the dynamic system optimal (DSO) principle at the upper level. We show how to discretize the OCPEC to a mathematical programming with equilibrium constraints (MPEC) and discuss its properties and solution techniques. The non-convex and non-smooth properties of the MPEC make it hard to be efficiently solved. To overcome this disadvantage, we develop a decomposition based heuristic model predictive control (HMPC) method by decomposing the original MPEC problem into two separate problems: one IDUE problem for HDVs and one DSO problem for CAVs. The experiment results show that, compared with the scenario that all vehicles are HDVs, the proposed methods can significantly improve the network performance under the mixed traffic flow of HDVs and CAVs.

Keywords: Human driven vehicles (HDVs); Connected and autonomous vehicles (CAVs); Mixed traffic flow; Dynamic user equilibrium (DUE); Instantaneous dynamic user equilibrium (DUE); Dynamic system optimal (DSO); dynamic bi-level model; Model predictive control (MPC)

1. INTRODUCTION

In the last decade or so, connected and automated vehicles (CAVs) have been one of the major disruptive mobility technologies in the transportation landscape. Recent developments on CAVs indicate that, despite tremendous advances that have been achieved on CAVs so far, it will take a relatively long time to reach full automation (i.e., level 5 as defined by SAE international (2016)) as well as a high market penetration of CAVs. Thus, it is expected that in the near future, we will see both human driven vehicles (HDVs) and CAVs on the roads. CAVs moving along with HDVs in the traffic stream will certainly bring opportunities to improve the traffic flow (e.g., a single CAV may be used to dampen traffic shockwaves; see Cui et al. (2017)). At the same time, it is important to develop methodologies to model the behavior and interactions of HDVs and CAVs, and to understand the overall network-wide effect of this mixed traffic flow. Human drivers are expected to minimize their own travel costs with little or no consideration to improve system level performance of the network. Their behavior is often assumed to follow the dynamic user equilibrium (DUE, see Peeta and Ziliaskopoulos (2001); Ran and Boyce (1996)) principle on a transportation network. For CAVs, they may be modeled in the same way as HDVs to follow DUE. However, thanks to their added communication and automation features, CAVs may be leveraged to implement certain strategies (such as routing and dispatching by transportation system mangers or ride-sourcing companies) to improve the

system performance of the entire network (e.g., the total system travel time and/or total fuel consumption). Such behavior can often be modeled by the dynamic system optimal (DSO) principle (Peeta and Ziliaskopoulos, 2001; Ran and Boyce, 1996). In this paper, we aim to explore such a system level control mechanism of CAVs to help improve the performance of a transportation network which consists of both CAVs and HDVs.

There are two key challenges to achieve this goal: First, how to properly model the impact of CAVs on the mixed traffic flow of HDVs and CAVs; and second, having the model of mixed traffic flow, how to model the behavior and interaction of HDVs and CAVs on the network level, based on which to improve the system performance by leveraging the automation and communication capabilities of CAVs. A detailed literature review of these two topics is presented in Section 2 . The first challenge has been studied recently in the literature, e.g., investigations on how traffic flow characteristics (e.g., capacity) are connected with the penetration of CAVs (Levin and Boyles, 2016; Liu et al., 2018). On the other hand, research on the second challenge has just started. Current studies have focused on static traffic assignment (Bagolee et el., 2016; Bahrami and Roorda, 2020), assumed 100% CAV penetration (Levin, 2017), or focused on local optimization, e.g., at traffic intersections (Levin and Boyles, 2016; Patel et al., 2016). Improving network level performance under mixed traffic flow by dynamically controlling the routes of CAVs remains an untapped area. This paper aims to fill this gap by designing a network level framework to model the dynamic interaction between HDVs and CAVs, based on which to develop a dynamic CAVs route controller under mixed traffic flow to improve the overall network performance.

We design a double queue (DQ) based mixed flow model to describe the dynamics of HDVs and CAVs. This mixed flow model explicitly considers the impact of CAVs to the characteristics of traffic flow (such as the flow capacity). The First-In-First-Out (FIFO) condition of the mixed-flow DQ model is also investigated. We then model the HDVs' behavior by the instantaneous dynamic user equilibrium (IDUE, see Ban et al. (2012a)) principle and the CAVs' behavior by DSO. Meanwhile, we model the interaction between HDVs and CAVs at the network level as a leader-follower game, in which CAVs are the leader and HDVs are the follower. Such a framework mathematically leads to a dynamic bi-level formulation. The whole problem is then formulated as an optimal control problem with equilibrium constraints (OCPEC), where CAVs' routes are the decision variable, system performance is the objective, mixed traffic flow dynamics are the constraints, and HDVs' route choice behavior (i.e., IDUE) is the lower-level problem. The OCPEC is hard to solve due to the non-convex and non-smooth characteristics introduced by the HDVs' behavior. We discretize the OCPEC to a mathematical program with equilibrium constraints (MPEC) that may be solved by certain relaxation methods. However, such a relaxation-based solution method cannot be applied to even small-size networks with long time horizons. We then propose a decomposition based heuristic model predictive control (HMPC) method to address this problem. HMPC decomposes the original OCPEC problem to two sub-problems that can be solved relatively easily.

The key contributions of this paper are that we:

1) Propose a DQ based mixed traffic flow model to describe the traffic dynamics of the mixed HDV/CAV traffic flow.
2) Develop a dynamic bi-level framework to capture the behavior and interactions of HDVs and CAVs at the network level, based on which to formulate an OCPEC to find the system performance oriented route for CAVs under the mixed traffic flow to improve the network performance.
3) Develop a decomposition based HMPC method to efficiently solve the proposed OCPEC.
4) Demonstrated the challenges of modeling mixed HDV/CAV flow on dynamic transportation networks, including how to model the capacity of the mixed flow, the delay terms in the model (i.e., the shockwave speed), and the FIFO conditions.

This paper is organized as follows. In Section 2, we provide a detailed review on the two challenges mentioned earlier. Section 3 presents the proposed methodology, starting with the discussion of a general modeling framework for DUE in Section 3.1 and the key assumptions in Section 3.2. In Section 3.3, we propose our mixed traffic dynamic model, including the basic DQ model and the FIFO condition (Section 3.3.1), the maximum sending and receiving flows (Section 3.3.2), and the nodal model (Section 3.3.3). We illustrate how to model the HDVs' behavior (Section 3.4) and the CAVs' behavior and formulate this problem into a dynamic bi-level problem (i.e., the OCPEC) in Section 3.5. We show how to discretize the OCPEC and discuss its properties in Section 4.1. We then introduce the HMPC method in Section 4.2. In Section 5, we show the numerical experiments on a small network with short time duration and discuss in detail how the proposed method can improve system performance (Section 5.1). We also test the proposed method on the small network with a longer time horizon, and on the Sioux Falls network (Section 5.2) to demonstrate the effectiveness of our methods. We conclude this paper in Section 6.

2. LITERATURE REVIEW

It is anticipated that the introduction of CAVs into transportation networks will impact the traffic flow dynamics, i.e., the parameters of traffic flow (e.g., capacity, jam density, free-flow speed, and shockwave propagation speed, etc.) and the deployment of CAVs (e.g., the market penetration and level of automation) are potentially correlated. Existing studies have shown positive impacts associated with the introduction of CAVs into the traffic flow. Such impacts have been found nonlinear from most of the studies. Most studies are either analytical or simulation based since CAVs have not been widely deployed in the real-world implementation. In this section, we discuss the literature on the impact of CAVs on the mixed traffic flow and the network modeling techniques to assess the system impact of CAVs on a transportation network.

2.1. Modeling the Impact of CAVs on the Mixed Traffic Flow of HDVs and CAVs

The study on the impact of CAVs on traffic flow initially rooted from the adaptive cruise control (ACC) and cooperative adaptive cruise control (CACC) modeling. The feature of desired speed and spacing in CACC and ACC motivated researchers to model CAVs in a similar manner. The majority of the research (Liu et al., 2018; Li et al., 2020; Zhou et al., 2020; Zheng et al., 2018, 2020) modeled CAVs at the microscopic level, including the intelligent driver model (IDM, Treiber et al., 2000), ACC (Milanés and Shladover, 2014), and so on. These studies showed that one of the most important impact of CAVs on the mixed traffic flow is that the capacity can be increased due to the introduction of CAVs. The capacity improvements are highly dependent on the market penetration of CAVs while the trend is nonlinear, i.e., moving from 10% to 30% CAV market share may not have the same impact as the transition from 60% to 80% market share (Li et al. 2020; Liu et al., 2018; Ye and Yamamoto, 2018). Ye and Yamamoto (2018) developed a two-state safe-speed model and a two-lane cellular automaton model to describe the mixed flow of HDVs and CAVs, and numerically investigated the impact of CAVs on the road capacity under different market penetration. The simulation results on a two-lane road showed that the road capacity increased almost linearly up to a CAVs penetration rate of 30%. When the penetration rate exceeded 30%, the capacity would still increase, but the growth rate was significantly affected by the CAVs' capability (i.e., the desired net time gap of ACC). Liu et al. (2018) explored the impact of CAVs equipped with CACC on multi-lane freeway traffic. They used a six-mode driver model, which combined the Newell (2002) model, Gipps (1981) model and IDM, to mimic the behavior of HDVs, and designed a rule-based CACC controller to model the behavior of CAVs. They also integrated lane-changing rules to make the CACC's behavior more realistic at on/off ramp areas. The simulation on a 4-lane real-world freeway showed that, for both on-ramp and off-ramp areas, the bottleneck capacity increased quadratically as the CAV penetration increases from 0 to 100%. Li et al. (2020) investigated the impact of CAVs on the capacity of a two-lane road under four different Right-of-Way (RoW) reallocation strategies (i.e., two mixed lane; one mixed lane and one CAV dedicated lane; one mixed lane and one HDV dedicated lane; one CAV dedicated lane and one HDV dedicated lane). They used the Krauss model (Krauss, 1998) to

describe the HDVs' behavior and used the ACC system to model the bahavior of CAVs. The simulation results showed that the road capacity can be significantly improved with appropriate RoW reallocation strategies at low or medium CAV penetration rates, compared with the do-nothing RoW strategy (i.e., two mixed-flow lanes). These studies revealed that, high level penetration of CAVs, high level and high capability automation, and exogenous technological assistance play important roles in improving the capacity of the mixed traffic flow of HDVs and CAVs. Apart from the impact on road capacity, CAVs may also influence the stability of the mixed traffic flow; see Wang et al. (2013), Zheng et al. (2020), Zhou et al (2020) for further discussions.

While most mixed traffic flow models were developed in simulation, only a few studies applied link-level methods to study the impact of CAVs on traffic flow dynamics. Levin and Boyles (2016) proposed a multi-class cell transmission model (CTM, Daganzo, 1992) for the mixed traffic of HDVs and CAVs. They derived the model by investigating the microscopic behavior of traffic flow (i.e., a collision avoidance car following model) assuming different vehicle reaction times of HDVs and CAVs, which leads to an explicit mixed HDV/CAV flow model. The model revealed that the capacity of a link could be formulated as a reciprocal function of the penetration of CAVs. Similar capacity model was also proposed by Lazar et al. (2017).

In summary, as reviewed in this section, although there are some commonly agreed effects of introducing CAVs to mixed traffic flow on traffic dynamics, many remains unclear, especially there is not a commonly agreed modeling framework to capture such effects in traffic dynamics. In this paper, we consider the impact of CAVs on the link capacity from the macroscopic level. Based on the few current studies on link level mixed traffic flow modeling (Bahrami and Roorda, 2020; Lazar et al., 2017; Levin, 2017; Levin and Boyles, 2016; Patel et al., 2016), we adopt the mixed flow model in Levin and Boyles (2016) to model the flow capacity of a link as a reciprocal function of the CAVs' penetration; see Section 3.2 for more details.

2.2. Improving Network Performance under the Mixed Traffic Flow of HDVs and CAVs

Having the model of mixed traffic flow, how to improve the system performance by leveraging the automation and communication features of CAVs remains another challenge. For examples, Bagolee et el. (2016) tested mixed CAVs and HDVs using static traffic assignment. They generated different CAV penetration rates from a fixed number of vehicle fleet, then modeled HDVs to follow static user equilibrium and CAVs to follow static system optimal. Next, they integrated these two optimization problems as a mixed user equilibrium and system optimal problem by directly combining the objectives and constraints, which led to a nonlinear program (NLP). Experiments on a five-node network showed that as the CAV penetration increased, the total travel time of the whole network decreased. However, in their study, no interaction between CAVs and HDVs was modeled. Bahrami and Roorda (2020) evaluated different policies (i.e., exclusive CAV and HDV links) with the UE principle. They modeled the capacity of a link as a quadratic function of the CAVs penetration and designed a non-linear complementarity problem (NCP) to formulate the equilibrium conditions. Experiment results showed that simple policies such as CAV exclusive links can improve network performance under the mixed traffic flow. However, only static UE condition was applied in this study without considering traffic dynamics.

Based on the multi-class CTM, Levin and Boyles (2016) designed an intelligent intersection control algorithm under the mixed traffic flow. They modified the conflict regions model (Levin and Boyles, 2015) for HDVs and used a tile-based reservation model (Dresner and Stone, 2004) for CAVs. Combing the multi-class CTM and the intersection model, they studied the mixed traffic flow in a dynamic traffic assignment (DTA) framework. Numerical results showed that, under high CAVs penetration, the proposed intersection management method could improve the system performance compared with traditional traffic light control. Using the same multi-class CTM and the same reservation-based intersection control method in Levin and Boyles (2016), Patel et al. (2016) analyzed the effects of CAVs

on congested arterial and freeway networks. The experiment results showed that the reduced headway of CAVs could significantly improve the capacity of freeway and arterial networks. Levin (2017) applied the link transmission model (LTM, Yperman, 2007) to model the shared autonomous vehicle (SAV) routing problem while considering congestion. They assumed that SAVs follow the system optimal principle to build a SAV DTA model. Experiment results showed that this method could reduce SAV travel times and determine the optimal fleet size to minimize congestion or maximize service. However, they assumed all personal vehicle trips were replaced by SAV trips, i.e., the SAV penetration is 100%.

It can be summarized that several issues exist for current studies on developing network models to capture the behavior and interactions of HDVs and CAVs to improve the network performance under the mixed traffic flow. First, many of them focused on static traffic assignment (Bagolee et el., 2016; Bahrami and Roorda, 2020), which only reflected the steady pattern of the mixed traffic flow and did not consider the dynamic interaction between HDVs and CAVs. Second, studies that did consider dynamic traffic flow, however, usually either assumed 100% CAV penetration (Levin, 2017) or focused on local optimization (e.g., intelligent intersection control under mixed traffic flow, as studied in Levin and Boyles (2016) and Patel et al. (2016)), without controlling the network-level behavior (such as routing) of CAVs.

3. METHODOLOGY

In this section, we first present a general modeling framework, called the differential complementarity system (DCS), to model DUE of mixed traffic flow on dynamic transportation networks (Section 3.1). This also identifies a key challenge when modeling mixed traffic flow at the network level. Next, we discuss the key assumptions of our developed model (Section 3.2). Sections 3.3 and 3.4 introduce the mixed traffic flow model, and mathematically formulate the behaviors of the HDVs. Finally, we present the behavior of CAVs and propose a dynamic bi-level optimal control problem to find the routes for CAVs to improve the system performance (Section 3.5). **Appendix A** lists the notation used hereafter in this paper.

3.1. DCS for Modeling DUE of Mixed HDV/CAV Flow on Dynamic Transportation Networks

This subsection presents the DCS framework to model DUE, following the work in Ban et al. (2012a) and Ma et al. (2018). This also helps to illustrate the challenge of balancing the mathematical rigor and capturing realm when modeling mixed HDV/CAV flow on dynamic transportation networks. To start, we show below the DCS-based DUE model for a single vehicle class (e.g., HDVs) and a single destination:

$$
\begin{aligned}
&0 \le p_{ij}(t) \perp \tau_{ij}(t) + \eta_j(t) + \eta_i(t) \ge 0 && (1-1)\\
&0 \le \eta_i(t) \perp \sum_{j:j\in K_i} p_{ij}(t) - \sum_{j:j\in I_i} v_{ji}(t) - d_i(t) \ge 0 && (1-2)\\
&\dot{q}^U_{i,j}(t) = p_{i,j}(t) - v_{i,j}\left(t - \tau^w_{i,j}\right) && (1-3)\\
&\dot{q}^D_{i,j}(t) = p_{i,j}\left(t - \tau^0_{i,j}\right) - v_{i,j}(t) && (1-4)\\
&\tau_{ij}(t) = \tau^0_{ij} + \frac{q^D_{ij}(t)}{\bar{C}_{ij}(t)} && (1-5)
\end{aligned}
\qquad (1)
$$

Here $\perp$ indicates "perpendicular", i.e., $x \perp y \Leftrightarrow x^T y = 0$ for two vectors x and y. Equation (1-1) describes the instantaneous DUE (IDUE) route choice condition in which the inflow to link (i,j) at time t, $p_{ij}(t)$, will be nonnegative (i.e., the flow may choose link (i,j)) if link (i,j) is on the minimum travel time path from i to the destination (i.e., $\tau_{ij}(t) + \eta_j(t) - \eta_i(t) = 0$), with $\tau_{ij}(t)$ denoting the link travel time and $\eta_i(t)$ the minimum travel time from node i to the destination. Otherwise, $p_{ij}(t) = 0$, i.e., flow will not choose link (i,j) if the link is not on the minimum travel time path (i.e., $\tau_{ij}(t) + \eta_j(t) - \eta_i(t) > 0$).

Equation (1-2) indicates flow conservation with $v_{ji}(t)$ the exit flow rate from link (j,i) at time t, $d_i(t)$ the demand generated at node i at time t, and I_i and K_i respectively the list of incoming nodes (with links directly to node i) and the list of outgoing nodes (i.e., with links directly from node i) of node i. Since $\eta_i(t) > 0$ always holds, the right-hand side of (1-2) always hold as an equation, representing the flow conservation at node i. Equations (1-3) and (1-4) define the double queue model with $q_{ij}^U(t)$ and $q_{ij}^D(t)$ denoting the upstream queue and downstream queue, respectively (see Section 3.3.1 and also Ma et al. (2014)). In these two equations, $\tau_{i,j}^0 = l_{ij}/v_{ij}^f$ and $\tau_{i,j}^w = l_{ij}/w_{ij}^w$ denote respectively the free flow travel time and the shockwave travel time of link (i,j). $l_{ij}, v_{ij}^f, w_{ij}^w$ denote respectively the length, free flow speed, and shockwave speed of the link. $\bar{C}_{ij}(t)$ denotes the capacity of link (i,j).

In the above DUE model, we apply IDUE in the DUE route choice and the double queue model for traffic flow, which will be extended later in this paper to model the DUE with mixed traffic flow of HDVs and CAVs. Readers can refer to Section 3.3.1 and Section 3.4 for more details. The DUE model presented above should also include certain nodal model to describe how flow is transferred at a junction node (see e.g., the nodal model in Ma et al. (2018)). We defer the presentation of such a nodal model to Subsection 3.3.3 since it is irrelevant to the modeling challenge that we will discuss next.

