
01 Jan 2021

Enhancing Mixed Traffic Flow Safety Via Connected and Autonomous Vehicle Trajectory Planning with a Reinforcement Learning Approach

Yanqiu Cheng

Chenxi Chen

XianBiao Hu

Missouri University of Science and Technology, xbhu@mst.edu

Kuanmin Chen

et. al. For a complete list of authors, see https://scholarsmine.mst.edu/civarc_enveng_facwork/2056

Follow this and additional works at: https://scholarsmine.mst.edu/civarc_enveng_facwork



Part of the [Civil Engineering Commons](#)

Recommended Citation

Y. Cheng et al., "Enhancing Mixed Traffic Flow Safety Via Connected and Autonomous Vehicle Trajectory Planning with a Reinforcement Learning Approach," *Journal of Advanced Transportation*, vol. 2021, article no. 6117890, Hindawi, Jan 2021.

The definitive version is available at <https://doi.org/10.1155/2021/6117890>



This work is licensed under a [Creative Commons Attribution 4.0 License](#).

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Civil, Architectural and Environmental Engineering Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

Research Article

Enhancing Mixed Traffic Flow Safety via Connected and Autonomous Vehicle Trajectory Planning with a Reinforcement Learning Approach

Yanqiu Cheng,^{1,2} Chenxi Chen,² Xianbiao Hu ,² Kuanmin Chen,¹ Qing Tang,² and Yang Song²

¹Department of Traffic Engineering, College of Transportation Engineering, Chang'an University, Xi'an 710064, Shaanxi, China

²Department of Civil, Architectural and Environmental Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA

Correspondence should be addressed to Xianbiao Hu; xbhu@mst.edu

Received 16 April 2021; Revised 23 May 2021; Accepted 4 June 2021; Published 14 June 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Yanqiu Cheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The longitudinal trajectory planning of connected and autonomous vehicle (CAV) has been widely studied in the literature to reduce travel time or fuel consumptions. The safety impact of CAV trajectory planning to the mixed traffic flow with both CAV and human-driven vehicle (HDV), however, is not well understood yet. This study presents a reinforcement learning modeling approach, named Monte Carlo tree search-based autonomous vehicle safety algorithm, or MCTS-AVS, to optimize the safety of mixed traffic flow, on a one-lane roadway with signalized intersection control. Crash potential index (CPI) is defined to quantitatively measure the safety performance of the mixed traffic flow. The CAV trajectory planning problem is firstly formulated as an optimization model; then, the solution procedure based on reinforcement learning is proposed. The tree-expansion determination module and rollout termination module are developed to identify and reduce the unnecessary tree expansion, so as to train the model more efficiently towards the desired direction. The case study results showed that the proposed algorithm was able to reduce the CPI by 76.56%, when compared with a benchmark model without any intelligence, and 12.08%, when compared with another benchmark model that the team developed earlier. These results demonstrated the satisfactory performance of the proposed algorithm in enhancing the safety of the mixed traffic flow.

1. Introduction

Connected and automated vehicles (CAVs) have been demonstrated to have great potentials for future transportation systems [1–4]. Compared with human-driven vehicles (HDVs), CAVs behave accurately as they are controlled by the computer algorithms, and their trajectories can be adjusted with predefined intelligence to achieve objectives such as minimizing delays and/or fuel consumptions at roadway intersections. This process is named longitudinal trajectory planning and is an important task to realize the full potentials of CAVs. Data from on-board equipment (e.g., in-vehicle sensors, radar, camera, and lidar)

and remote facilities (e.g., DSRC/Cellular, GNSS/IMU, and priori map) can be utilized to schedule CAV trajectory [5].