Equations (1-3) and (1-4) contain time delay terms, $\tau_{i,j}^0$ and $\tau_{i,j}^w$. These two terms are constant (for a given link) for the above DUE model, which applies to either 100% HDVs or 100% CAVs. As shown in the literature (Ban et al., 2012a; Ma et al., 2018), DCS with constant time delays can be mathematically analyzed for solution existence and convergence and can be readily solved by discretization and time-stepping methods; more details on this can be seen in Ban et al. (2012a). However, as we will show in the next subsection, for mixed flow of HDVs and CAVs, the shockwave speed may change, leading to varied (and state-dependent) shockwave travel time. This will result in a DCS model with time-varying, state dependent delays that are of the type of functional differential equations (rather than ordinary differential equations as shown in Equations (1-3) and (1-4) above) and much harder to deal with mathematically for both analysis and solution. This is a key challenge when modeling the DUE problem of mixed HDVs/CAVs on dynamic transportation networks. It is about how to balance the mathematical rigor of the developed model (e.g., constant or time-varying delay terms) and the model's ability to accurately capture the underlying physical property/process of the problem (e.g., constant or varying shockwave speed). Nevertheless, identifying and understanding such challenge can help provide useful insights on model development. It is often desirable to develop a model that can best balance these two aspects (i.e., mathematical rigor and realm), which often requires certain approximation schemes or creative modeling techniques under specific situations. We discuss more about this aspect when presenting the third assumption in the next subsection.

We discuss two remarks to close this subsection. First, although we use the DCS framework here to illustrate DUE, delay terms in DUE exist in other modeling frameworks and are one of the key challenges of modeling dynamic transportation networks (Friesz et al., 2001; Friesz and Mookherjee, 2006; Friesz et al., 2010; Ma et al., 2015; Ma et al., 2018). Therefore, the above identified challenge is general to DUE modeling. Second, if one has to apply time-varying, state-dependent delay terms in DUE models (e.g., the free flow or shockwave travel times have to be considered as state-dependent), the proposed modeling framework in this paper may still apply (Ma et al., 2018), which also requires certain approximation techniques (Friesz and Mookherjee, 2006; Friesz et al., 2010; Ma et al., 2015) to generate and solve sub-problems that are similar to (1) above. That is, the mixed HDV/CAV network model proposed in this paper, by assuming fixed shockwave speed, can be considered as the key sub-problem of the more general mixed flow network model. We omit detailed discussions on this general model in this paper to save space, which is summarized as one of the future research directions in Section 6.

3.2. Assumptions

We have four major assumptions: 1) Both HDVs and CAVs can access all information of the traffic network to make their route choices; 2) HDVs select their routes based on the IDUE principle and CAVs can be fully controlled and deployed to help improve the network performance; 3) The capacity of a link under mixed traffic flows is modeled following the work in Levin and Boyles (2016), i.e., the capacity is a dynamic reciprocal function of the penetration of CAVs; and 4) the shockwave speed of traffic on a link is assumed to be constant. Next, we will discuss these assumptions in detail.

First, we assume that both HDVs and CAVs can access all information of the traffic network, e.g., the number of vehicles and queue length of each link. In reality, CAVs may differ from HDVs in many aspects such as the information perception, decision making, and driving operations. For examples, CAVs collect the information of the surrounding environment by onboard sensors and V2V techniques, while HDV drivers collect such information through human sensing; the trajectories and motions of CAVs can also be controlled by automation algorithms, while HDVs are controlled by humans; CAVs tend to travel in vehicular platoons while HDVs are more independently controlled and well mixed. Since this paper focuses on network level traffic flow control (i.e., the routing of CAVs), the differences between CAVs and HDVs in information perception and decision making are neglected. Instead, we model traffic as "flow" and do not capture individual vehicles. Therefore, the "micro-level" vehicle dynamics (e.g., vehicle trajectory and motion control) will not be modeled.

Second, based on the first assumption, HDVs select their routes to minimize their own travel costs (i.e., they are "selfish"). In this paper, we apply DUE to model HDVs' route choice behavior. There are essentially two modeling approaches for DUE: predictive DUE (PDUE) and instantaneous DUE (IDUE). PDUE (Ma et al., 2018; Ran and Boyce, 1996) assumes that travelers select their routes based on the predicted (future) traffic conditions at the beginning of their trips and stick to these routes during the entire trips. IDUE (Ban et al., 2012a; Ran and Boyce, 1996) assumes that travelers make route choice decisions based on the current prevailing traffic conditions and thus can change their routes during their trips. In this paper, we use IDUE to model the route choice behavior of HDVs due to two reasons. First, PDUE requires that one could accurately predict the future traffic conditions, which is difficult in practice. On the contrast, IDUE only needs the current traffic information, which is more practical through various means of intelligent transportation systems (ITS). Second, compared with PDUE that does not revise routes along a trip, IDUE allows en-route revisions of routes, which can better capture the dynamic traffic conditions and reflect the interactions between HDVs and CAVs. IDUE is reactive, i.e., the flow assigned to a path at a time instant does not impact the cost (travel time) of the path at that time instant but will impact the path travel times at later time instants. This reactive nature of IDUE also have drawbacks, noticeably that flow is often assigned to the single minimum-cost path and in extreme situations flow may go back to the origin and select a different route (see some discussions on this in Ban et al. (2012a)).

On the other hand, it is reasonable to assume that CAVs can be fully controlled and deployed to help resolve traffic congestion and related problems. The routes of CAVs are thus assumed to be controllable and we assume CAVs (are controlled to) choose their routes to improve the system performance. Thus, in this paper, CAVs' behavior is modeled by the DSO principle (Ma et al., 2014; Merchant and Nemhauser, 1978; Shen and Zhang, 2018; Ziliaskopoulos, 2000) to improve (minimize) the system objective (e.g., the total system travel time spent by all travelers). Note that here we just aim to explore the possibility of leveraging the communication and automation capability of CAVs to improve the transportation network performance. To actually make this work, there can be other related issues that need to be carefully investigated, e.g., how to make sure the routing strategies generated by the proposed model are "fair" to the passengers of CAVs. We leave this for future research as summarized in Section 6.

Third, similar to existing studies on link level mixed traffic flow model (Lazar et al., 2017; Levin and Boyles, 2016; Levin, 2017; Patel et al., 2016), we adopt the model in Levin and Boyles (2016) and model the flow capacity of a link as a dynamic reciprocal function of the penetration of CAVs. Such a model is

derived from the microscopic car-following model by assuming that CAVs can reduce the reaction time and time headway between vehicles. Specificlly, let $n_{i,j}^{hdv}(t)$ and $n_{i,j}^{cav}(t)$ be the number of of HDVs and CAVs respectively on link (i,j) at time t, which yields the relative penetration of CAVs as $\rho_{i,j}(t) = \frac{n_{i,j}^{cav}(t)}{n_{i,j}^{cav}(t)+n_{i,j}^{hdv}(t)}$. Let $\bar{C}_{i,j}^{hdv}$ and $\bar{C}_{i,j}^{cav}$ be the link flow capacities under 100% HDVs and CAVs respectively ($\bar{C}_{i,j}^{cav} > \bar{C}_{i,j}^{hdv}$). The capacity of link (i,j) at time t can be calculated by (Levin and Boyles, 2016):

$$\bar{C}_{i,j}(t) = \frac{\bar{C}_{i,j}^{hdv}\bar{C}_{i,j}^{cav}}{\bar{C}_{i,j}^{cav} + \rho_{i,j}(t)(\bar{C}_{i,j}^{hdv} - \bar{C}_{i,j}^{cav})} \tag{2}$$

Notice that Equation (2) was not explicitly given in Levin and Boyles (2016), which however can be straightforwardly derived from their model, e.g., Equation (21) in Levin and Boyles (2016). Equation (2) above shows that, as the CAV penetration increases, the capacity $\bar{C}_{i,j}(t)$ of link (i,j) increases from $\bar{C}_{i,j}^{hdv}$ to $\bar{C}_{i,j}^{cav}$. We show this capacity function under different CAV saturation headways (h_t) and penetration rates (ρ) in Figure 1. For illustration purposes, the time headway of HDVs is assumed to be 1.8s and the free flow speed is assumed to be 50 mph, which indicate that the link capacity when all vehicles are HDVs is $\bar{C}_{i,j}^{hdv} = 2000$ vph. Figure 1 shows that, if CAV penetration is fixed, as the CAVs' time headway decreases, the link capacity increases due to the reduction of time headway. Such increment is also nonlinear. For a fixed time headway, as the CAV penetration increases, the capacity also increases.

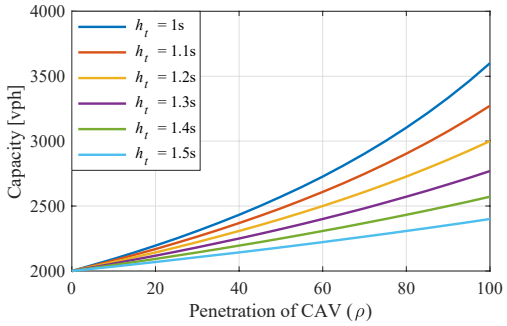


Figure 1. Capacity function under different CAV headways using Equation (2)

We need to point out that existing studies in the literature found (mostly using simulation data) that the capacity of mixed HDV/CAV flow increases quadratically with the penetration of CAVs (Liu et al., 2018). Equation (2) here actually is consistent with this finding. To see this, we approximate the right-hand side of (1) using its Taylor series expansions (around zero) up to the second order.

$$\bar{C}_{i,j}(t) = \bar{C}_{i,j}^{hdv}\left[1 + \rho_{i,j}(t)\left(1 - \frac{\bar{C}_{i,j}^{hdv}}{\bar{C}_{i,j}^{cav}}\right) + \rho_{i,j}^2(t)\left(1 - \frac{\bar{C}_{i,j}^{hdv}}{\bar{C}_{i,j}^{cav}}\right)^2 + O[\rho_{i,j}^3(t)]\right] \tag{3}$$

Equation (3) indicates that a quadratic relationship approximately holds between CAV penetration and the capacity of the mixed HDV/CAV flow.

Fourth, as shown in Section 3.1 above, it is also imporant to model the impact of CAVs on the shockwave speed. In most existing studies, the introduction of CAVs to the traffic flow (of HDVs) is assumed to change the shockwave speed (that depends on the actual penetration of CAVs), which however does not change the jam density (Levin and Boyles, 2016; Patel et al., 2016). Note here that these are largely assumptions with limited or no field observations. It is our understanding that the introduction of CAVs may change

both the shockwave speed and the jam density (i.e. under CAVs the spacing between vehicles may be reduced at the jammed state due to vehicle connectivity and automation). To see the latter impact, Figure 2 shows the fundamental diagrams of a link under different CAV penetrations assuming fixed shockwave speed (and varying jam density) of a link. Note that each fundamental diagram is corresponding to a set of fixed capacity (i.e., fixed CAV penetration), free-flow speed and shockwave speed. Here the shockwave speed is set as 15 mph and the CAV time headway is 1.4s. Then as CAV penetration increases from 0 to 100%, the jam density increases from 174 veh/mi to 223 veh/mi. Assume the average effective vehicle length in a typical traffic flow is 20 feet, the constant shockwave speed indicates that the average gap between consecutive vehicles are reduced from 10.34 feet to 3.67 feet. Such reduction may be achieved by CAV techniques considering that traffic is at standstill when jam density is considered, which implies that the constant shockwave speed may hold.

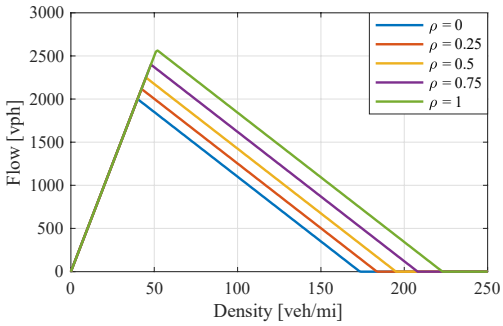


Figure 2. Fundamental diagram of a link under different CAV penetrations

We notice here that assuming fixed jam density or fixed shockwave speed just represents one extreme (and simplified) case of the general situation. As discussed in Section 3.1 above, the varying shockwave speeds will lead to a DCS with time-varying, state-dependent delays that are much harder to analyze and solve. For the sake of clearly presenting how to model HDVs/CAVs at the network level, in this paper, we assume fixed shockwave speed and varying jam density (for a given link) with respect to different CAV penetration to simplify the model analysis and solution process. This also presents an alterantive (and simplified) way to model the impact of CAVs on jam density and shockwave speed. The fixed shockwave speed assumption can be relaxed by using the approximations schemes proposed in Ma et al. (2015, 2018); see the discussions in Section 3.1 and also in Section 6.

Also noteworthy is that the current findings about mixed HDV/CAV flow are either based on simulation data (e.g. how capacity increases with CAV penetration) or largely assumptions (e.g., fixed jam density or shockwave speed) since CAVs are not widely deployed (especially high level CAVs) and field observations are quite limited. One should certainly keep a close eye on field testing/deployment of CAVs to properly update the properties of mixed HDV/CAV flow and the resulting DUE models. We present more discussions on this in the Section Section 6.

3.3. Mixed Traffic Flow Model of HDVs and CAVs

We use the LTM (Yperman, 2007) in this paper as the basis to formulate link level traffic dynamics. LTM is similar to CTM, but only requires temporal discretization (and without spatial discretization) of a link, which can help reduce the dimension of the problem. In particular, we apply the equivalent "double queue" (DQ) model first proposed in Osorio et al. (2011) to simplify the link-flow dynamics. DQ describes the dynamics of a link with two queues: an upstream queue and a downstream queue. The downstream queue is similar to the point queue model, which can model the free-flow time delay and the queuing process and exit flow (from the link) at the downstream of a link. The upstream queue can

capture the backward shockwave delay and the congestion propagation process and the inflow (to the link) by a pseudo queue at the upstream of a link. In this paper, we use the continuous-time DQ model developed in Ma et al. (2014) as the underlying traffic flow model.

3.3.1. The basic double queue model

The DQ of a link (i,j) can be formulated as the following ODEs (Ma et al., 2014):

$$\begin{aligned} \dot{q}_{i,j}^{U}(t) &= p_{i,j}(t) - v_{i,j}\left(t - \tau_{i,j}^{w}\right) \\ \dot{q}_{i,j}^{D}(t) &= p_{i,j}\left(t - \tau_{i,j}^{0}\right) - v_{i,j}(t) \end{aligned} \tag{4}$$

where $q_{i,j}^{U}(t)$ is the upstream queue, which is upper-bounded by the queue storage capacity of the link, i.e., $0 \le q_{i,j}^{U}(t) \le \bar{Q}_{i,j}(t)$. $q_{i,j}^{D}(t)$ is the downstream queue, $p_{i,j}(t)$ is the inflow rate of link (i,j), $v_{i,j}(t)$ is the exit flow rate of link (i,j), $\tau_{i,j}^{0}$ is the free flow travel time, and $\tau_{i,j}^{w}$ is the shockwave propagation travel time.

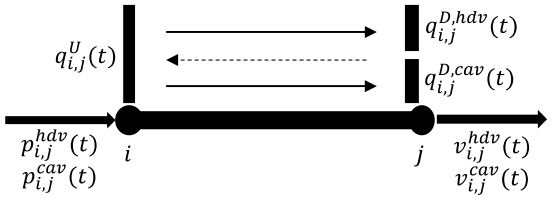


Figure 3. The double queue model

As shown in Figure 3, the basic DQ model (4) can be extended to capture the dynamics of HDVs and CAVs at the link level:

$$\begin{aligned} \dot{q}_{i,j}^{D,hdv}(t) &= p_{i,j}^{hdv}\left(t - \tau_{i,j}^{0}\right) - v_{i,j}^{hdv}(t) \\ \dot{q}_{i,j}^{D,cav}(t) &= p_{i,j}^{cav}\left(t - \tau_{i,j}^{0}\right) - v_{i,j}^{cav}(t) \end{aligned} \tag{5}$$

We do not need to separate the upstream queue to HDVs and CAVs since the upstream queue is used to calculate the total receiving capacity of both classes of vehicles, as indicated in Equation (13) later. Hereafter in this paper, we use asterisk $(*)$ to represent either HDVs or CAVs. So, we have $p_{i,j}(t) = \sum_{*} p_{i,j}^{*}(t)$ and $v_{i,j}(t) = \sum_{*} v_{i,j}^{*}(t)$. The dynamics of the total number of vehicles of a link (i,j) is

$$\dot{n}_{i,j}^{*}(t) = p_{i,j}^{*}(t) - v_{i,j}^{*}(t) \tag{6}$$

The instantaneous link travel time is defined as the sum of the free-flow travel time and the instantaneous exit time:

$$\tau_{i,j}(t) = \tau_{i,j}^{0} + \frac{q_{i,j}^{D}(t)}{\bar{C}_{i,j}(t)} \tag{7}$$

Here $\bar{C}_{i,j}(t)$ is the instantaneous capacity defined in Equation (2).

FIFO is an important requirement for modeling traffic flow on dynamic transportation networks. For the DQ model presented here, we have the following proposition regarding FIFO:

Proposition 1: FIFO holds for link (i,j) if the following condition is satisfied for all t:

$$\left(\dot{\rho}_{i,j}(t) q_{i,j}^{D}(t) + \rho_{i,j}(t) \dot{q}_{i,j}^{D}(t)\right)\left(\frac{1}{\bar{C}_{i,j}^{cav}} - \frac{1}{\bar{C}_{i,j}^{hdv}}\right) + \frac{\dot{q}_{i,j}^{D}(t)}{\bar{C}_{i,j}^{hdv}} + 1 > 0 \tag{8}$$

The proof of Proposition 1 can be done straightforwardly by taking the derivatives of both sides of (7) with respect to time t and requiring $\dot{\tau}_{i,j}(t) > -1$. The theorem shows that FIFO holds if the downstream queue, the CAV penetration, and the derivatives of these two follow Equation (8).

We can show that FIFO holds under certain situations based on Proposition 1. For example, when the length of the downstream queue does not change much, i.e., $\dot{q}_{i,j}^D(t) \sim 0$, FIFO will hold if $\dot{\rho}_{i,j}(t) < 1/\left[q_{i,j}^D(t)\left(\frac{1}{\bar{C}_{i,j}^{hdv}} - \frac{1}{\bar{C}_{i,j}^{cav}}\right)\right]$, i.e., when the increase of the CAV penetration on the link is relatively small. On the other hand, if the CAV penetration remains approximately constant (e.g. 100% HDVs, 100% CAVs, or relatively constant CAV penetration on the link), $\dot{\rho}_{i,j}(t) \sim 0$, (8) implies $\dot{q}_{i,j}^D(t) < 1/\left[\frac{\rho_{i,j}(t)}{\bar{C}_{i,j}^{cav}} + \frac{1-\rho_{i,j}(t)}{\bar{C}_{i,j}^{hdv}}\right]$, i.e., the downstream queue of the link does not change too dramatically. This is similar to the FIFO conditions of existing DUE models (for only HDVs). In the numerical experiments in Section 5, we also show that FIFO holds for the testing cases in this paper.

The condition in Equation (8) however may not be satisfied readily, implying that FIFO may be violated for the DQ model presented here. Since FIFO is a basic assumption of DTA so that dynamics such as mass balance constraints and flow propagation can be properly modeled mathematically (as ODEs or PDEs), FIFO violation will make the formulations of those dynamics less accurate (or only hold approximately). As a result, errors or unpractical dynamic behaviors may be introduced if FIFO is violated. Therefore, future research should strive to resolve this FIFO issue; see Section 6. We also note here that FIFO can normally be guaranteed approximately even one models single-class vehicles (Blumberg and Bar-Gera, 2009), left alone modeling multi-class vehicles as we do here. This further indicates that to model the mixed HDV/CAV flow, additional challenges (e.g. FIFO) may be introduced, which need to be addressed properly.

The above DQ model is only at the link level. To extend it to the network level, we need to model the flow interactions and transfers among neighboring links. We next present the maximum sending/receiving flows of a link (Section 3.3.2) and the nodal model (Section 3.3.3) to describe how flow is transferred from a link to its downstream links.

3.3.2. Maximum sending flow and maximum receiving flow

The maximum sending flow of a link (i,j), denoted as $S_{i,j}(t)$, is the maximum flow that can exit from the downstream queue of the link without considering the congestion of the downstream links. In LTM (Yperman, 2007), the sending flow is:

$$S_{i,j}(t)\Delta t = \min\left(\left(N\left(x_{i,j}^{in}, t + \Delta t - \tau_{i,j}^0\right) - N\left(x_{i,j}^{out}, t\right)\right), \quad \bar{C}_{i,j}^s(t)\Delta t\right) \tag{9}$$

where $N(x,t)$ is the cumulative vehicle number on place x at time t, $x_{i,j}^{in}$ is the entrance point of the link (i,j), $x_{i,j}^{out}$ is the exit point of the link (i,j). Since we use the double queue model in this paper, we can derive the sending flow rate for DQ as

$$S_{i,j}(t) = \begin{cases} \min\left(p_{i,j}\left(t - \tau_{i,j}^0\right), \quad \bar{C}_{i,j}^s(t)\right), & if\ q_{i,j}^D(t) = 0 \\ \bar{C}_{i,j}^s(t), & if\ q_{i,j}^D(t) > 0 \end{cases} \tag{10}$$

Equations (9) and (10) are the same as shown in Lemma B.1 in **Appendix B**. The queueing dynamics in (7) is similar to the point queue model (Ban et al., 2012b; Daganzo, 1995; Nie and Zhang, 2005; Vickrey, 1969). To make the sending flow continuous, we adopt the modified point queue idea proposed in Ban et al. (2012b) and express Equation (10) as

$$S_{i,j}(t) = \min\left(\bar{C}_{i,j}^s(t), \quad p_{i,j}\left(t - \tau_{i,j}^0\right) + \alpha q_{i,j}^D(t)\right) \tag{11}$$

where $\alpha > 0$ is a smoothing parameter. For the class-specified sending flow, it is common to assign them proportionally based on the downstream queue length:

$$S_{i,j}^{*}(t) = S_{i,j}(t)\frac{q_{i,j}^{D,*}(t)}{q_{i,j}^{D}(t)} \tag{12}$$

For the maximum receiving flow, we have

$$R_{i,j}(t) = \begin{cases} \min\left(\bar{C}_{i,j}^{r}(t),\ \ v_{i,j}\left(t-\tau_{i,j}^{w}\right)\right), & if\ q_{i,j}^{U}(t) = \bar{Q}_{i,j}(t) \\ \bar{C}_{i,j}^{r}(t) & if\ q_{i,j}^{U}(t) < \bar{Q}_{i,j}(t) \end{cases} \tag{13}$$

Here $\bar{C}_{i,j}^{r}(t)$ is the capacity of receiving flow of link (i,j) at time t. Equation (13) shows that, if the upstream queue has reached the queue storage capacity of the link (i.e., $q_{i,j}^{U}(t) = \bar{Q}_{i,j}(t)$), the possible maximum receiving flow is the minimum between the receiving flow capacity $\bar{C}_{i,j}^{r}(t)$ and the exit flow of the link at an earlier time (i.e. $\tau_{i,j}^{w}$ based on the upstream queue dynamics Equation (4)), so that the upstream queue would not increase anymore. If the upstream queue is less than the queue capacity (i.e., $q_{i,j}^{U}(t) < \bar{Q}_{i,j}(t)$), the possible maximum receiving flow is the receiving flow capacity $\bar{C}_{i,j}^{r}(t)$. Note that in the numerical section of this paper, the sending capacity and receiving capacity are set as the same, i.e., $\bar{C}_{i,j}^{s}(t) = \bar{C}_{i,j}^{r}(t) = \bar{C}_{i,j}(t)$. Having the maximum sending flow $S_{i,j}^{*}(t)$ of the incoming links and the maximum receiving flow $R_{j,k}(t)$ of the outgoing links, we need to assign the sending flow from an incoming link to a downstream link, i.e., to determine the flow transfer $G_{(i,j)\to(j,k)}^{*}(t)$ for every adjacent link pair. This will be shown in the nodal model next. Let I_j be the set of the incoming nodes that are directly linked to node j, and K_j be the set of outgoing nodes that are directly linked from node j. Given $G_{(i,j)\to(j,k)}^{*}(t)$, we can connect the nodal model with the double queue model by

$$\begin{aligned} p_{j,k}^{*}(t) &= \sum_{i\in I_j} G_{(i,j)\to(j,k)}^{*}(t) \\ v_{i,j}^{*}(t) &= \sum_{k\in K_j} G_{(i,j)\to(j,k)}^{*}(t) \end{aligned} \tag{14}$$

3.3.3. Nodal model

There could be multiple incoming and outgoing links associated with node j. The nodal model determines the flow transfer $G_{(i,j)\to(j,k)}(t)$ from an incoming link (i,j) to an outgoing link (j,k) where $i \in I_j$ and $k \in K_j$. We present the nodal model under different cases of node j based on the nodal model introduced in Yperman (2007).

(i) Origin node

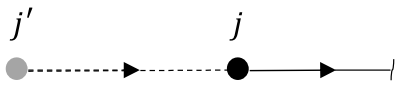


Figure 4. Origin node

We add a pseudo origin node to the network to guarantee that there is only one outgoing link (j',j). The free-free travel time and the shockwave propagation time of this pseudo link is 0 (i.e., $\tau_{j',j}^{0} = \tau_{j',j}^{\omega} = 0$), and there is no limit on the upstream queue of the pseudo link (i.e., $\bar{Q}_{j',j}(t) = +\infty$). Thus, the receiving capacity of the pseudo link is unlimited (i.e., $R_{j',j}(t) = +\infty$). The demand flow $d_{j'}^{*}(t)$ gets into the network from the pseudo node j'. The flow transfer from the origin node is then

$$\boldsymbol{G}_{(j',j)}^{*}(\boldsymbol{t}) = \boldsymbol{d}_{j'}^{*}(\boldsymbol{t}) \tag{15}$$

(ii) Destination node

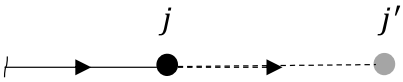


Figure 5. Destination node

Similarly, we add a pseudo destination node to the network to guarantee that there is only one incoming link (j, j'). The pseudo link shares the same property as discussed in (i), i.e., $\tau^0_{j,j'} = \tau^{\omega}_{j,j'} = 0, \bar{Q}_{j,j'} = +\infty, R_{j,j'}(t) = +\infty$. The flow transfer to the pseudo destination node is

$$\boldsymbol{G^*_{(j,j')}(t) = S^*_{j,j'}(t)} \tag{16}$$

(iii) General multi-in and multi-out node

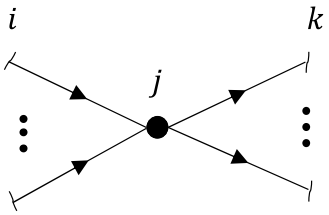


Figure 6. A general node

The flow transfers are usually discussed separately for a diverge node and a merge node (e.g., Yperman (2007), Jin and Zhang (2003)). We here combine these two types of nodes and build a general nodal model for a multi-in and multi-out node, which is also integrated with our optimal control based network model (see Section 3.5). The following five conditions define the rules and constraints that the flow transfer $G^*_{(i,j)\to(j,k)}(t)$ needs to satisfy. Since we are developing a network model that integrates both behavior of HDVs and CAVs and the traffic flow dynamics, the actual values of $G^*_{(i,j)\to(j,k)}(t)$ will need to be determined jointly by solving both the traffic flow model and the behavior model, i.e., the network wide OCPEC as presented in Section 3.5.

1) As described in Section 3.2, we model the route choice behavior of HDVs by IDUE, which is mathematically formulated in Section 3.4. Given the network information and the maximum sending flow of incoming links $S^{hdv}_{i,j}(t)$, the IDUE determines the routes for HDV and assign $S^{hdv}_{i,j}(t)$ to the outgoing links that are on the shortest paths, which gives $S^{hdv}_{(i,j)\to(j,k)}(t)$, the maximum HDV flow that may be transferred from the incoming link (i, j) to the outgoing link (j, k).
2) The flow transfer from a specific incoming link should be upper-bounded by the maximum sending flow. For HDVs, we have the maximum sending flow from (i, j) to (j, k), thus

$$\boldsymbol{G^{hdv}_{(i,j)\to(j,k)}(t) \le S^{hdv}_{(i,j)\to(j,k)}(t)} \tag{17}$$

For CAVs, we only have the maximum sending flow from (i, j), thus

$$\sum_{\boldsymbol{k\in K_j}} \boldsymbol{G^{cav}_{(i,j)\to(j,k)}(t) \le S^{cav}_{i,j}(t)} \tag{18}$$

3) The total flow transfer of HDVs and CAVs to a specific outgoing link should be upper-bounded by the receiving capacity of link (j, k), thus

$$\sum_{i\in I_j} \boldsymbol{G}^{\boldsymbol{cav}}_{(i,j)\to(j,k)}(\boldsymbol{t}) + \sum_{i\in I_j} \boldsymbol{G}^{\boldsymbol{hdv}}_{(i,j)\to(j,k)}(\boldsymbol{t}) \le \boldsymbol{R}_{j,k}(\boldsymbol{t}) \tag{19}$$

4) The flow transfer of HDVs from each incoming link should be proportionally distributed based on the maximum sending flow, thus

$$\frac{\boldsymbol{G}^{\boldsymbol{hdv}}_{(i,j)\to(j,k)}(\boldsymbol{t})}{\sum_{i\in I_j} \boldsymbol{G}^{\boldsymbol{hdv}}_{(i,j)\to(j,k)}(\boldsymbol{t})} = \frac{\boldsymbol{S}^{\boldsymbol{hdv}}_{(i,j)\to(j,k)}(\boldsymbol{t})}{\sum_{i\in I_j} \boldsymbol{S}^{\boldsymbol{hdv}}_{(i,j)\to(j,k)}(\boldsymbol{t})} \tag{20}$$

5) According to Equation (12), the flow transfers of HDVs and CAVs from the same incoming link should be proportionally distributed, thus

$$\frac{\sum_{k\in K_j} \boldsymbol{G}^{\boldsymbol{cav}}_{(i,j)\to(j,k)}(\boldsymbol{t})}{\sum_{k\in K_j} \boldsymbol{G}^{\boldsymbol{hdv}}_{(i,j)\to(j,k)}(\boldsymbol{t})} = \frac{\boldsymbol{q}^{\boldsymbol{D,cav}}_{\boldsymbol{i,j}}(\boldsymbol{t})}{\boldsymbol{q}^{\boldsymbol{D,hdv}}_{\boldsymbol{i,j}}(\boldsymbol{t})} \tag{21}$$

3.4. Route Choice Behavior of HDVs

For HDVs, the objective is to minimize their individual travel cost (travel time is used in this paper). As discussed in Section 3.1, we can formulate the route choice behavior of HDVs as IDUE by the following complementarity conditions (Ban et al., 2012a) (denoted as **P1**):

(i) Route choice (at a junction node j)

$$0 \le \hat{p}^{hdv}_{j,k}(t) \perp \tau_{j,k}(t) + \eta^s_k(t) - \eta^s_j(t) \ge 0 \tag{22}$$

(ii) Flow conservation (at a junction node j)

$$0 \le \eta^s_j(t) \perp \sum_{k\in K_j} \hat{p}^{hdv}_{j,k}(t) - \sum_{i\in I_j} S^{hdv}_{i,j}(t) \ge 0 \tag{23}$$

Similar to (1-1) – (1-2) in Section 3.1, Equation (22) shows that travelers at node j will select to enter link (j,k) (i.e., $\hat{p}^{hdv}_{j,k}(t) > 0$) if link (j,k) is on a route with the minimum instantaneous travel time from node j to destination s ($s \neq j$), i.e., $\tau_{j,k}(t) + \eta^s_k(t) - \eta^s_j(t) = 0$. Otherwise, the traveler will not enter link (j,k) (i.e., if $\tau_{j,k}(t) + \eta^s_k(t) - \eta^s_j(t) > 0$, $\hat{p}^{hdv}_{j,k}(t) = 0$). This follows exactly the IDUE principle. For Equation (23), as it can be easily shown that the instantaneous route travel time $\eta^s_j(t) > 0$ for $s \neq j$, the equality always holds for the right-hand side of (23). That is, Equation (23) indicates "flow conservation", which guarantees that the sum of all sending flows of HDVs exiting from upstream links of node j (represented by $\sum_{i\in I_j} S^{hdv}_{i,j}(t)$) is equal to all outgoing flows from node j to downstream links (represented by $\sum_{k\in K_j} \hat{p}^{hdv}_{j,k}(t)$). From the point of view of a specific node, all traffic flows from that node to the same destination will be assigned to the outgoing links that are on the paths with the minimum travel time at time t. These flows will be re-assigned based on the traffic condition at that time when they arrive other downstream nodes.

The IDUE model above is slightly different from the "regular" IDUE models in the literature (e.g., the one in Ban et al. (2012a)). First, instead of using the actual demands of HDVs from a node, we use $S^{hdv}_{i,j}(t)$, the maximum sending flow from the upstream link of a node j (as defined in the mixed traffic flow model in Section 3.3.2). Thus, the IDUE model here is to determine the "optimal" maximum sending flow that can be assigned to a downstream link of the same node j. This is represented by the variable $\hat{p}^{hdv}_{j,k}(t)$ in the IDUE model, which also indicates which outgoing links HDVs will select as their routes. Here $\hat{p}^{hdv}_{j,k}(t)$ is not the actual HDV flow $p^{hdv}_{j,k}(t)$ in the DQ model (i.e., Equation (5) and (14)). The actual HDV flow have

to follow the mixed flow DQ dynamics and nodal constraints. But the two are closely related; see below after we present Equation (24).

Given $\hat{p}_{j,k}^{hdv}(t)$, we can calculate $S_{(i,j)\to(j,k)}^{hdv}(t)$, the maximum sending flow of HDVs from incoming links (i,j) to outgoing links (j,k), by

$$S_{(i,j)\to(j,k)}^{hdv}(t) = S_{i,j}^{hdv}(t)\frac{\hat{p}_{j,k}^{hdv}(t)}{\sum_{k'}\hat{p}_{j,k'}^{hdv}(t)}. \tag{24}$$

Notice that $S_{(i,j)\to(j,k)}^{hdv}(t)$ is used in the nodal model in Section 3.3.3. We next show that $\hat{p}_{j,k}^{hdv}(t) > 0$ implies $p_{j,k}^{hdv}(t) > 0$. If $\hat{p}_{j,k}^{hdv}(t) > 0$, i.e., HDVs select (j,k) as the next link of their routes, by Equation (24), the maximum sending flow should follow the proportional distribution, which yields $S_{(i,j)\to(j,k)}^{hdv}(t) > 0$. Based on the nodal constraints Equation (17)-(20), if there is certain receiving capacity of the outgoing link (j,k), the flow transfer of HDVs from (i,j) to (j,k) should also satisfy $G_{(i,j)\to(j,k)}^{hdv}(t) > 0$. Thus, based on Equation (14), the actual HDV flow should satisfy $p_{j,k}^{hdv}(t) > 0$. This guarantees that the HDVs route $\hat{p}_{j,k}^{hdv}(t)$ calculated from the IDUE principle (i.e., Equation (22) and (23)) will indeed be followed by HDVs as their route choices (as represented by $p_{j,k}^{hdv}(t)$).

3.5. Route Choice Behavior of CAVs and the Dynamic Bi-level Problem.

The route choice behavior of CAVs follows DSO and is designed to improve the performance from the system perspective. With this objective in mind, designing the routes for CAVs is essentially an optimal control problem (OCP), which can be cast as a dynamic bi-level problem. The CAVs are at the upper level and the HDVs are at the lower level. Our main OCP (i.e., the upper-level problem) is to generate routes for CAVs, of which the HDVs should follow the IDUE principle (i.e., the lower-level problem, **P1**). Meanwhile, HDVs and CAVs should follow the DQ dynamics and nodal constraints as defined in the mixed traffic flow model earlier. Together, we can formulate this dynamic bi-level problem as an OCPEC, by considering the objective of CAVs as to minimize the total system travel time. The OCPEC (denoted as **P2**) is formulated as follows.

(i) Objective

$$\min_{G^{cav}} \int_0^{t_f} \sum_{(i,j)} \sum_{*} n_{i,j}^{*}(t)\, dt \tag{25}$$

(i) DQ dynamics: (4)-(7),
(ii) Maximum sending/receiving flows: (11)-(14),
(iii) Nodal model (at node j): (15) - (21),
(iv) Equilibrium constraints (i.e., the lower-level problem **P1**): (22)-(24)
(v) Boundary conditions: $n_{i,j}^{cav}(t_0) = q_{i,j}^{D,cav}(t_0) = n_{i,j}^{hdv}(t_0) = q_{i,j}^{D,hdv}(t_0) = 0$. (26)

The above continuous-time OCPEC consists of non-linear constraints (i.e., the DQ dynamics and the nodal constraints) and equilibrium constraints (i.e., the route choice of HDVs). It integrates both the mixed traffic dynamics model and the behavior models of HDVs and CAVs, which provides a modeling framework to study the behavior, interaction, and impact of HDVs and CAVs on the network level. Solving this OCPEC model will jointly determine (optimally) the inflow/exit flow of each link for both CAVs and HDVS, for which the inflows to links also represent the route choices of HDVs and CAVs.

OCPEC contains constant time delays in the DQ dynamics (i.e., $p_{i,j}\left(t-\tau_{i,j}^{0}\right)$ and $v_{i,j}\left(t-\tau_{i,j}^{w}\right)$ in Equation (4)). In fact, the time delays will become time-varying if we assume a varying shockwave speed, making the OCPEC framework, while still applicable, more complicated to analyze and solve. As

discussed earlier, this is one of the major reasons why we assume constant shockwave speed in Section 3.2. The nonlinearity, time delays, and especially the equilibrium constraints make the OCPEC hard to be solved analytically. Classical methods aiming to provide the analytical optimality conditions for an OCP, such as the Pontryagin's maximum principle and dynamic programming, can hardly deal with such state-constrained problems with time delays either. In addition, the complementarity constraints (i.e., the equilibrium constraints for HDVs), even being reformulated as inequality constraints, make the classical necessary optimality conditions for OCP with only equality and inequality constraints non-applicable (Guo and Ye, 2016). Nevertheless, one can still numerically solve such a problem by using appropriate time discretization to convert the continuous OCPEC to a finite-dimensional optimization problem. This way, the OCPEC can be discretized as a non-convex and non-smooth mathematical programming with equilibrium constraints (MPEC). In the next section, we show how to discretize the OCPEC to MPEC and discuss the numerical solution techniques.