Plenty of studies on CAV longitudinal trajectory planning have been conducted. For example, Chen et al. [6] proposed a centralized control method for CAVs by using a cost function which included CAV safety, efficiency, and ride comfort as the minimization objective. The robust platooning was formulated as a Min-Max Model Predictive Control (MM-MPC) problem, where optimal accelerations were generated to minimize this cost function. Wu et al. [7] presented an optimal longitudinal control strategy for a homogeneous CAV platoon. A linear-quadratic optimal controller was designed considering a comprehensive

perspective, including driving safety, efficiency, and ride comfort, with three performance indicators including vehicle gap error, relative speed, and desired acceleration. Malikopoulos et al. [8] provided a decentralized theoretical framework for coordination of CAVs. Rear-end, speed-dependent safety constraint had been taken into account. Research studies with similar objectives can also be found in [9–15].

While significant progress on CAV longitudinal trajectory planning can be observed in the abovementioned literature, one thing that is largely missing is the impact of CAV longitudinal trajectory planning algorithms to the safety of traffic flow, and subsequently, how should we design CAV longitudinal trajectory planning algorithms to minimize the probability of crash occurrence. To clarify, in most abovementioned works, the objective of trajectory planning is usually to reduce travel time or fuel consumption, and CAV safety is usually built in the model as a constraint, rather than an objective. In addition, the consideration of driving safety is usually limited to the CAV itself, instead of the other HDVs in the traffic flow. However, learning from the driving safety and human behavior research, traffic crash happens most frequently when the vehicles are changing speed, e.g., accelerating or decelerating at intersections. In a mixed traffic flow environment with both CAV and HDV, the CAV control algorithm will not only impact the movement of the CAV, but, through traffic flow shockwave propagation, will also influence the driving behavior of the HDVs at upstream locations. As such, it should be noted that the safety impact of CAV is not only limited to the CAV itself but also to the surrounding HDVs as well, and a good longitudinal trajectory planning algorithm needs to consider all of these and aims to minimize the crash potential of the entire traffic flow.

Methodologically, CAV trajectory scheduling is still a sophisticated problem, considering the great challenges from the highly stochastic nature of human driving behaviors and almost infinite decision-making states in real-world mixed traffic context. One common and effective approach to simplify the above complicate problem is to divide a vehicle trajectory into several segments. In other words, vehicles are usually set to the same cruising speed, or with constant acceleration/deceleration, at each stage. For example, He et al. [16] proposed a multistage approximation control model to solve the optimal trajectory problem. First, the vehicle cruised at the speed calculated by their algorithm and then accelerated/decelerated to a final speed when passing through the intersection. Wu et al. [17] divided the whole vehicle control process into a sequence of control stages and each control stage was formulated as an individual optimal control problem involving spatial and temporal constraints induced by the presence of vehicle queues. In [18], the vehicle was supposed to accelerate to different optimal cruising speeds by few speed guidance which also divided the roadway. In [19], the roadway was separated into three segments by two individual variable speed limits (IVSL). After those IVSLs, vehicle speed was adjusted to a final constant value so that their trajectories are smooth.

Similar method can be found in [20, 21], in which each vehicle trajectory was broken into a few sections to decompose the originally hard trajectory design problem to a simple one. Although the abovementioned approach does make the model analytically solvable and help reduce the computational burden, such assumptions sacrificed the modeling realism and were not flexible to account for the uncertainty of human driver behaviors in a mixed traffic environment.

Considering the modeling techniques of trajectory planning, the computational complexity and algorithm runtime are directly related to modeling realism and the market penetration rate (MPR) of CAV. One way to reduce the complexity of the model is to only consider the pure CAV traffic, i.e., a traffic environment without any HDVs. In fact, large amounts of research studies on CAV trajectory planning were under this assumption. For example, Lee and Park [22] developed a CVIC algorithm for manipulating individual automated vehicle into crossing an intersection without colliding with other vehicles in a 100% MPR AVs environment. Wang et al. [13] proposed a rolling horizon control framework to control all vehicles' trajectory, which were equipped with driver assistance systems by optimizing a cost function reflecting different control objectives. Under the same assumption, Ahn et al. [23] developed an eco-drive system that combines an eco-cruise control algorithm and state-of-the-art car-following models. Zhou et al. [15] proposed a reinforcement learning-based approach to train a CAV platoon to pass through the intersection with a steady speed. The same research context can be found in [24–29]. In the above research studies, although it was able to simplify the model and improve calculation efficiency under the pure CAV environment, the applicability of the models was greatly reduced.