4. DISCRETIZATION AND SOLUTION METHODS

4.1. Discretization

The continuous OCPEC needs to be discretized by carefully selecting the discrete time step. Previous studies (Ban et al., 2012a; Ma et al., 2018) suggested that the time step should be small enough such that the free-flow travel times and shockwave propagation times of all regular links should be multiples of the time step. We use a time step $h > 0$ to discretize the time interval $[t_0, t_f]$ into $N_h = (t_f - t_0)/h$ (assumed to be integer) time steps, each of equal length h:

$$t_0 \triangleq t_0^h \le t_1^h \le \cdots \le t_{N_h}^h \triangleq t_f \tag{27}$$

We use the implicit backward-Euler difference method to discretize the whole problem. Ban et al. (2012b) provided more discussions on why the implicit discretization scheme should be used. We notice here that the implicit scheme is similar to defining the link travel time at the end of a discrete time interval used in Ban et al. (2008) or the concept of predictive DUE in Heydecker and Verlander (1999). After discretization, the continuous OCPEC **P2** can be converted to a MPEC, denoted as **P3**:

(i) Objective function:

$$\min \sum_{r=1}^{N_h} \sum_{(i,j)\in\mathcal{L}} \sum_{*} n_{i,j}^{*,h,r}(t) \tag{28}$$

with the decision variables

$$\left\{ G_{(i,j)\to(j,k)}^{cav,h,r} \right\}_{r=1}^{N_h}, \quad i \in I_j, k \in K_j, j \in \mathcal{N} \tag{29}$$

(ii) DQ dynamics

- Upstream queue dynamics

$$q_{i,j}^{U,h,r} - q_{i,j}^{U,h,r-1} = h\left(p_{i,j}^{h,r} - v_{i,j}^{h,r-\tau_{i,j}^{w,h}} \right) \tag{30}$$

with $v_{i,j}^{h,r-\tau_{i,j}^{w,h}} = 0$ for $r = 1,2,\ldots \tau_{i,j}^{w,h}$.

- Downstream queue dynamics

$$q_{i,j}^{D,h,r} - q_{i,j}^{D,h,r-1} = h\left(p_{i,j}^{h,r-\tau_{i,j}^{0,h}} - v_{i,j}^{h,r} \right) \tag{31}$$

with $p_{i,j}^{h,r-\tau_{i,j}^{0,h}} = 0$ for $r = 1,2,\ldots \tau_{i,j}^{0,h}$. Consider vehicle types, we have

$$q_{i,j}^{D,*,h,r} - q_{i,j}^{D,*,h,r-1} = h\left(p_{i,j}^{*,h,r-\tau_{i,j}^{0,h}} - v_{i,j}^{*,h,r}\right) \tag{32}$$

with $p_{i,j}^{*,h,r-\tau_{i,j}^{0,h}} = 0$ for $r = 1,2,\ldots \tau_{i,j}^{0,h}$.

- Total flow dynamics

$$n_{i,j}^{*,h,r} - n_{i,j}^{*,h,r-1} = h\left(p_{i,j}^{*,h,r} - v_{i,j}^{*,h,r}\right) \tag{33}$$

- Link travel time

$$\tau_{i,j}^{h,r} = \tau_{i,j}^{0,h} + q_{i,j}^{D,h,r}/\bar{C}_{i,j}^{h,r} \tag{34}$$

(iii) DQ dynamics

- Sending flow

$$S_{i,j}^{h,r} = \min\left(\bar{C}_{i,j}^{h,r},\ \ p_{i,j}^{h,r-\tau_{i,j}^{0,h}} + \alpha q_{i,j}^{D,h,r}\right) \tag{35}$$

- Type-specified sending flow

$$\frac{S_{i,j}^{*,h,r}}{S_{i,j}^{h,r}} = \frac{q_{i,j}^{D,*,h,r}}{q_{i,j}^{D,h,r}} \tag{36}$$

- Receiving flow

$$R_{i,j}^{h,r} = \min\left(\bar{C}_{i,j}^{h,r},\ \ \bar{Q}_{i,j}^{h,r} - q_{i,j}^{U,h,r} + v_{i,j}^{h,r-\tau_{i,j}^{w,h}}\right) \tag{37}$$

- Connection to nodal model

$$p_{j,k}^{*,h,r} = \sum_{i \in I_j} G_{(i,j)\to(j,k)}^{*,h,r},\quad v_{i,j}^{*,h,r} = \sum_{k \in K_j} G_{(i,j)\to(j,k)}^{*,h,r} \tag{38}$$

(iv) Nodal model:

- Origin node

$$G_{(j',j)}^{*,h,r} = d_{j'}^{*,h,r} \tag{39}$$

- Destination node

$$G_{(j,j')}^{*,h,r} = S_{j,j'}^{*,h,r} \tag{40}$$

- The upper bound of the transition flow of HDVs and CAVs

$$\begin{aligned} G_{(i,j)\to(j,k)}^{hdv,h,r} &\le S_{i,j}^{hdv,h,r} \\ \sum_{k \in K_j} G_{(i,j)\to(j,k)}^{cav,h,r} &\le S_{i,j}^{cav,h,r} \end{aligned} \tag{41}$$

- The upper bound of total HDVs and CAVs flow to outgoing links

$$\sum_{i \in I_j} G_{(i,j)\to(j,k)}^{cav,h,r} + \sum_{i \in I_j} G_{(i,j)\to(j,k)}^{hdv,h,r} \le R_{j,k}^{h,r} \tag{42}$$

- The proportional relationship between transition flows and sending flows of HDVs

$$\frac{G_{(i,j)\to(j,k)}^{hdv,h,r}}{\sum_{i\in I_j} G_{(i,j)\to(j,k)}^{hdv,h,r}} = \frac{S_{(i,j)\to(j,k)}^{hdv,h,r}}{\sum_{i\in I_j} S_{(i,j)\to(j,k)}^{hdv,h,r}} \tag{43}$$

- The proportional relationship between HDVs and CAVs

$$\frac{\sum_{k\in K_j} G_{(i,j)\to(j,k)}^{cav,h,r}}{\sum_{k\in K_j} G_{(i,j)\to(j,k)}^{hdv,h,r}} = \frac{q_{i,j}^{D,cav,h,r}}{q_{i,j}^{D,hdv,h,r}} \tag{44}$$

(v) Equilibrium constraints (i.e., route choice behavior of HDVs)

- Route choice

$$0 \le \hat{p}_{j,k}^{hdv,h,r+1} \perp \tau_{j,k}^{h,r} + \eta_k^{s,h,r} - \eta_j^{s,h,r} \ge 0 \tag{45}$$

- Flow conservation

$$0 \le \eta_j^{s,h,r} \perp \sum_{k\in K_j} \hat{p}_{j,k}^{hdv,h,r+1} - \sum_{i\in I_j} S_{i,j}^{hdv,h,r+1} \ge 0 \tag{46}$$

- Routes assignment

$$S_{(i,j)\to(j,k)}^{hdv,h,r+1} = S_{(i,j)}^{hdv,h,r+1} \frac{\hat{p}_{j,k}^{hdv,h,r+1}}{\sum_{k'} \hat{p}_{j,k'}^{hdv,h,r+1}} \tag{47}$$

(vi) Boundary conditions

$$n_{i,j}^{cav,h,0} = q_{i,j}^{D,cav,h,0} = n_{i,j}^{hdv,h,0} = q_{i,j}^{D,hdv,h,0} = 0 \tag{48}$$

The equilibrium constraints (45) and (46) violates the Magasarian-Fromovitz constraint qualification (MFCQ) (Luo et al., 1996), implying that **P3** cannot be analyzed or solved by classical methods such as the KKT conditions that usually require certain constraint qualifications. Specific first order and second order optimality conditions may be derived for bilevel problems (Luo et al., 1996; Ralph and Wright, 2004), which however require rather restrictive conditions and cannot be applied to the model in this paper. Therefore, **P3** cannot be solved directly as an NLP. In the literature, some relaxation ideas (Ban and Liu, 2009; Ban et al., 2006) were applied to iteratively solve this problem. The relaxed problem has been proven to satisfy the MFCQ (Ralph and Wright, 2004) and thus can be solved by standard NLP solvers. Readers can refer to Ban and Liu (2009) and Ban et al. (2006) for details of the relaxation method that are omitted here. As shown in the numerical results in Section 5, it turns out that the relaxation method can only solve **P3** on small networks with a short time horizon. Therefore, we propose a decomposition based heuristic model predictive control (HMPC) method next to deal with larger networks with longer time horizons.

4.2. Heuristic Model Predictive Control (HMPC) Method

The difficulty of solving **P3** lies in the non-smooth and non-convex nature of the equilibrium constraints. The key idea of the proposed HMPC method is to decompose the dynamic bi-level problem to separate sub-problems that are solved in an iterative manner. At a specific time, the IDUE problem of HDVs is a mixed complementarity problem (MCP) which alone can be readily solved (Ban et al., 2012b). Considering that CAVs should follow DSO to improve the network performance, we establish such a DSO problem in a forward time horizon. However, in this DSO problem, we do not formulate the HDVs as equilibrium constraints for reasons discussed above; instead, we consider HDVs as CAVs, i.e., HDVs will also follow the DSO. The DSO problem established in this way is indeed an NLP problem, which

can be solved relatively easily. The DSO reflects the system optimal routes under ideal conditions but may conflict with the actual IDUE routes of HDVs or violate the traffic dynamics. We then design a route adjustor that takes the IDUE routes of HDVs and the ideal DSO routes of both HDVs and CAVs as input and generates the adjusted IDUE routes for HDVs and the system performance oriented routes for CAVs to satisfy the DQ dynamics and nodal constraints. The overall idea of HMPC is shown in Figure 7.

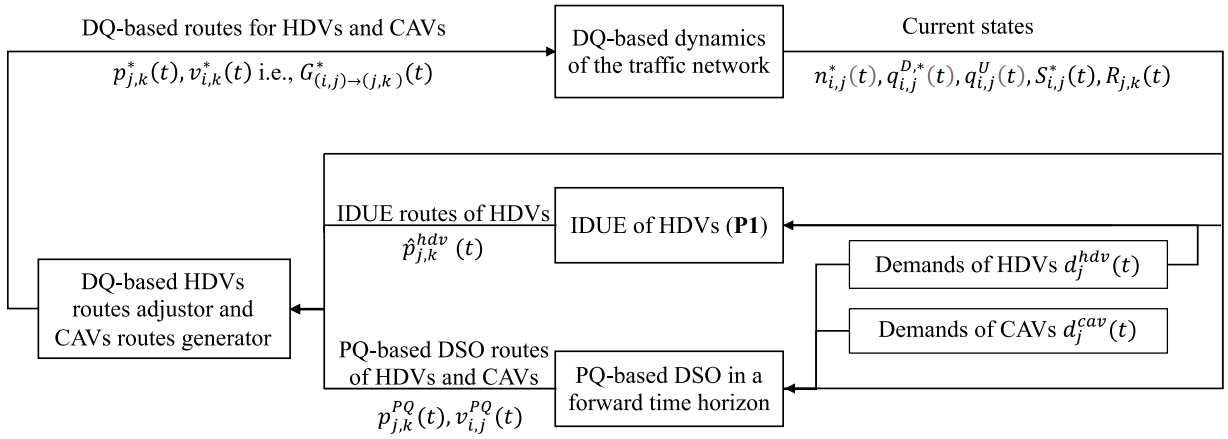


Figure 7. The framework of HMPC

Consider a specific time, given network states and the demand of HDVs from each node to the destination, we find the routes for HDVs based on the IDUE principle. So far, the DQ nodal constraints have not been considered. Meanwhile, given current network states and the future demands of both HDVs and CAVs, we can formulate a DSO problem in a forward time horizon by assuming both HDVs and CAVs will follow the DSO principle. Note that for this DSO, we use the point queue (PQ) to model the dynamics of the network since the PQ-based DSO problem can be solved more efficiently. Having the DSO solution in a forward time horizon, we take the DSO routes of HDVs and CAVs at the first time step as the output, which will be used to generate the real routes for HDVs and CAVs. We then design a route adjustor to adjust the HDVs' routes and generate the CAVs' routes to make it satisfy DQ dynamics. Thus, we have three main components of the proposed HMPC: (a) the IDUE problem of HDVs, (b) the DSO problem in a forward time horizon, and (c) the route adjustor. The IDUE principle of HDVs is the same as **P1**. The next sections will discuss the rest two components.

4.2.1. PQ-based DSO in a forward time horizon

The PQ-based DSO is similar to the OCPEC **P2** except two characteristics. First, in **P2**, HDVs follow the IDUE condition, while CAVs follow the DSO principle. In the PQ-based DSO, both HDVs and CAVs will follow the DSO principle. That means there is no difference between HDVs and CAVs regarding the route choice behavior. Second, the network dynamics are described by DQ in **P2**. In the PQ-based DSO, the network dynamics are modeled by PQ, which means that there are no upper queue bounds. Although the PQ cannot capture the queue spillback, it can be solved more efficiently compared with the DQ. The PQ-based DSO can be formulated as follows (noted as **P4**).

(i) Objective

$$\min_{p^{PQ}_{i,j}} \int_t^{t+\Delta T} \sum_{(i,j)} n_{i,j}(t')\, dt' \tag{49}$$

(ii) Network dynamics

$$\begin{aligned} \dot{q}_{i,j}^{D}(t') &= p_{i,j}^{PQ}\left(t' - \tau_{i,j}^{0}\right) - v_{i,j}^{PQ}(t') \\ \dot{n}_{i,j}(t') &= p_{i,j}^{PQ}(t') - v_{i,j}^{PQ}(t') \\ v_{i,j}^{PQ}(t') &= \min\left(\overline{C}_{i,j}^{s}(t'),\ \ p_{i,j}^{PQ}\left(t' - \tau_{i,j}^{0}\right) + \alpha q_{i,j}^{D}(t')\right) \end{aligned} \tag{50}$$

(iii) Nodal constraints (at node j):

$$\sum_{k \in K_j} p_{j,k}^{PQ}(t') - \sum_{i \in I_j} v_{i,j}^{PQ}(t') - d_j^{hdv}(t') - d_j^{cav}(t') = 0 \tag{51}$$

(iv) Initial conditions

$$\begin{aligned} n_{i,j}(t) &= n_{i,j}^{cav}(t) + n_{i,j}^{hdv}(t) \\ q_{i,j}^{D}(t) &= q_{i,j}^{D,cav}(t) + q_{i,j}^{D,hdv}(t) \end{aligned} \tag{52}$$

Compared with **P2**, there are no equilibrium constraints since both HDVs and CAVs are considered as controllable. **P4** can be discretized as an NLP problem and can be solved by any standard NLP algorithm. Solving **P4** gives $p_{j,k}^{PQ}(t\colon t + \Delta T)$ and $v_{i,j}^{PQ}(t\colon t + \Delta T)$, which represent the DSO routes of the mixed flows of both HDVs and CAVs for a forward time horizon ΔT. Only the first terms $p_{j,k}^{PQ}(t)$ and $v_{i,j}^{PQ}(t)$ are adopted as the control variables. Considering that the DQ is used to model the real traffic dynamics, we need to adjust the DSO routes generated from **P4** and the IDUE routes generated from **P1** so they can follow the DQ dynamics.

4.2.2. Adjusting HDV routes and generating the CAV routes

At a time t, we have the IDUE routes of HDVs (i.e., $\hat{p}_{j,k}^{hdv}(t)$ calculated from Equation (22) and (23)) and the DSO routes of both HDVs and CAVs (i.e., $p_{j,k}^{PQ}(t)$ and $v_{i,j}^{PQ}(t)$). However, these routes and flows are generated without considering the DQ nodal constraints. In addition, the routes for CAVs haven't been assigned yet. In this section, we design a rule-based routes adjustor to generate CAVs routes to improve the system performance, meanwhile adjust CAVs and HDVs routes such that they satisfy the real DQ traffic dynamics and nodal models. To make the following discussion clear, we first illustrate the inputs and outputs of this route generator.

Input	
$S_{(i,j)}^{*}(t)$	The maximum sending flow of HDVs and CAVs from the incoming link (i,j), see Equation (12).
$\hat{p}_{j,k}^{hdv}(t)$	The pseudo IDUE routes of HDVs, see **P1**.
$p_{j,k}^{PQ}(t), v_{i,j}^{PQ}(t)$	PQ-based DSO routes of both CAVs and HDVs, see **P4**.
$R_{j,k}(t)$	Receiving capacity of the outgoing link (j,k), see Equation (10).
Intermediate variables	
$\hat{p}_{j,k}^{cav}(t)$	The pseudo routes for CAVs, an auxiliary variable.
$\hat{S}_{j}^{hdv}(t)$	Total sending flow of HDVs related to node j.
$\hat{S}_{j,hdv}^{cav}(t)$	Total sending flow of CAVs coupled with HDVs related to node j.
$\hat{S}_{j,only}^{cav}(t)$	Total sending flow of CAVs not coupled with HDVs related to node j.
Output	
$p_{j,k}^{*}(t), v_{i,j}^{*}(t)$	Routes of HDVs and CAVs that satisfy DQ dynamics

The first step is to assign CAVs routes using a heuristic method and adjust HDVs and CAVs routes so that they do not exceed the receiving capacity of the outgoing links. Considering that, at a specific time t, the HDVs usually select one and only one outgoing link, we define this link as $\left(j, k^{hdv}\right)$ where

$\hat{p}_{j,k^{hdv}}^{hdv}(t) > 0$. Having the HDVs routes, we fist modify these routes based on the limitation of the receiving capacity, i.e., $\hat{p}_{j,k^{hdv}}^{hdv}(t) = \min\left\{R_{j,k^{hdv}}(t),\ \hat{p}_{j,k^{hdv}}^{hdv}(t)\right\}$. For CAVs, there are two conditions considering different values of $p_{j,k^{hdv}}^{PQ}(t)$.

(i) First, $p_{j,k^{hdv}}^{PQ}(t) > \hat{p}_{j,k^{hdv}}^{hdv}(t)$, which means the solution of the DSO problem **P4** requires more vehicles than the current assigned HDVs to go to link (j, k^{hdv}). Under this condition, we can assign $\hat{p}_{j,k^{hdv}}^{cav}(t) = \min\left\{p_{j,k^{hdv}}^{PQ}(t) - \hat{p}_{j,k^{hdv}}^{hdv}(t),\quad R_{j,k^{hdv}}(t) - \hat{p}_{j,k^{hdv}}^{hdv}(t)\right\}$ CAV flow to link (j, k^{hdv}), and assign the rest of CAV flows exactly according to the optimal routes $\hat{p}_{j,k}^{cav}(t) = \min\left\{p_{j,k}^{PQ}(t),\quad R_{j,k^{hdv}}(t)\right\}$ for $k \in K_j$ and $k \neq k^{hdv}$.

(ii) Second, $p_{j,k^{hdv}}^{PQ}(t) \leq \hat{p}_{j,k^{hdv}}^{hdv}(t)$, meaning that the current HDVs flows to link (j, k^{hdv}) has exceeded the solution of the DSO problem **P4**. Under this condition, we set $\hat{p}_{j,k^{hdv}}^{cav}(t) = 0$, and assign the total CAV flows to the links (j, k) where $k \in K_j$ and $k \neq k^{hdv}$ proportionally by $\hat{p}_{j,k}^{cav}(t) = \min\left\{\left[\sum_{k' \in K_j} p_{j,k'}^{PQ}(t) - \hat{p}_{j,k^{hdv}}^{hdv}(t)\right]\frac{p_{j,k}^{PQ}(t)}{\sum_{k' \in K_j, k' \neq k^{hdv}} p_{j,k'}^{PQ}(t)},\quad R_{j,k^{hdv}}(t)\right\}$

The second step is to modify the HDVs and CAVs routes based on the DQ dynamics and nodal constraints. We first calculate the total sending flow of HDVs from the incoming links $\hat{S}_j^{hdv}(t)$, total sending flow of CAVs coupled with HDVs from the incoming links $\hat{S}_{j,hdv}^{cav}(t)$, and total sending flow of CAVs not coupled with HDVs from the incoming links $\hat{S}_{j,only}^{cav}(t)$. If there are no HDVs sending flows coming from a specific link, the CAVs sending flows from that link contribute to $\hat{S}_{j,only}^{cav}(t)$. Otherwise, the HDVs sending flows are added up as $\hat{S}_j^{hdv}(t)$ and the coupled CAVs flows are added up as $\hat{S}_{j,hdv}^{cav}(t)$. Then, the conditions can be categorized as two types based on whether there are HDVs flows going to the outgoing links.

(i) First, after satisfying the receiving capacity constraint, there are still some HDVs flows going to the outgoing links. In other words, we can find a k^{hdv} such that $\hat{p}_{j,k^{hdv}}^{hdv}(t) > 0$. We calculate the proportion $\rho = \hat{S}_{j,hdv}^{cav}(t)/\hat{S}_j^{hdv}(t)$. Then there are also two types of conditions:

- If $\sum_k \hat{p}_{j,k}^{cav}(t) > \rho\hat{p}_{j,k^{hdv}}^{hdv}(t)$, which means that the outgoing CAVs flow is greater than the incoming CAVs flows coupled with the HDVs flows, the incoming CAVs flows coupled with HDVs are safe. We have $p_{j,k^{hdv}}^{hdv}(t) = \hat{p}_{j,k^{hdv}}^{hdv}(t)$ and $p_{j,k}^{cav}(t) = \hat{p}_{j,k}^{cav}(t)$ for all $k \in K_j$. For an incoming link (i, j), if $S_{i,j}^{hdv}(t) > 0$ (i.e., there is HDVs sending flow), we assign the exit flows of HDVs and CAVs proportionally by $v_{i,j}^{hdv}(t) = \hat{p}_{j,k^{hdv}}^{hdv}(t)\frac{S_{i,j}^{hdv}(t)}{\hat{S}_j^{hdv}(t)}$ and $v_{i,j}^{cav}(t) = \hat{p}_{j,k^{hdv}}^{hdv}(t)\frac{S_{i,j}^{cav}(t)}{\hat{S}_j^{hdv}(t)}$. If $S_{i,j}^{hdv}(t) = 0$, which means there are no HDVs (i.e., only CAVs) sending flows coming from this link, we can assign the rest CAVs flows (i.e., $\sum_k \hat{p}_{j,k}^{cav}(t) - \hat{S}_{j,hdv}^{cav}(t)$) proportionally by $v_{i,j}^{cav}(t) = S_{i,j}^{cav}(t)\frac{\sum_k \hat{p}_{j,k}^{cav}(t) - \hat{S}_{j,hdv}^{cav}(t)}{\hat{S}_{j,only}^{cav}(t)}$.
- If $\sum_k \hat{p}_{j,k}^{cav}(t) \leq \rho\hat{p}_{j,k^{hdv}}^{hdv}(t)$, which means that the outgoing CAVs flows are smaller than the incoming CAVs flows coupled with the HDVs flows (i.e., the receiving capacity is limited such that the outgoing HDVs and the CAVs are not proportionally to the coupled incoming flows), we need to re-assign the HDVs flows by $p_{j,k^{hdv}}^{hdv}(t) =$

$\left(\sum_k \hat{p}_{j,k}^{cav}(t) + \hat{p}_{j,k^{hdv}}^{hdv}(t)\right) \frac{\hat{S}_j^{hdv}(t)}{\hat{S}_j^{hdv}(t)+\hat{S}_{j,hdv}^{cav}(t)+\hat{S}_{j,only}^{cav}(t)}$ and assign the CAVs by $p_{j,k}^{cav}(t) = \hat{p}_{j,k}^{cav}(t) + \hat{p}_{j,k}^{hdv}(t) - p_{j,k}^{hdv}(t)$. Then, we assign the exit flows of HDVs and CAVs from the incoming links based on the proportional principle, i.e., $v_{i,j}^{hdv}(t) = p_{j,k^{hdv}}^{hdv}(t) \frac{S_{i,j}^{hdv}(t)}{\hat{S}_j^{hdv}(t)}$ and $v_{i,j}^{cav}(t) = p_{j,k}^{hdv}(t) \frac{S_{i,j}^{cav}(t)}{\hat{S}_j^{hdv}(t)}$.

(ii) Second, there is no HDVs flow going to the outgoing links, either because there is no receiving capacity of the outgoing links or there are no HDVs sending flows from the incoming links. Under this condition, if there are outgoing flows, they must be CAVs. Therefore, for the outgoing flows, we set $p_{j,k}^{cav}(t) = \hat{p}_{j,k}^{cav}(t)$ and $p_{j,k}^{hdv}(t) = 0$. For the incoming flows of link (i,j), there are two conditions

- If $S_{i,j}^{hdv}(t) > 0$, which means there are HDVs supplies but they cannot go to the outgoing links due to the receiving capacity limitation. Since HDVs and CAVs are assumed to be proportionally coupled, no CAVs flows can exit from this link. Therefore, we set $v_{i,j}^{hdv}(t) = v_{i,j}^{cav}(t) = 0$.
- If $S_{i,j}^{hdv}(t) = 0$, which means there are only CAVs exiting from this link. We assign the HDVs flows by $v_{i,j}^{hdv}(t) = 0$ and CAVs exit flows by $v_{i,j}^{cav}(t) = \sum_k p_{j,k}^{cav}(t) \frac{S_{i,j}^{cav}(t)}{\hat{S}_{j,only}^{cav}(t)}$.

The pseudo codes of this HDV route adjustor and CAV route generator are shown as Algorithm-1 in **Appendix C**. The overall algorithm integrating the IDUE-based HDV route generator, PQ-based DSO route generator and the route adjustor is shown as Algorithm-2 in **Appendix C**.

5. NUMERICAL EXPERIMENT

We first test the relaxation-based and the HMPC solution methods on a five-node network for a short time horizon (30 minutes). We compare the HMPC solution with the solution obtained from the relaxation method. Then, we extend the time horizon to 300 minutes. Under such long horizon, the relaxation method cannot be directly applied to solve the MPEC because of the increased dimension of the problem. Thus, we only apply the HMPC method to this scenario. Finally, we test the HMPC method on the Sioux Falls network to illustrate the proposed model and algorithm on a more general (and larger size) network.

5.1. A Five-node Network

Figure 8 shows the five-node multi-origin, single-destination network. Travelers enter the network from node 1, 2 and 3, and exit the network from node 5. The free-flow travel time $\tau_{i,j}^0$ (minutes), shockwave travel time $\tau_{i,j}^w$ (minutes), capacity when all vehicles are HDVs $\bar{C}_{i,j}^{hdv}$ (vehicles/minute), and capacity when all vehicles are CAVs $\bar{C}_{i,j}^{cav}$ (vehicles/minute) of each link are given in Figure 8. The demand profile is shown in Figure 9. We first solve this problem under 0% and 100% CAV penetrations, respectively. Under 0% CAV penetration, the original problem is reduced to a differential complementarity system (DCS) based IDUE problem (i.e., with no objective), since there are no CAVs and All vehicles are HDVs (that follow the IDUE principle). Under 100% CAV penetration, the original problem is reduced to a DSO problem without the complementarity constraints, which is a regular NLP after discretization. The DSO solution can serve as the ideal case, which provides a lower bound of the system objective (i.e., the total travel time). We then test the proposed relaxation scheme (abbreviated as "RELAX" in this section) and HMPC under different CAV penetrations ranging from 0% to 100%. The overall performance of RELAX and HMPC under different CAV penetration rates and different travel time functions is shown in Figure 10 (a). We then calculate the improvement of network performance of

the RELAX, HMPC and DSO solutions compared with that of the IDUE, the result of which is shown in Figure 10 (b). The improvement of network performance is calculated as the percentage of reduction regarding total travel times of DSO, RELAX, and HMPC compared with IDUE. Note that the total vehicle time of IDUE (calculated when CAV penetration is zero) or DSO (calculated when CAV penetration is 100%) in Figure 10 (a) does not depend on CAV penetration. We show them as two extreme scenarios so that readers can easily compare the performance of different control methods.

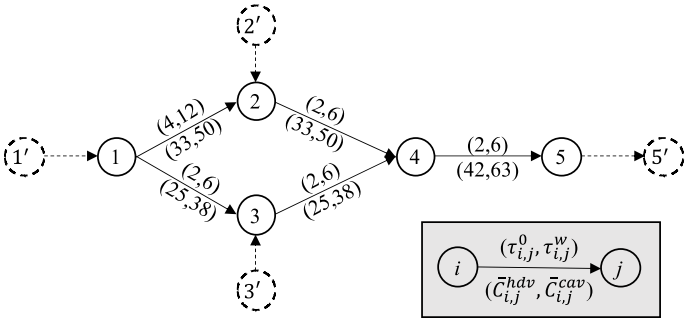


Figure 8. A five-node network (more explanation of the items in the figure needed)

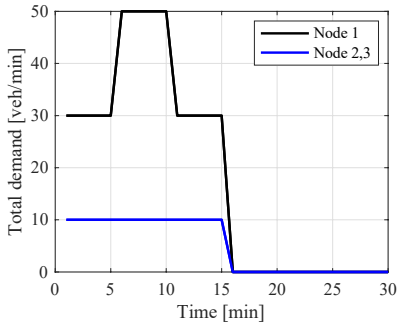
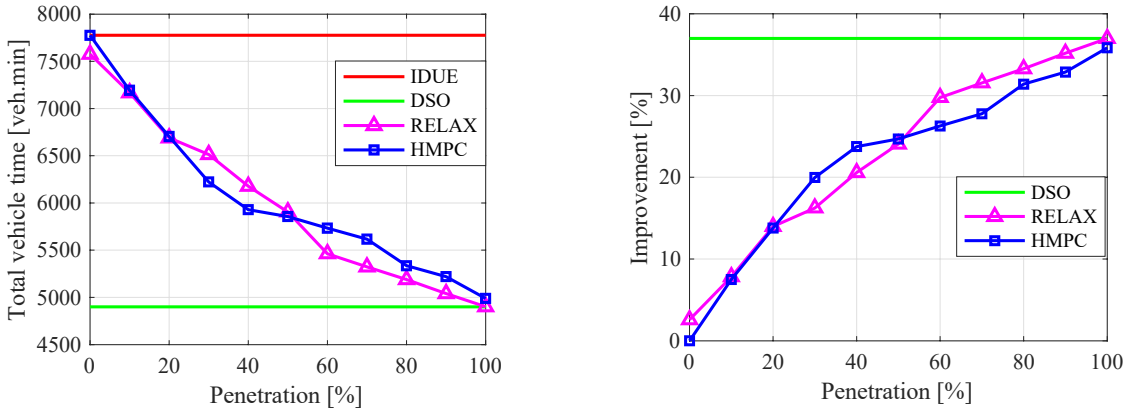


Figure 9. Demand profile



(a) Network performance

(b) Improvement

Figure 10. Performance of different methods

The key observation is, as the CAV penetration increases from 0% to 100%, the system performance obtained by both RELAX and HMPC increases from 0% to 37%. At 0% CAV penetration, all vehicles are HDVs that follow the IDUE principle, which yields a higher system cost. At 0% CAV penetration, the performance of HMPC matches exactly as the IDUE case, while the performance of RELAX is slightly better than the IDUE case. For HMPC, at every time step, we calculate the travel time of each link, based on which we accurately solve the IDUE problem. Then, we solve the DSO problem in a forward time horizon (20 minutes in this paper) and assign flows and update the mixed flow dynamics correspondingly. For RELAX, the whole-time domain problem is solved at once, in which the IDUE is approximated by certain relaxation scheme. Such relaxation-based approximation makes the HDVs' behavior not exactly following the IDUE. This can be used by RELAX to improve the system objective of MPEC, leading to slightly better performance (i.e. smaller objective value) of RELAX compared with that of the IDUE. Under 100% CAV penetration, there is no equilibrium constraints in the MPEC. Thus, RELAX reduces to the DSO problem and yields the same performance as DSO. For HMPC, we cannot guarantee that it can

achieve exactly the DSO performance since it is a heuristic decomposition method. It is observed that the difference of the HMPC objective value and DSO objective value is very small (around 1%). Another observation is that the relationship between performance improvement and CAV penetration is not exactly same to the relationship between link capacity and CAV penetration, i.e., Equation (2). This is due to the fact that performance improvement is related to not only capacity increment (with respect to CAV penetration) but also the route choice behavior of HDVs and CAVs. HDVs' routes, CAVs' routes, and capacity are mutually dependent with each other. The relationship between capacity and CAV penetration is not necessary the relationship between the system performance improvement and CAV penetration. Apart from these overall analyses, in the rest of this section, we firstly validate the mixed flow dynamics and then discuss the route choice behavior of HDVs and CAVs under different methods.

5.1.1. DQ based mixed flow dynamics

The DQ based mixed flow framework consists of link dynamics and the nodal model. We first analyze the link dynamics under 0% CAV penetration (i.e., the IDUE case) where the capacity and upstream queue capacity are fixed since all vehicles are HDVs.

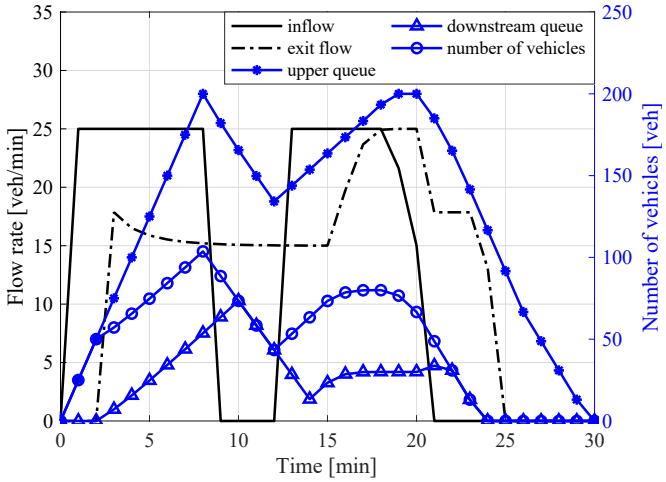
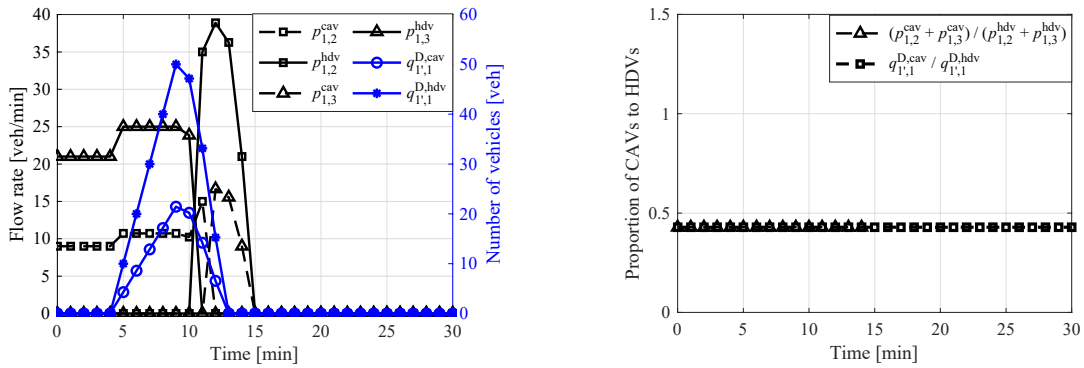


Figure 11. DQ dynamics of link (1,3) in the five-node network

Figure 11illustrates the DQ link dynamics by using link $(1,3)$. The inflow and exit flow rates are shown by the left axis while the upstream queue, downstream queue, and the number of vehicles on the link is shown by the right vertical axis. At the beginning, link $(1,3)$ is on the shortest path from node 1 to the destination. Following the IDUE condition, HDVs are assigned to this link as many as possible. Therefore, the inflow increases to the capacity of link $(1,3)$, i.e., 25 veh/min. Based on the DQ link dynamics Equation (4) and (6), the upstream queue increases since there are no exit flow at the beginning and all inflows contribute to the upstream queue. The same condition holds for the number of vehicles on the link. So, they share a same curve during the interval $[0,2]$. After that, the first inflow arrives at the end of the link, but the exit flow rate is less than the first inflow (since the exit flow from $(3', 3)$ also needs to enter the outgoing link $(3,4)$, but the receiving capacity of link $(3,4)$ cannot allow all exit flows from $(3', 3)$ and $(1,3)$ to get in), which results in a downstream queue. Meanwhile, during $[2,8]$, the upstream queue keeps increasing because the dissipation of the downstream queue has not arrived at the upstream queue (Equation (4)). At $t = 8$, travelers select link $(1,2)$ as their routes since link $(1,3)$ is not on the shortest path anymore due to the increasing travel time. Therefore, the inflow drops to 0 and the upstream queue stops increasing. Also, the number of vehicles deceases and the downstream queue also decreases after 2 minutes delay. After $t = 8$, there is no inflow, but the dissipation has arrived the upstream queue (i.e., $v_{i,j}\left(t - \tau_{i,j}^{w}\right) > 0$ in Equation (4)), the upstream queue starts to decrease. After $t = 12$, travelers select link $(1,3)$ again due to the increasing time of another path. Thereafter, the similar process happens

again. One difference is that the upper queue reaches its capacity at $t = 19$. Therefore, at $t = 20$, inflow to link $(1,3)$ is no longer determined by the link capacity. Instead, based on Equation (4), only $v_{i,j}\left(20 - \tau_{i,j}^{w}\right) = v_{i,j}(14)$ flow can get into link $(1,3)$. This can be observed from Figure 11 that the inflow at $t = 20$ is equal to the exit flow at $t = 14$. As a short summary, the numerical experiments match our DQ-based link dynamics.