To deal with the abovementioned issues, this study proposes a model-free trajectory planning approach for improving the safety of mixed traffic flow of HDV and CAV, named Monte Carlo tree search-based autonomous vehicle safety algorithm, or MCTS-AVS. We quantify the safety level of the mixed traffic flow by using crash potential index (CPI) as the minimization objectives. The CAV trajectory planning problem is firstly formulated as an optimization model, and then, a solution procedure based on reinforcement learning is proposed. The tree-expansion determination module and rollout termination module are developed to identify and reduce the unnecessary tree expansion, so as to train the model more efficiently towards the desired direction. These modeling efforts lead to the improvement of algorithm solution quality and safety performance. Finally, the proposed algorithm was implemented and tested in a one-lane roadway with signalized intersection control.

2. Notations

As a convenient reference, the mathematical notations used in this section are presented below.

- t, T : discrete time step, and the time horizon
 $s(t)$: state at time t
 $a_{CAV}(t), \mathcal{A}_{s(t)}$: action of CAV, and set of all actions at state s and time t
 $y_i(t)$: the distance of vehicles i from roadway entrance, at time t
 $Y(t)$: an array that stores vehicles' distance from roadway entrance, at time t
 $v_i(t)$: the speed of vehicle i , at time t
 $V(t)$: an array that stores vehicles' speed, at time t
 $d_{hi}(t)$: the distance headway of vehicle i , at time t
 $D(t)$: an array that stores vehicles' distance headway, at time t
 t_g, t_y, t_r, t_c : durations of green, yellow, and red signals, and cycle length
 Δt : the shortest time interval
 v_s : speed limit
 l_{seg} : length of the roadway segment
 l_a : average vehicle length

3. Methodology

3.1. Model Formulation

3.1.1. Problem Setting and Decomposition. We believe that, in the near future, the mixed traffic flow that composes of multiple HDVs and CAVs traveling on arterial segment will be a general scenario, as opposed to pure CAV traffic flow. This is because transitioning to fully CAV traffic might be a time-consuming process. It also implies that we will have a mixture of CAV and HDV in the mixed traffic flow, and the traffic dynamics become complex. To simplify the CAV control problem, this mixed traffic flow is firstly decomposed into several "basic interactive unit (BIU)," as illustrated in Figure 1. After the decomposition, each CAV is involved in one BIU, and the rest of the vehicles in the platoon are HDVs. In the Figure 1, the number of HDVs might be one or multiple, or there might be no CAV at all. As such, the mixed traffic flow problem can be converted into a trajectory optimization problem for each BIU, which significantly reduced the total computational complexity.

There are two reasons for such decomposition. First, if an HDV is driving in front of a CAV, due to the human nature, it will drive according to speed limit or prevailing cruising speed, and as a result, its behavior is not impacted by the CAV behind it. Second, CAV is subjected to the speed limit or current traffic conditions, as it cannot drive faster than a typical HDV. On the contrary, when it slows down to a speed that is lower than HDV, it becomes a moving bottleneck, and all HDVs behind it are forced to slow down and follow this CAV. To summarize, for mixed traffic flow control problem, we will always have a CAV that is leading the platoon and potentially multiple HDVs behind the CAV, in each basic interactive unit. Such decomposition is also frequently used in the previous literatures.