The key characteristic of the nodal model is the proportional distribution of HDVs and CAVs from incoming links to outgoing links, i.e., Equation (20) and (21). Specifically, Equation (20) is for a merge node and Equation (21) is for a diverge node. In this section, we show that Equation (21) holds for a diverge node (i.e., node 1). The validation process of Equation (20) is similar (e.g., analyzing node 4) and is omitted here for brevity. We validate Equation (21) by looking into the flow transitions at node 1. To show this, the CAV penetration cannot be 0% or 100%. Here we use the HMPC solution with 30% CAV penetration (randomly selected) as an example. Figure 12 (a) shows the downstream queue of HDVs and CAVs on link $(1', 1)$ and inflow rates of HDVs and CAVs to the outgoing links $(1,2)$ and $(1,3)$. Figure 12 (b) shows the proportion of CAVs to HDVs for the downstream queue as well as the total outgoing flows. It can be shown that the proportion of CAVs to HDVs for the outgoing flows are the same as the proportion of CAVs to HDVs in the downstream queues of the incoming link, which means Equation (21) holds. It should be noted that the sum of inflow rates into links $(1,2)$ and $(1,3)$ is equal to the total demand when the demand is less than the total receiving capacity of links $(1,2)$ and $(1,3)$. If the demand is larger than the total receiving capacity, either due to the high demand or the congestion of outgoing links, the sum of the inflow rates into the two links will be less than the demand at node 1. The remaining demand will be queued at the origin and will be discharged at a later time. This is a typical (also realistic) phenomenon when capacitated physical queues are applied to model traffic flow; see also Ma et al. (2014) and Ma et al. (2018).



(a) Flows and queues at node 1 (b) Proportion of CAVs to HDVs at node 1

Figure 12. Flows and downstream queues at node 1

In addition, we show the FIFO condition of each link (under HMPC, 30% CAV penetration) in Figure 13. The vertical axis of each sub-figure represents the value of the derivative of the link travel time $\dot{\tau}_{i,j}$ and -1 is marked to clearly show the threshold. As discussed in **Proposition 1**, FIFO will hold if $\dot{\tau}_{i,j} > -1$ holds. Although we cannot theoretically guarantee the FIFO condition holds, as shown in Figure 13, such a condition is satisfied for the four links of the five-node network (note that the FIFO condition also holds for link 4 to 5, which is omitted here to save space).

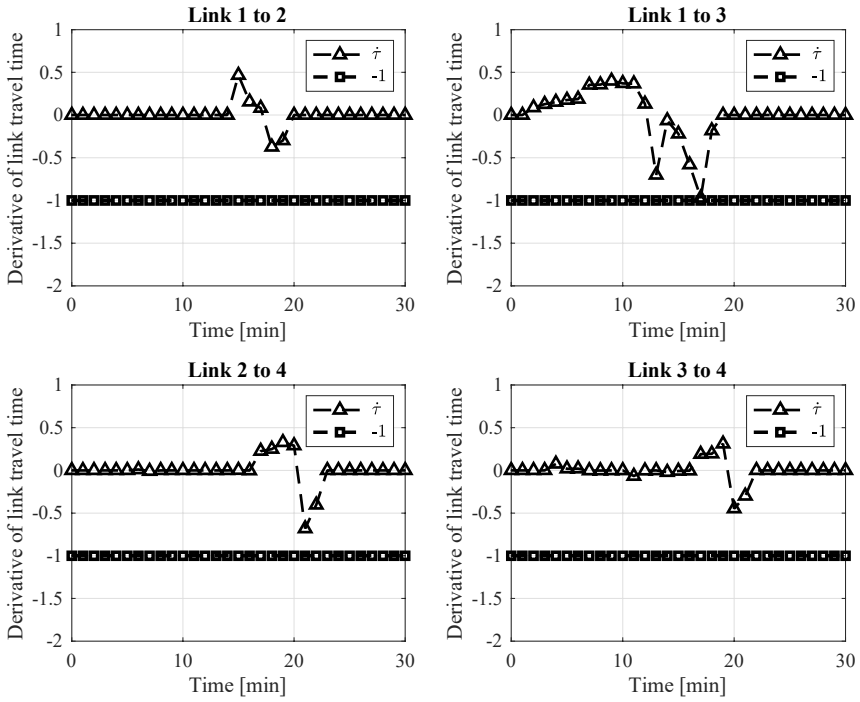


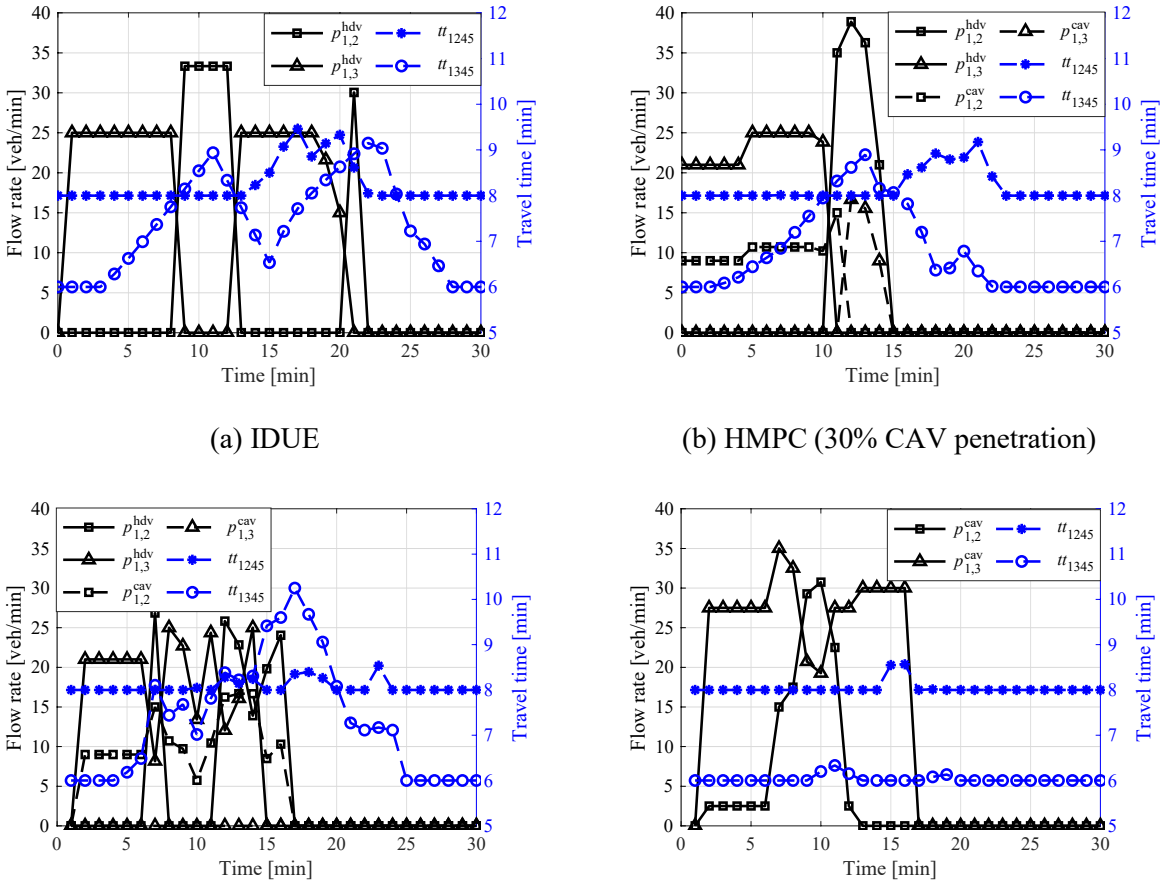
Figure 13. FIFO condition of each link (under HMPC, 30% CAV penetration)

5.1.2. Route choice behavior under different solution methods

There are only two paths from node 1 to the destination, we can analyze the route choice behavior of HDVs and CAVs by analyzing the outgoing flows at node 1. Figure 14 shows how the flows of HDVs and CAVs exiting from node 1 select their routes under different solution methods. $p_{i,j}^{*}$ represents the inflow to outgoing link (i,j) of different vehicle types (i.e., $*$ is either HDV or CAV). tt_{1245} and tt_{1345} represent the instantaneous travel times of path $1 \to 2 \to 4 \to 5$ and $1 \to 3 \to 4 \to 5$, respectively. Under IDUE, Figure 14 (a) clearly shows that HDVs follow the IDUE principle. If the travel time of path $1 \to 2 \to 4 \to 5$ is less than that of $1 \to 3 \to 4 \to 5$, HDVs choose link $(1,2)$ as the next link (i.e., $p_{1,2}^{hdv} > 0$). Otherwise, HDVs select link $(1,3)$ as the next link. The same IDUE route choice behavior of HDVs can be observed from the HMPC by Figure 14 (b). Under RELAX, as shown in Figure 14 (c), HDVs follow the IDUE principle at most time. However, when the travel times of the two candidate routes are close (e.g., $t \in [7,13]$), HDVs may be assigned to both outgoing links $(1,2)$ and $(1,3)$. This is resulted by the relaxation solution method. Such a method relaxes the complementarity equilibrium constraints to nonlinear constraints by introducing relaxation parameters, which will cause errors, especially when the travel times of the two paths are close. Please see Ban and Liu (2009) and Ban et al. (2006) for detailed discussion of the relaxation method. Under DSO, all vehicles are considered as CAVs so there is no relation between the flows and path travel times.

There are mainly two ways by which CAVs improve system performance. First, CAVs can be assigned to paths that are not with the minimum travel time, which can however help reduce network congestion. As shown in Figure 14 (a), under IDUE, HDVs always select the path with minimum travel time and try to select the shortest path with maximum capacity. When CAVs are introduced into the network, as shown in Figure 14 (b) and (c), although HDVs still follow the IDUE principle, the CAVs can be assigned to other paths. These paths are generated based on the objective to improve the network performance. Although the travel times of some CAVs are increased, the network wide performance is improved, as

shown in Figure 10. Second, CAVs can increase link capacity thus improve the network performance. Two extreme conditions are the IDUE where all vehicles are HDVs and the DSO where all vehicles are CAVs. It is shown from Figure 14 (a) and (d) that the capacities of link (1,2) and (1,3) increase when HDVs are replaced by CAVs. As a result, the congestion of the network is reduced (which is why the travel times of the two paths almost don't change in Figure 14 (d)).



(a) IDUE

(b) HMPC (30% CAV penetration)

(c) RELAX (30% CAV penetration)

(d) DSO

Figure 14. Routes of HDVs and CAVs under different methods at node 1

The first way discussed above by which CAVs improve the network performance also brings the fairness (i.e., equity) issue between CAVs and HDVs. As shown in Figure 14 (b), during [0,10] min, CAVs are assigned to route "1-2-4-5", which has longer instantaneous travel time than the route "1-3-4-5". This indicates the CAVs "sacrifice" their instantaneous travel times to help reduce the system-wide congestion of the network. At 11 min, CAVs are assigned to route "1-2-4-5", which is the shortest route. Then, during (11,15] min, CAVs are assigned to route "1-3-4-5" thus again "sacrifice" their instantaneous travel times for improving the system performance. The difference of the actual CAV travel time and the minimum path travel time may be considered as how "unfair" the control scheme imposes to CAVs. We provide more discussions of this later in Section 5.2.

5.1.3. Five-node network with long time horizon

Real world traffic is usually spanned widely at the temporal domain. Thus, we extend the time horizon to 300 minutes for the five-node network, of which the demand profile is shown in Figure 15. For this scenario, solving the MPEC directly by the relaxation method is difficult. We apply the HMPC method to

generate routes for CAVs and compare the network performance with the IDUE and DSO cases. The result is shown in Figure 16, which indicates similar results as those for the short time horizon scenario. As the CAV penetration increases from 0% to 100%, the performance of HMPC increases from 0% to 75%. The difference between the performance of DSO and the HMPC under 100% CAV penetration is only 1.01%, which indicates that the HMPC can achieve almost the same performance as the DSO at 100% CAV penetration. It is shown the maximum improvement of the network performance under this long time horizon scenario is greater than that under the short time horizon scenario. Under the IDUE case, when one path becomes congested, HDVs will not immediately change their routes. Instead, they will stick to the congested path until the instantaneous travel time of that path becomes larger than that of another path. Such phenomenon becomes more dominant if the free-flow travel times of different paths differ. For example, a path with longer free-flow travel time may stay empty while the path with shorter free-flow travel time is heavily congested, as long as the congestion-increased travel time is less than the difference between the free-flow travel times. This lies in the nature of IDUE, since the "instantaneous" travel time is used to calculate travel cost. In addition, the longer time horizon (with travel demand), the heavier the congestion. On the other hand, HMPC could reduce the congestion by assigning CAVs to paths that can help improve network system performance.

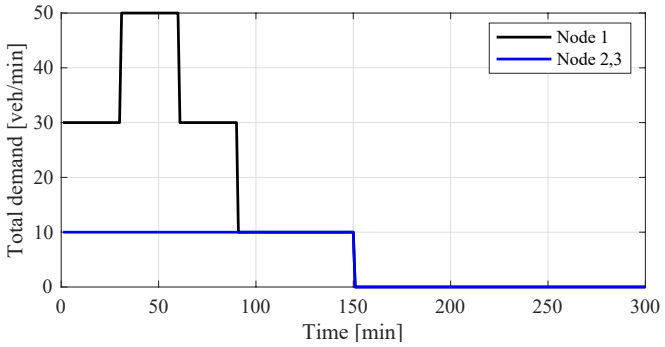
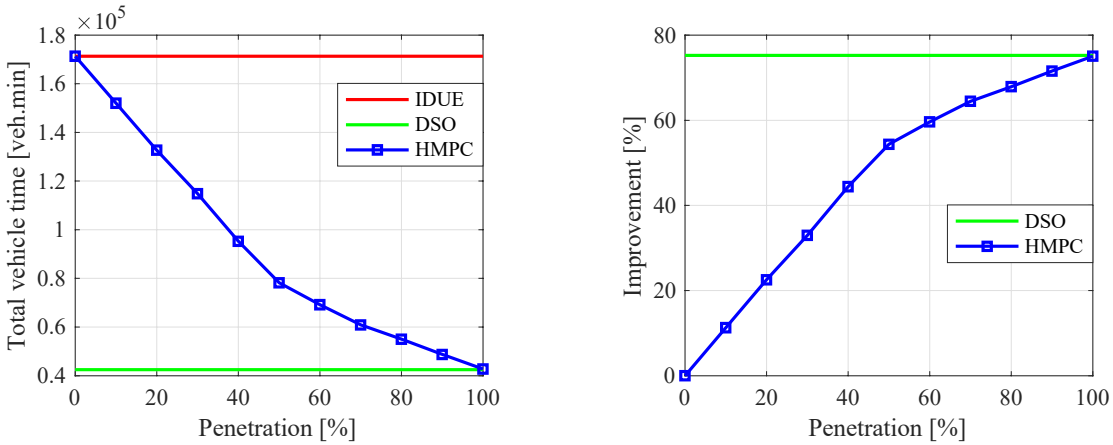


Figure 15. Demand profile of the five-node network with a long time horizon



(a) (b)

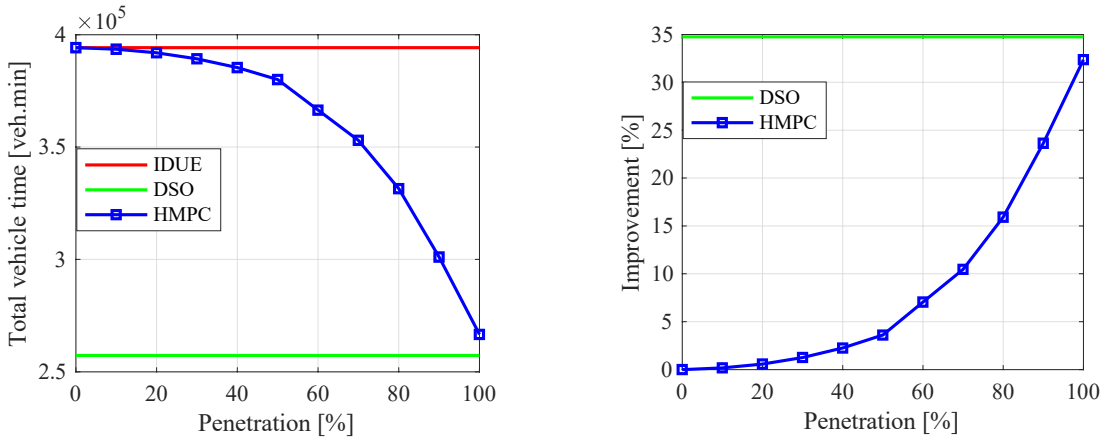
Figure 16. Performance of HMPC on the five-node network

5.2. The Sioux Falls Network

We also test the HMPC on the Sioux Falls network, which is a larger network (compared with the five-node network) with a more general network structure. The geometry of the Sioux Falls network is shown in **Appendix D**. The reader can refer to Ma et al. (2018) and many other published works for the detail of the Sioux Falls network. We set node 15 as the (single) destination and all other 23 nodes as the origins. The demands are adopted from Lablanc (1975), which is shown in Table 1. The total simulation time is 90 minutes. We distribute the daily demand uniformly and apply such demand in the first 30 minutes. Same as the five-node network experiment, we leave a 60-minute clearance time. Figure 17 shows the performances of the HMPC algorithm under different CAV penetration rates, compared with both the performances of IDUE and DSO. The figure shows similar results as that in Figure 16 except for two aspects. First, the curves of HMPC performances in the Sioux Falls network show more "nonlinear" trends compared with those of the five-node network. This may be caused by the different network structure of the two scenarios. Second, the maximum improvement of the Sioux Falls network is far less than the maximum improvement of the five-node network with long time horizon, but similar to the maximum improvement of the five-node network with short time horizon.