3.1.2. State Transition. To describe state transition, we use $s(t) = \{Y(t), V(t)\}$ to represent the state of mixed traffic flow at time, where $Y(t) = \{y_i(t), i = 1, 2, \dots, k\}$, $V(t) = \{v_i(t), i = 1, 2, \dots, k\}$, and k is the total number of vehicles. Then, $v_i(t)$ is updated by $v_i(t+1) = v_i(t) + a_i(t)$. CAV moves with action $a_{CAV}(t) \in \mathcal{A}_{s(t)}$ at time t . For the HDVs in the traffic flow, there are two distinct scenarios: (1) when HDVs are relatively far away from the intersection, their behaviors are mostly car-following (CF) and can be described by the classic CF model; (2) when HDVs are getting close to the intersection, the vehicle behaviors are subject to the signal lights. In other words, vehicles will drive through the intersection when the light is green or if they cannot come to a safe stop when the yellow light is on. Otherwise, it will slowdown and stop before the stop line. The HDVs behavior of these two scenarios are illustrated in Figures 2(a) and 2(b), and both of them follow the vehicle constraints, including collision avoidance and speed limit, as well as vehicle kinematics.

To describe the velocity decision-making of the HDVs for the first scenario, the general GM model considering stochastic HDVs behavior is employed. Compared with the classic intelligent driver model (IDM) which was introduced in [30], the GM model has the following advantages. (1) Human perception reaction time, speed difference, and space headway were involved in this model as a simple structure. It enables HDV trajectories' simulation rapidly but without losing too much detail. (2) A random term to reveal the uncertain factors of human drivers behavior was also been considered. This makes the model closer to the real scenario and a higher applicability. The specific formulation of GM model is shown as

$$a_i(t) = \alpha * v_i(t)^\beta \frac{\Delta v_i(t - t_{reaction})}{\Delta x_i(t - t_{reaction})^\gamma} + \epsilon_i(t), \quad (1)$$

where $a_i(t)$ is the acceleration value of the human drive vehicle i at time t , $v_i(t)$ is the vehicle's speed, $t_{reaction}$ denotes the human perception reaction time, $\Delta v_i(t - t_{reaction})$ is the speed difference between the target vehicle and its leading vehicle at time $(t - t_{reaction})$, $\Delta x_i(t - t_{reaction})$ is the space headway, α , β , and γ are the parameters to calibrate, and $\epsilon_i(t)$ is a random term associated with vehicle i at time t . Several researchers (e.g., [31]) calibrated these parameters with collected data in real world. After Δt , state transition $s(t) \rightarrow^{\Delta t} s(t+1)$ is realized by

$$y_i(t+1) = y_i(t) + v_i(t) + \frac{1}{2} a_i(t) t^2, \quad i = 1, 2, \dots, k. \quad (2)$$

3.1.3. Crash Potential Index Function. Considering the movement of the vehicles in the traffic, we divide the traffic flow states into two types to further evaluate the safety performance of the current state. In general, when the vehicle velocity is less than the rear vehicle, two vehicles tend to be close, and the traffic flow has potential crash risk. We define this kind of state as a crash potential state, as shown

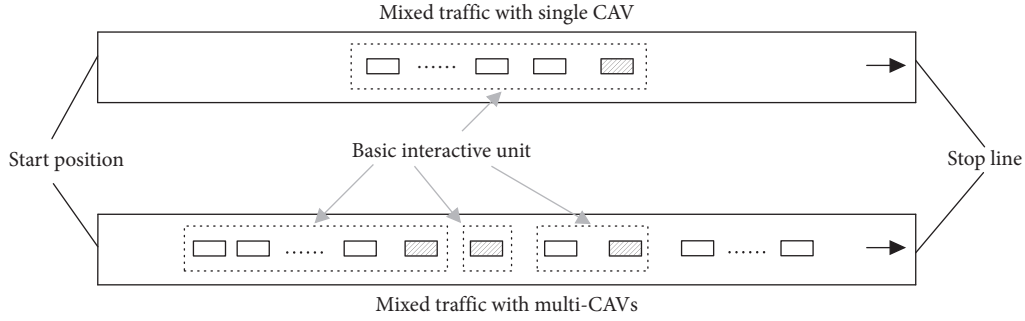


FIGURE 1: Decomposition of mixed traffic flow into basic interactive units.

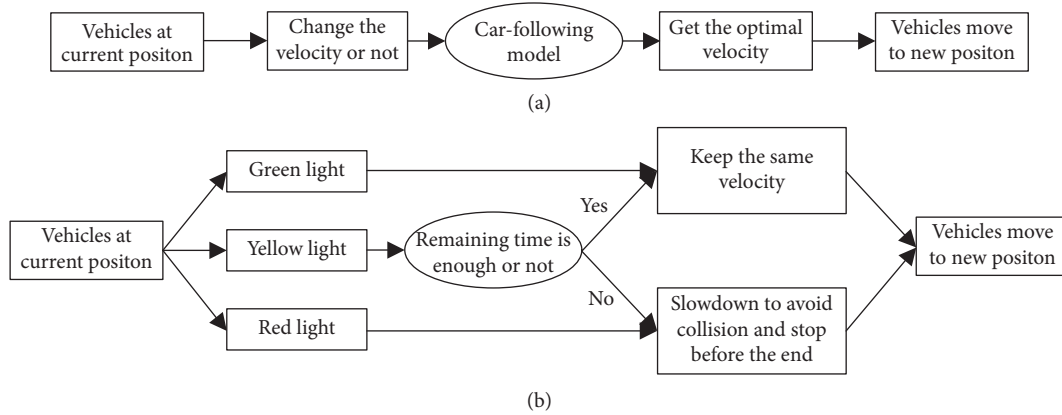


FIGURE 2: HDVs' behavior in roadway with signal control intersection. (a) HDVs are far from the signal light and (b) HDVs are close to the signal light.

on the left side of Figure 3. For example, when the signal light changes from green to yellow, the leading vehicle slows down and the traffic flow gets dense. On the contrary, when the vehicle velocity is greater than or equal to the rear vehicle, the distance headway will remain the same or increase, and there is less risk of collision in this traffic. This kind of state is defined as a safe state. For example, when the signal light changes from red to green, the leading vehicle begins to accelerate and the distance headway increases gradually, as shown on the right side of Figure 3.

To quantify the safety degree of a traffic flow, we defined a crash potential index (CPI) function as

$$X(t) = \begin{cases} \sum_{i=1}^k (v_{i+1}(t) - v_i(t)), & i = 1, 2, \dots, k, \text{ if } v_i(t) < v_{i+1}(t) \\ 0, & \text{else,} \end{cases} \quad (3)$$

where $X(t)$ is the CPI value of this traffic flow at time t and k is total number of vehicles. The cumulative value considers the above two states: the speed difference between two adjacent vehicles is calculated when they are close to each other or zero when the two adjacent vehicles are far away or

relatively slow. This value directly reflects the overall crash potential degree of the traffic flow.

3.1.4. Optimization Model. The overall optimization problem is represented by

$$\text{minimize } \sum_t^T X(t). \quad (4)$$

The feasible region for CAV action $a_{CAV}(t)$ at time t is subjected to

$$y_{CAV}(t+1) = y_{CAV}(t) + v_{CAV}(t) + \frac{1}{2} a_{CAV}(t) t^2, \quad t = 0, 1, \dots, T, \quad (5a)$$