Table 1. Demands to node 15 of the Sioux Falls network (thousand vehicles per day)

Node	Demand	Node	Demand	Node	Demand	Node	Demand
1	5	7	5	13	7	19	8
2	1	8	6	14	13	20	11
3	1	9	9	15	0	21	8
4	5	10	40	16	12	22	26
5	2	11	14	17	15	23	10
6	2	12	7	18	2	24	4

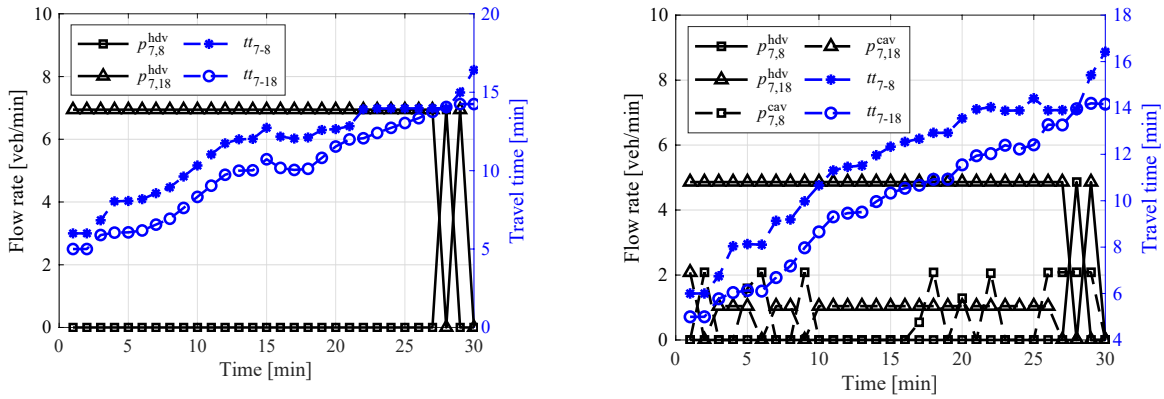


(a) Performance (b) Improvement

Figure 17. Performance of HMPC on Sioux Falls network

As shown in Figure 18, we illustrate the route choices behavior for Sioux Falls network by analyzing the routes of HDVs and CAVs entering the network from node 7, since there are only two candidate routes for this node which could make the analysis clear. Note that in Figure 18, tt_{i-j} represents the sum of the travel time of link (i,j) and the minimum travel time from node j to the destination. As shown in Figure 18 (a), if all vehicles are HDVs following IDUE, the HDVs select the link on the minimum travel time

path as the route. Specifically, during the first 27 minutes, link (7,18) is on the minimum travel time path, therefore, all the HDVs flow into link (7,18). At $t = 28$ min, the travel time through link (7,8) is slightly shorter than that through (7,18), thus HDVs select (7,8) as their route. After that, link (7,18) is again the on the shortest travel time path, thus HDVs flow into link (7,18). Under HMPC, as shown in Figure 18 (b), HDVs still follow the IDUE condition, while CAVs' routes are not related to the minimum travel time path.



(a) IDUE (b) HMPC

Figure 18. Routes of HDV and CAV flows entering the network from node 7 under IDUE and HMPC

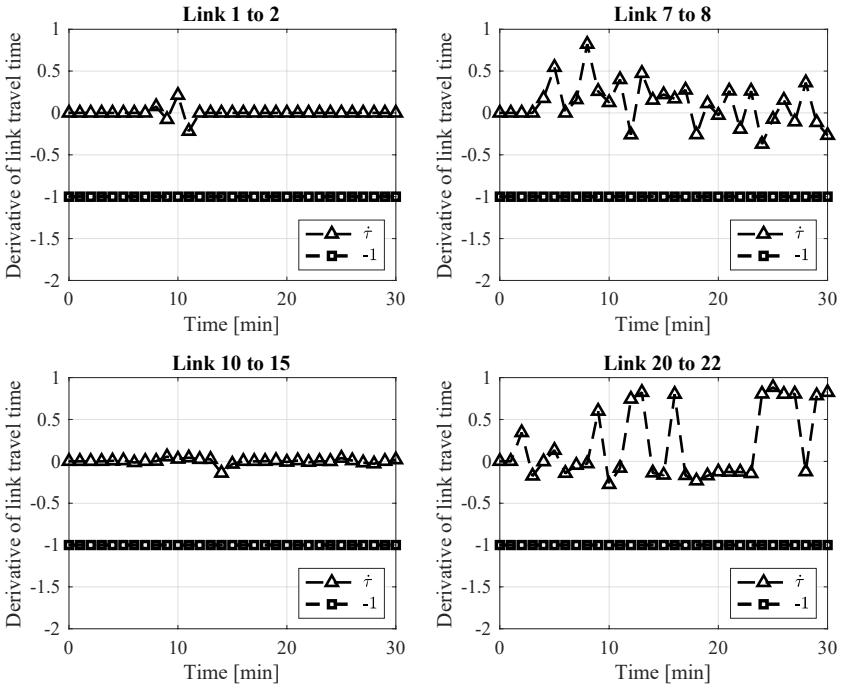
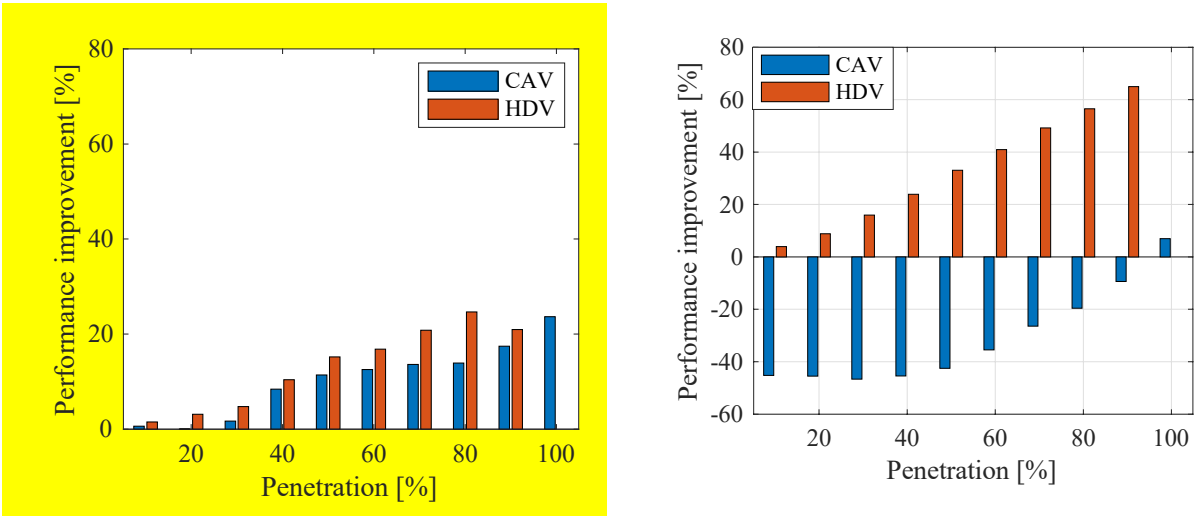


Figure 19. FIFO condition of four randomly selected links (under HMPC, 30% CAV penetration)

To check whether the FIFO condition is satisfied for the Sioux Falls network, Figure 19 shows the FIFO condition of four randomly selected links (under HMPC, 30% CAV penetration). It shows that the FIFO conditions (i.e., $\dot{\tau}_{i,j} > -1$) are all satisfied for the selected links.

Before concluding the numerical experiments, we present more results and discussions about the equity issues of the proposed control scheme to CAVs. First, for a given CAV penetration, for the HMPC case (i.e., with the control of CAVs), we can calculate the total travel times of all CAVs and all HDVs separately. Then, we assign both HDVs and CAVs to follow IDUE for the same CAV penetration (i.e., without the control of CAVs), for which the capacities of links will change according to the penetration of CAVs; see Equation (2) in the manuscript. We also collect the total travel times of all CAVs and all HDVs separately for this without-control case. Finally, we compare the performance (i.e., total travel times) of HDVs and CAVs, respectively, under the above two cases (i.e., with control and without control of CAVs) to analyze the equity issue. We test this for various penetration rates of CAVs. Figure 1 shows the performance improvements of HDVs and CAVs under the control of CAVs (i.e., HMPC) compared with the case of without control of CAVs.

For the five-node network (Figure 20 (a)), it is shown that, under any penetration rate, the performances of CAVs and HDVs are both improved, while the improvements for HDVs are generally higher. This indicates that there is "sacrifice" of CAVs in the control case (because their improvements are smaller than those of HDVs), whereas the sacrifice is not that dramatic. Although CAVs may sacrifice their instantaneous travel time at a specific time instant, the overall total travel time of CAVs can still be reduced due to the fact that CAVs can improve the capacities of the links. The actual travel times of CAVs might be reduced due to the increased capacities. For the Sioux Falls network (Figure 20 (b)), although the performances of HDVs are always improved under different CAV penetration rates for the case with CAV control, the performances of CAVs are degraded when the penetration is less than 90%. The network structure of Sioux Falls is much more complex than the five-node network. For CAVs, the benefits introduced by the increased capacities cannot overcome the sacrifice introduced by being assigned to the routes with longer travel times.



(a) Five-node Network (b) Sioux Falls Network

Figure 20. Performance improvements of HDVs and CAVs under HMPC compared with IDUE

The above results clearly show that, for the proposed scheme of CAV control, CAVs sacrifice themselves in general to improve the overall network performance, and thus may need to be compensated for their sacrifice. However, depending on the specific network structures or demand patterns of different networks, the level of sacrifice may be different, for which different compensation and incentive schemes may be designed to CAV passengers. This can be an important future research direction.

6. CONSLUSIONS AND FUTURE RESEARCH

This research proposed a double queue (DQ)-based mixed traffic flow model to describe the link dynamics and nodal flow transitions of the mixed HDV/CAV traffic flow. Based on the DQ model, we developed a dynamic bi-level framework to capture the behavior and interaction of HDVs and CAVs at the network level. The CAVs' route control problem is the upper level and the HDVs' route choice problem is the lower level. We developed mathematical models to describe the HDVs' and CAVs' route choice behavior by the IDUE principle and the DSO principle, respectively. An optimal control problem with equilibrium constraints (OCPEC) model was formulated to find the system performance oriented routes for CAVs in the network. We discretized the OCPEC to an MPEC and proposed the solution techniques. The nonlinear and non-smooth properties of the MPEC make it hard to be efficiently solved directly (e.g., by the relaxation method). To overcome this shortcoming, we developed a heuristic decomposition based MPC (HMPC) method by decomposing the original MPEC problem into two separate problems (i.e., the IDUE problem and the DSO problem).

Experiment results on a five-node network showed that the OCPEC model could improve the network performance compared with the scenario where all vehicles are HDVs, and the two solution methods can achieve similar results. Specifically, as the CAV penetration increases from 0% to 100%, the two solution methods reduce the network-wide total travel time dramatically. The maximum gain can achieve 37% under the studied scenario. However, for larger networks or longer time horizons, the relaxation method does not work well, while HMPC can still apply. We then tested the HMPC method on the five-node network with a longer time horizon and on the Sioux Falls network. The results further showed the effectiveness of the HMPC method.

There are several limitations of the proposed model and the HMPC solution method, which merit further investigations. First, we assume that CAVs always follow the assigned routes to improve the network performance. Although in theory this could be done by controlling the CAVs routes directly via communication and automation, doing so may be "unfair" to the CAV passengers. Therefore, this equity issue needs to be further investigated so that compensation and incentive schemes for CAV passengers can be developed to compensate the extra travel times (and/or other related costs) they may experience. To this end, what incentives and how to implement them should be carefully studied. Second, we assumed a constant shockwave speed for each link. Although the constant shockwave speed makes it easier to model and solve the studied problem, both shockwave speed and jam density may change in real-world scenarios and thus the constant shockwave speed may only represent a special (and simplified) case. Relaxing such constant shockwave speed to varying shockwave speed is one of the future research directions. Changing shockwave speeds will lead to a dynamic network model with time-varying, state-dependent delays. For such models, the proposed modeling framework in this paper still apply which however requires certain approximation schemes, e.g., those proposed in Ma et al. (2014) and Friesz and Mookherjee (2006), to generate and solve sub-problems similar to the network model in this paper. Deep learning and AI models may also hold great promise in developing such approximation schemes (Song et al., 2017), which is worthwhile for future investigations. Third, the possible FIFO violation issues of the DQ-based traffic dynamics model needs to be further studied. For this, one may need to design improved traffic flow models that guarantee the condition in Proposition 1 directly or design the control scheme of CAVs (e.g., by adding constraints) so that the increase of the CAV penetration on a link is not dramatic so that the FIFO condition in Proposition 1 can be satisfied. Fourth, we assume full observability of the

network information. In real-world traffic, such an assumption may not hold, and only limited information can be accessed. Developing models and algorithms that can work under partial observability is an interesting topic. Fifth, this paper focuses on network level flow control, while neglecting the control and optimization of individual vehicles such as vehicle trajectory and motion planning. It is an interesting future research direction to integrate network level flow control and micro-level individual vehicle control in one modeling framework. Last but not least, the proposed OCPEC model and the HMPC solution method need to be further tested and validated on larger, real-world transportation networks. The authors plan to work on these research topics and results may be reported in subsequent papers.

ACKNOWLEDGEMENT

The authors thank the four anonymous reviewers for their constructive comments that helped significantly improve the quality of earlier versions of the manuscript. This research is partially supported by the National Science Foundation under grant CMMI-1825053. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Bahrami, S. and Roorda, M.J., 2020. Optimal traffic management policies for mixed human and automated traffic flows. *Transportation Research Part A: Policy and Practice*, 135, pp.130-143.

Bagloee, S.A., Tavana, M., Asadi, M. and Oliver, T., 2016. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of modern transportation*, 24(4), pp.284-303.

Ban, X.J. and Liu, H.X., 2009. A link-node discrete-time dynamic second best toll pricing model with a relaxation solution algorithm. *Networks and Spatial Economics*, *9*(2), p.243.

Ban, J.X., Liu, H.X., Ferris, M.C. and Ran, B., 2006. A general MPCC model and its solution algorithm for continuous network design problem. *Mathematical and Computer Modelling*, *43*(5-6), pp.493-505.

Ban, X.J., Pang, J.S., Liu, H.X. and Ma, R., 2012a. Modeling and solving continuous-time instantaneous dynamic user equilibria: A differential complementarity systems approach. *Transportation Research Part B: Methodological*, *46*(3), pp.389-408.

Ban, X.J., Pang, J.S., Liu, H.X. and Ma, R., 2012b. Continuous-time point-queue models in dynamic network loading. *Transportation Research Part B: Methodological*, *46*(3), pp.360-380.

Blumberg, M. and Bar-Gera, H., 2009. Consistent node arrival order in dynamic network loading models. Transportation Research Part B: Methodological, 43(3), pp.285-300.

Cui, S., Seibold, B., Stern, R. and Work, D.B., 2017, June. Stabilizing traffic flow via a single autonomous vehicle: Possibilities and limitations. In *2017 IEEE Intelligent Vehicles Symposium (IV)* (pp. 1336-1341). IEEE.

Daganzo, C., 1992. The cell transmission model. Part I: A simple dynamic representation of highway traffic.

Daganzo, C.F., 1995. Properties of link travel time functions under dynamic loads. *Transportation Research Part B: Methodological*, *29*(2), pp.95-98.

Dresner, K. and Stone, P., 2004, July. Multiagent traffic management: A reservation-based intersection control mechanism. In Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2 (pp. 530-537).

Friesz, T.L. and Mookherjee, R., 2006. Solving the dynamic network user equilibrium problem with state-dependent time shifts. Transportation Research Part B: Methodological, 40(3), pp.207-229.

Friesz TL, Bernstein D, Suo Z, Tobin R (2001) Dynamic network user equilibrium with state-dependent time lags. Networks and Spatial Economics 1(3-4), 319–347.

Friesz TL, Kim T, Kwon C, Rigdon M (2010) Approximate network loading and dual-time-scale dynamic user equilibrium. Transportation Research Part B 45(1), 176–207

Gipps, P.G., 1981. A behavioural car-following model for computer simulation. Transport. Res. Part B: Methodol. 15 (2), 105–111.

Guo, L. and Ye, J.J., 2016. Necessary optimality conditions for optimal control problems with equilibrium constraints. *SIAM Journal on Control and Optimization*, *54*(5), pp.2710-2733.

Heydecker, B.G., Verlander, N., 1999. Calculation of dynamic traffic equilibrium assignments. In: Proceedings of the European Transport Conferences, Seminar F, pp. 79–91.

Jin, W.L. and Zhang, H.M., 2003. On the distribution schemes for determining flows through a merge. *Transportation Research Part B: Methodological*, 37(6), pp.521-540.
Krauss, S., 1998. *Microscopic modeling of traffic flow: Investigation of collision free vehicle dynamics* (Doctoral dissertation).
Lazar, D.A., Coogan, S. and Pedarsani, R., 2017, December. Capacity modeling and routing for traffic networks with mixed autonomy. In 2017 IEEE 56th Annual Conference on Decision and Control (CDC) (pp. 5678-5683). IEEE.
Leblanc, L.J., 1975. An algorithm for the discrete network design problem. Transportation Science, 9(3), pp.183-199.
Levin, M.W., 2017. Congestion-aware system optimal route choice for shared autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *82*, pp.229-247.
Levin, M.W. and Boyles, S.D., 2015. Intersection auctions and reservation-based control in dynamic traffic assignment. Transportation Research Record, 2497(1), pp.35-44.
Levin, M.W. and Boyles, S.D., 2016. A multiclass cell transmission model for shared human and autonomous vehicle roads. *Transportation Research Part C: Emerging Technologies*, 62, pp.103-116.
Li, T., Guo, F., Krishnan, R., Sivakumar, A., Polak, J., 2020. Right-of-way reallocation for mixed flow of autonomous vehicles and human driven vehicles. *Transp. Res. Part C Emerg. Technol*. 115, 102630. doi:10.1016/j.trc.2020.102630
Liu, H., Kan, X. (David), Shladover, S.E., Lu, X.Y., Ferlis, R.E., 2018. Modeling impacts of Cooperative Adaptive Cruise Control on mixed traffic flow in multi-lane freeway facilities. *Transp. Res. Part C Emerg. Technol*. 95, 261–279. doi:10.1016/j.trc.2018.07.027
Luo, Z.Q., Pang, J.S. and Ralph, D., 1996. *Mathematical programs with equilibrium constraints*. Cambridge University Press.
Ma, R., ***Ban, X***., Pang, J.S., Liu, X., 2015. Approximating time delays in solving continuous-time dynamic user equilibria. *Networks and Spatial Economics* 15(3), 443-463.
Ma, R., Ban, X.J. and Pang, J.S., 2014. Continuous-time dynamic system optimum for single-destination traffic networks with queue spillbacks. *Transportation Research Part B: Methodological*, *68*, pp.98-122.
Ma, R., Ban, X., Pang, J.S., 2018. A Link-Based Dynamic Complementarity System Formulation for Continuous-time Dynamic User Equilibria with Queue Spillbacks. Transportation Science 52(3), 564-592.
Merchant, D.K. and Nemhauser, G.L., 1978. A model and an algorithm for the dynamic traffic assignment problems. *Transportation science*, *12*(3), pp.183-199.
Milanés, V., Shladover, S.E., 2014. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. Transport. Res. Part C: Emerg. Technol. 48, 285–300.
Newell, G.F., 2002. A simplified car-following theory: a lower order model. Transport. Res. Part B: Methodol. 36 (3), 195–205.
Nie, Y., 2011. A cell-based Merchant–Nemhauser model for the system optimum dynamic traffic assignment problem. *Transportation Research Part B* 45 (2), 329–342.
Nie, X. and Zhang, H.M., 2005. A comparative study of some macroscopic link models used in dynamic traffic assignment. *Networks and Spatial Economics*, *5*(1), pp.89-115.
Osorio, C., Flötteröd, G. and Bierlaire, M., 2011. Dynamic network loading: a stochastic differentiable model that derives link state distributions. *Procedia-Social and Behavioral Sciences*, *17*, pp.364-381.
Patel, R., Levin, M.W. and Boyles, S.D., 2016. Effects of autonomous vehicle behavior on arterial and freeway networks. Transportation Research Record, 2561(1), pp.9-17.
Peeta, S. and Ziliaskopoulos, A.K., 2001. Foundations of dynamic traffic assignment: The past, the present and the future. Networks and spatial economics, 1(3-4), pp.233-265.
Ralph, D. and Wright, S.J., 2004. Some properties of regularization and penalization schemes for MPECs. *Optimization Methods and Software*, *19*(5), pp.527-556.
Ran, B., Boyce, D. (1996) Modeling Dynamic Transportation Networks: An Intelligent Transportation Systems Oriented Approach. *Springer-Verlag*, New York.
SAE international, 2016. Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. *SAE International,(J3016)*.
Shen, W., Zhang, H.M., 2008. What do different traffic flow models mean for system-optimal dynamic traffic assignment in a many-to-one network? *Transportation Research Record* 2088, 157–166.
Song, W., Han, K., Wang, Y., Friesz, T. and Del Castillo, E., 2017. Statistical metamodeling of dynamic network loading. Transportation research procedia, 23, pp.263-282.