$$v_{CAV}(t+1) = v_{CAV}(t) + a_{CAV}(t), \quad t = 0, 1, \dots, T, \quad (5b)$$

$$|a_{CAV}(t)| \leq A, \quad t = 0, 1, \dots, T, \quad (5c)$$

$$0 \leq v_{CAV}(t) \leq v_s, \quad t = 0, 1, \dots, T, \quad (5d)$$

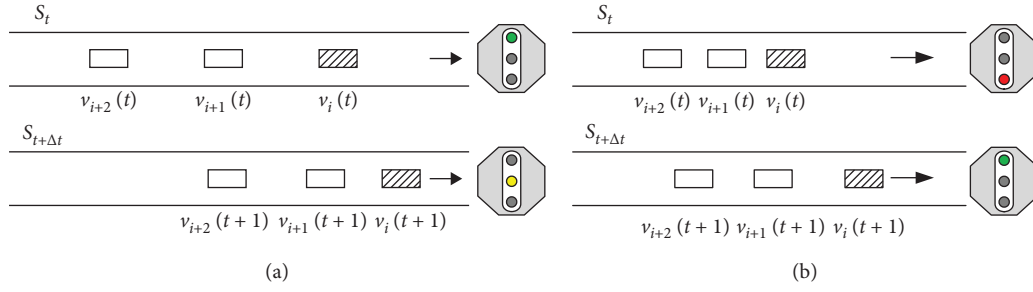


FIGURE 3: Two types state for a traffic flow: (a) crash potential state and (b) safe state.

where T is the time at the end of mixed traffic flow travel (e.g., get through an intersection). A is the upper limit of the absolute value of CAV acceleration.

3.2. Solution Algorithm

3.2.1. UCT Formulation. The problem in equations (4) and (5) is a challenging nonlinear program (NLP) with a huge state space, which makes the problem computationally intractable. This is because, at a given time t , the state of this problem is defined by a list of specific input features to describe the current system status and is required for any reinforcement learning algorithm. For a mixed traffic flow, many variables can be used to describe the state, for example, vehicle's distance from roadway entrance, vehicle velocity, accelerations, spacing/time headways between vehicles, elapsed time, and signal light color and their remaining duration. Obviously, when more features are selected, more details of the state will be captured. However, excessive number of state elements may directly lead to an exponential growth of the state space and lead to the "curse of dimensionality." As a result, a huge state space will come with a higher memory requirement and computational burden. Therefore, the features have to be chosen carefully.

In this study, we choose to use a combination of time, vehicle location, and vehicle speed to represent the time, in which the vehicle location and speed are two arrays that include information of all vehicles in the traffic flow. However, even with these 3 limited variables, once we discretize the time, space, and speed dimensions, this model becomes high-dimensional in state and is very challenging to solve and as such we have to rely on the reinforcement learning approach. In this study, we developed a heuristic algorithm, Monte Carlo tree search-based autonomous vehicle safety algorithm, or MCTS-AVS, to solve this problem by searching near-optimum action at every time step for CAV.

Typical MCTS algorithm consists of four steps: selection, expansion, simulation, and backpropagation [32, 33]. UCT algorithm (upper confidence bounds for trees) is employed to the first step of MCTS-AVS, as it can well balance the dilemma between exploration and exploitation part of a selection policy. The underlying mechanism for UCT, which is denoted by π_{UCT} , is described by the following formula:

$$\pi_{UCT} = \arg \max \left\{ Q''_{UCT}(s, a) = Q_{UCT}(s, a) + \mathbb{C} \sqrt{\frac{\ln(n(s))}{n(s, a)}}, a \in \mathbb{A} \right\}, \quad (6)$$

where π_{UCT} is the selected policy, s is system state, a is action, \mathbb{A} is the set for all actions, $n(s)$ is the total number of times a state s has been visited, $n(s, a)$ is the number of times action a has been selected in state s , $Q_{UCT}(s, a)$ is the empirical cumulative reward, averaged over all iterations, when action a has been selected in state s , and \mathbb{C} is a problem-dependent parameter to control the balance between exploitation and exploration. Equation (7) is defined to calculate the value of reward $Q(s, a)$:

$$Q(s, a) = \frac{\sum_{i=1}^{n(s, a)} X_i}{n(s, a)}, \quad (7)$$

where X_i denotes the reward of i th simulation associated with action a . The safety objective functions were modeled by equation (3). This objective is focused on the crash potential index. The expectation was that, by adjusting the movement of CAV, the crash potential of the mixed traffic flow can be reduced.