Treiber, M., Hennecke, A., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E 62 (2), 1805–1824.
Vickrey, W.S., 1969. Congestion theory and transport investment. *The American Economic Review*, *59*(2), pp.251-260.
Ye, L., Yamamoto, T., 2018. Modeling connected and autonomous vehicles in heterogeneous traffic flow. *Phys. A Stat. Mech. its Appl*. 490, 269–277. doi:10.1016/j.physa.2017.08.015
Zheng, F., Jabari, S.E., Liu, H.X., Lin, D., 2018. Traffic state estimation using stochastic Lagrangian dynamics. Transportation Res. Part B: Methodological 115, 143–165.
Zheng, F., Liu, C., Liu, X., Jabari, S.E., Lu, L., 2020. Analyzing the impact of automated vehicles on uncertainty and stability of the mixed traffic flow. Transp. Res. Part C Emerg. Technol. 112, 203–219. doi:10.1016/j.trc.2020.01.017.
Zhou, Y., Ahn, S., Wang, M., Hoogendoorn, S., 2020. Stabilizing mixed vehicular platoons with connected automated vehicles: An H-infinity approach. Transp. Res. Part B Methodol. 132, 152–170. doi:10.1016/j.trb.2019.06.005.
Ziliaskopoulos, A.K., Peeta, S., 2000. A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation science*, *34*(1), pp.37-49.

APPENDIX A: NOTATION LIST

Network notation	
$\mathcal{N}$	Set of nodes
$\mathcal{L}$	Set of links
I_j	Incoming nodes that are directly linked to node j
K_j	Outgoing nodes that are directly linked from node j
$\bar{C}_{i,j}^{hdv}$	Capacity of link (i,j) when all vehicles are HDVs
$\bar{C}_{i,j}^{cav}$	Capacity of link (i,j) when all vehicles are CAVs
$\bar{Q}_{i,j}$	Queue storage capacity of link (i,j)
α, β	The parameter for the modified point queue model
$\tau_{i,j}^{0}$	The free flow travel time of link (i,j)
$\tau_{i,j}^{w}$	The shockwave travel time of link (i,j)
ΔT	The prediction time horizon of the predictive controller
Time-dependent variables for network dynamics (continuous-time)	
$\bar{C}_{i,j}(t)$	Capacity of link (i,j) at time t
$\bar{C}_{i,j}^{s}(t)$	Sending flow capacity of link (i,j) at time t
$\bar{C}_{i,j}^{r}(t)$	Receiving flow capacity of link (i,j) at time t
$Q_{i,j}(t)$	Upstream queue capacity of link (i,j) at time t
$d_i^{cav}(t)$	Demand of CAV from node i to destination at time t
$d_i^{hcv}(t)$	Demand of HDV from node i to destination at time t
$n_{i,j}(t)$	Number of all vehicles of link (i,j) at time t
$n_{i,j}^{hdv}(t)$	Number of HDVs of link (i,j) at time t
$n_{i,j}^{cav}(t)$	Number of CAVs of link (i,j) at time t
$p_{i,j}(t)$	Total inflow rate to ink (i,j) at time t
$p_{i,j}^{hdv}(t)$	Inflow rate of HDVs to link (i,j) at time t
$p_{i,j}^{cav}(t)$	Inflow rate of CAVs to link (i,j) at time t
$\hat{p}_{i,j}^{hdv}(t)$	Pseudo inflow of HDVs to link (i,j) at time t, calculated by IDUE
$p_{j,k}^{PQ}(t)$	Inflow rate to ink (i,j) at time t when modeled by point queue
$q_{i,j}^{U}(t)$	Upstream queue length of link (i,j) at time t
$q_{i,j}^{D}(t)$	Downstream queue length of link (i,j) at time t
$q_{i,j}^{D,hdv}(t)$	Number of downstream HDVs on link (i,j) at time t
$q_{i,j}^{D,cav}(t)$	Number of downstream CAVs on link (i,j) at time t
$v_{i,j}(t)$	Total exit flow rate from link (i,j) at time t
$v_{i,j}^{hdv}(t)$	Exit flow rate of HDVs from link (i,j) at time t
$v_{i,j}^{cav}(t)$	Exit flow rate of CAVs from link (i,j) at time t
$v_{j,k}^{PQ}(t)$	Exit flow rate from link (i,j) at time t when modeled by point queue
$\tau_{i,j}(t)$	Travel time of link (i,j) at time t

$\eta_i^s(t)$	The instantaneous minimum travel time from link i to destination s at time t
$\rho_{i,j}(t)$	CAV penetration of link (j,k) at time t
$\Omega(t)$	The information of the whole network at time t
$G_{(i,j)\to(j,k)}(t)$	Total transition flow from link (i,j) to link (j,k) at time t
$G^{hdv}_{(i,j)\to(j,k)}(t)$	HDVs transition flow from link (i,j) to link (j,k) at time t
$G^{cav}_{(i,j)\to(j,k)}(t)$	CAVs transition flow from link (i,j) to link (j,k) at time t
$R_{i,j}(t)$	Receiving flow of link (i,j) at time t
$S_{i,j}(t)$	Sending flow from link (i,j) at time t
$S^{hdv}_{i,j}(t)$	Sending flow of HDVs from link (i,j) at time t
$S^{cav}_{i,j}(t)$	Sending flow of CAVs from link (i,j) at time t
Time-dependent variables for network dynamics (discrete-time)	
h	Time step of the discretization
$t_i^h, i = 0,1,2 \dots N_h$	Discretized time
$\bar{C}^{h,r}_{i,j}$	Capacity of link (i,j) at discrete time r
$\bar{Q}^{h,r}_{i,j}$	Upstream queue capacity of link (i,j) at discrete time r
$d_i^{cav,h,r}$	Demand of CAV from node i to destination at discrete time r
$d_i^{hdv,h,r}$	Demand of HDV from node i to destination at discrete time r
$n^{h,r}_{i,j}$	Number of all vehicles of link (i,j) at discrete time r
$n^{hdv,h,r}_{i,j}$	Number of HDVs of link (i,j) at discrete time r
$n^{cav,h,r}_{i,j}$	Number of CAVs of link (i,j) at discrete time r
$p^{h,r}_{i,j}$	Total inflow rate to ink (i,j) at discrete time r
$p^{hdv,h,r}_{i,j}$	Inflow rate of HDVs to link (i,j) at discrete time r
$p^{cav,h,r}_{i,j}$	Inflow rate of CAVs to link (i,j) at discrete time r
$\hat{p}^{hdv,h,r}_{j,k}$	Pseudo inflow of HDVs to link (i,j) at discrete time r calculated by IDUE
$q^{U,h,r}_{i,j}$	Upstream queue length of link (i,j) at discrete time r
$q^{D,h,r}_{i,j}$	Downstream queue length of link (i,j) at discrete time r
$q^{D,hdv,h,r}_{i,j}$	Number of downstream HDVs on link (i,j) at discrete time r
$q^{D,cav,h,r}_{i,j}$	Number of downstream CAVs on link (i,j) at discrete time r
$v^{h,r}_{i,j}$	Total exit flow rate from link (i,j) at discrete time r
$v^{hdv,h,r}_{i,j}$	Exit flow rate of HDVs from link (i,j) at discrete time r
$p^{cav,h,r}_{i,j}$	Exit flow rate of CAVs from link (i,j) at discrete time r
$\tau^{h,r}_{i,j}$	Travel time of link (i,j) at discrete time r
$\eta_i^{s,h,r}$	The instantaneous minimum travel time from link i to destination s at discrete time r

$\rho_{i,j}^{h,r}$	CAV penetration of link (j,k) at discrete time r
$G_{(i,j)\to(j,k)}^{h,r}$	Total transition flow from link (i,j) to link (j,k) at discrete time r
$G_{(i,j)\to(j,k)}^{hdv,h,r}$	HDVs transition flow from link (i,j) to link (j,k) at discrete time r
$G_{(i,j)\to(j,k)}^{cav,h,r}$	CAVs transition flow from link (i,j) to link (j,k) at discrete time r
$R_{i,j}^{h,r}$	Receiving flow of link (i,j) at discrete time r
$S_{i,j}^{h,r}$	Sending flow from link (i,j) at discrete time r
$S_{i,j}^{hdv,h,r}$	Sending flow of HDVs from link (i,j) at discrete time r
$S_{i,j}^{cav,h,r}$	Sending flow of CAVs from link (i,j) at discrete time r

APPENDIX B

Lemma B1: Equation (10) in DQ is equivalent to Equation (9) in LTM.

Proof: Here we show Equation (7) in DQ can be derived from Equation (6) in LTM.
First, divide both sides in Equation (9) by Δt, we have

$$S_{i,j}(t) = \min\left(\frac{N\left(x_{i,j}^{in}, t+\Delta t-\tau_{i,j}^{0}\right) - N(x_{i,j}^{out}, t)}{\Delta t}, \quad \overline{C}_{i,j}^{s}(t)\right) \tag{53}$$

Notice that $q_{i,j}^{D}(t) = N\left(x_{i,j}^{in}, t-\tau_{i,j}^{0}\right) - N(x_{i,j}^{out}, t)$. If $q_{i,j}^{D}(t) = 0$, i.e., $N\left(x_{i,j}^{in}, t-\tau_{i,j}^{0}\right) = N(x_{i,j}^{out}, t)$, let $\Delta t \to 0$, we have

$$\lim_{\Delta t\to 0} \frac{N\left(x_{i,j}^{in}, t+\Delta t-\tau_{i,j}^{0}\right) - N\left(x_{i,j}^{in}, t-\tau_{i,j}^{0}\right)}{\Delta t} = p_{i,j}\left(t-\tau_{i,j}^{0}\right) \tag{54}$$

This means

$$S_{i,j}(t) = \min\left(p_{i,j}\left(t-\tau_{i,j}^{0}\right), \quad \overline{C}_{i,j}^{s}(t)\right), \quad if\ q_{i,j}^{D}(t) = 0 \tag{55}$$

If $q_{i,j}^{D}(t) > 0$, we have

$$\begin{aligned} &\lim_{\Delta t\to 0} \frac{N\left(x_{i,j}^{in}, t+\Delta t-\tau_{i,j}^{0}\right) - N(x_{i,j}^{out}, t)}{\Delta t} \\ &= \lim_{\Delta t\to 0} \frac{N\left(x_{i,j}^{in}, t+\Delta t-\tau_{i,j}^{0}\right) - N\left(x_{i,j}^{in}, t-\tau_{i,j}^{0}\right) + q_{i,j}^{D}(t)}{\Delta t} \\ &= p_{i,j}\left(t-\tau_{i,j}^{0}\right) + \lim_{\Delta t\to 0} \frac{q_{i,j}^{D}(t)}{\Delta t} = +\infty \end{aligned} \tag{56}$$

Thus,

$$S_{i,j}(t) = \min\left(+\infty, \quad \overline{C}_{i,j}^{s}(t)\right) = \overline{C}_{i,j}^{s}(t), \quad if\ q_{i,j}^{D}(t) > 0 \tag{57}$$

This concludes the proof.

APPENDIX C

Algorithm-1: The HDV routes adjustor and CAV routes generator

Calculate the receiving capacity $R_{j,k}(t)$ of each outgoing link $(j,k),\ k \in K_j$

I. Adjust HDV routes, assign CAV routes, check receiving capacity

If $\hat{p}_{j,k^{hdv}}^{hdv}(t) \geq R_{j,k^{hdv}}(t)$

Set $\hat{p}_{j,k^{hdv}}^{hdv}(t) = R_{j,k^{hdv}}(t)$

If $p_{j,k^{hdv}}^{PQ}(t) \geq \hat{p}_{j,k^{hdv}}^{hdv}(t)$

Set $\hat{p}_{j,k^{hdv}}^{cav}(t) = \min\left\{p_{j,k^{hdv}}^{PQ}(t) - \hat{p}_{j,k^{hdv}}^{hdv}(t), \quad R_{j,k^{hdv}}(t) - \hat{p}_{j,k^{hdv}}^{hdv}(t)\right\}$

Set $\hat{p}_{j,k}^{cav}(t) = \min\left\{p_{j,k}^{PQ}(t), \quad R_{j,k^{hdv}}(t)\right\}$ for $k \in K_j$ and $k \neq k^{hdv}$

Else

Set $\hat{p}_{j,k^{hdv}}^{cav}(t) = 0$

Set $\hat{p}_{j,k}^{cav}(t) = \min\left\{\left[\sum_{k' \in K_j} p_{j,k'}^{PQ}(t) - \hat{p}_{j,k^{hdv}}^{hdv}(t)\right]\frac{p_{j,k}^{PQ}(t)}{\sum_{k' \in K_j, k' \neq k^{hdv}} p_{j,k'}^{PQ}(t)}, \quad R_{j,k^{hdv}}(t)\right\}$ for $k' \in K_j$ and $k' \neq k^{hdv}$

Calculate total assigned flows for both HDVs and CAVs: $P_j^{hdv} = \sum_{k^{hdv}} \hat{p}_{j,k^{hdv}}^{hdv}(t)$, $P_j^{cav} = \sum_k \hat{p}_{j,k}^{cav}(t)$

II. Calculate the proportion of HDVs and CAVs of the sending flows

Initialize $\hat{S}_j^{hdv}(t) = \hat{S}_{j,hdv}^{cav}(t) = \hat{S}_{j,only}^{cav}(t) = 0$.

For $i \in I_j$

If $\hat{S}_{(i,j)}^{hdv}(t) > 0$

$\hat{S}_j^{hdv}(t)$+=$\hat{S}_{(i,j)}^{hdv}(t)$; $\hat{S}_{j,hdv}^{cav}(t)$+=$\hat{S}_{(i,j)}^{cav}(t)$

Else

$\hat{S}_{j,only}^{cav}$+=$\hat{S}_{(i,j)}^{cav}(t)$

III. *Modify HDV and CAV routes to satisfy the DQ nodal model*)

If there is a k^{hdv} such that $\hat{p}_{j,k^{hdv}}^{hdv}(t) > 0$

Calculate the proportion of CAVs to HDVs of the coupled sending flow: $\rho = \frac{\hat{S}_{j,hdv}^{cav}(t)}{\hat{S}_j^{hdv}(t)}$

If $\sum_k \hat{p}_{j,k}^{cav}(t) > \rho \hat{p}_{j,k^{hdv}}^{hdv}(t)$

For $k \in K_j$

$p_{j,k}^{hdv}(t) = \hat{p}_{j,k}^{hdv}(t)$ and $p_{j,k}^{cav}(t) = \hat{p}_{j,k}^{cav}(t)$

For $i \in I_j$

If $S_{i,j}^{hdv}(t) > 0$

$v_{i,j}^{hdv}(t) = \hat{p}_{j,k^{hdv}}^{hdv}(t)\frac{S_{i,j}^{hdv}(t)}{\hat{S}_j^{hdv}(t)}$ and $v_{i,j}^{cav}(t) = \hat{p}_{j,k^{hdv}}^{hdv}(t)\frac{S_{i,j}^{cav}(t)}{\hat{S}_j^{hdv}(t)}$

Else

$v_{i,j}^{hdv}(t) = 0$ and $v_{(i,j)}^{cav}(t) = S_{i,j}^{cav}(t)\frac{\sum_k \hat{p}_{j,k}^{cav}(t) - \hat{S}_{j,hdv}^{cav}(t)}{\hat{S}_{j,only}^{cav}(t)}$

Else

For $k \in K_j$

If $k = k^{hdv}$

$p_{j,k^{hdv}}^{hdv}(t) = \left(\sum_k \hat{p}_{j,k}^{cav}(t) + \hat{p}_{j,k^{hdv}}^{hdv}(t)\right)\frac{\hat{S}_j^{hdv}(t)}{\hat{S}_j^{hdv}(t) + \hat{S}_{j,hdv}^{cav}(t) + \hat{S}_{j,only}^{cav}(t)}$ and $p_{j,k}^{cav}(t) = \hat{p}_{j,k}^{cav}(t) + \hat{p}_{j,k}^{hdv}(t) - p_{j,k}^{hdv}(t)$

Else

$p_{j,k}^{hdv}(t) = 0$ and $p_{j,k}^{cav}(t) = \hat{p}_{j,k}^{cav}(t)$

For $i \in I_j$

If $S_{i,j}^{hdv}(t) > 0$

$v_{i,j}^{hdv}(t) = p_{j,k}^{hdv}(t)\frac{S_{i,j}^{hdv}(t)}{\hat{s}_j^{hdv}(t)}$ and $v_{i,j}^{cav}(t) = p_{j,k}^{hdv}(t)\frac{S_{i,j}^{cav}(t)}{\hat{s}_j^{hdv}(t)}$

Else

$v_{i,j}^{hdv}(t) = v_{i,j}^{cav}(t) = 0$

Else

For $k \in K_j$

$p_{j,k}^{cav}(t) = \hat{p}_{j,k}^{cav}(t)$ and $p_{j,k}^{hdv}(t) = 0$

For $i \in I_j$

If $S_{i,j}^{hdv}(t) > 0$

$v_{i,j}^{hdv}(t) = v_{i,j}^{cav}(t) = 0$

Else

$v_{i,j}^{hdv}(t) = 0$ and $v_{i,j}^{cav}(t) = \sum_k p_{j,k}^{cav}(t)\frac{S_{i,j}^{cav}(t)}{\hat{s}_{j,only}^{cav}(t)}$

End

Algorithm-2: The overall algorithm

For $t = 1{:}N$, do

Collect the real transition flow of HDVs and CAVs from previous step $p_{j,k}^*(t-1), v_{i,j}^*(t-1)$

Move the DQ link dynamics 1 step forward

Collect current network states $n_{i,j}^*(t)$, $q_{i,j}^{D,*}(t), q_{i,j}^U(t), S_{i,j}^*(t), R_{j,k}(t)$ and obtain the demands of HDVs and CAVs $d_j^*(t)$

Calculate the routes of HDVs by IDUE principle (i.e., Equation (22) and (23)) and get $\hat{p}_{j,k}^{hdv}(t)$

Calculate the expected routes of both HDVs and CAVs by solving the PQ-based DSO (Equation (49)-(52)) and get $p_{j,k}^{PQ}(t), v_{i,j}^{PQ}(t)$

Adjust HDV routes and generate CAV routes using Algorithm-1 and get $p_{j,k}^*(t), v_{i,j}^*(t)$

End

APPENDIX D: SIOUX FALLS NETWORK