3.2.2. Tree-Expansion Determination Module. When CAV launches a general MCTS algorithm, it will run four steps at any time step. However, sometimes some operations were neither necessary nor helpful in improving the solution quality during the actual operation process. In other words, if the traffic condition was not much changed compared with the last moment, triggering of MCTS does not bring any new information to the simulation, but instead may introduce random noise and grow the tree towards an undesired direction. Additionally, such operation brings significant concerns to the algorithm run time and leads to a waste of memory and CPU resources.

To determine when should the tree expansion be prohibited, we analyze the "marginal impact" of a CAV movement. While CAV performs an action, the HDV that is immediately behind CAV would find a different time headway, and thus, its speed might be adjusted according to equation (8). To determine the degree of adjustment, we perform the partial derivative and can derive the acceleration/deceleration value as follows:

$$\begin{aligned}
\dot{V}(d_{hi}^t) &= \frac{\partial v}{\partial d} \\
&= 16.8 * \frac{\partial [\tanh 0.0860 (d_{hi}(t) - 25) + 0.913]}{\partial d} \\
&= 1.448(1 - \tanh^2 0.0860 (d_{hi}(t) - 25))d(d_{hi}(t)).
\end{aligned} \tag{8}$$

It should be noted that equation (8) merely quantifies the impact of CAV to the vehicle that follows immediately behind it. If multiple vehicles are following CAV, the impact would propagate to the upstream vehicles in the form of shockwave. As such, the total impact is the summation of all vehicles behind CAV, i.e.,

$$\begin{aligned}
\text{Sum}(V) &= \sum V(d_{hi}(t)) \\
&= \sum 1.448 * (1 - \tanh^2 0.0860 (d_{hi}(t) - 25))d(d_{hi}(t)),
\end{aligned} \tag{9}$$

and $\forall i$ behind the CAV vehicle.

3.2.3. Rollout Termination Module. In the simulation step, rapid rollout algorithm is employed to update $Q(s, a)$ value in equation (7) as follows. For a basic simulation, CAV moves with an action that is drawn randomly from the action set, until all vehicles successfully pass through the intersection. This final state is defined as the normal terminal state and thus terminates the simulation process. However, there are some special intermediate states, such as vehicle crash or other kinds of traffic rule violation, after which the simulation lost its practical significance. These final states are defined as the abnormal terminal state that will also terminate the simulation process. In order to further improve the expansion efficiency of Monte Carlo tree and accelerate the rollout algorithm, we create the rollout termination module as equation (10) to identify abnormal terminal state and to shorten the simulation period duration.

Simulation terminates if

$$\min Y(t) \geq l_s + l_a, \quad t = 0, 1, \dots, T, \tag{10a}$$

$$\min D(t) \leq 0, \quad t = 0, 1, \dots, T, \tag{10b}$$

$$\exists y_i(t) \in Y(t), t \in [n * t_c + t_g + t_y, (n + 1) * t_c], \tag{10c}$$

$$n = 0, 1, \dots,$$

$$\min V(t) < 0, \quad t = 0, 1, \dots, T, \tag{10d}$$

$$\max V(t) \geq v_s, \quad t = 0, 1, \dots, T. \tag{10e}$$

This module includes the following cases from equations (10a)–(10e): all vehicles pass the stop line, crash, running red light, reversing, and speeding. The module can avoid unnecessary simulation to reduce unnecessary expansion of the search tree to improve the efficiency of the algorithm. Figure 4 shows the influence of the rollout termination module on the structure of the search tree. It can be seen that

unnecessary tree expansion has been cut after filtering, and the width and depth of the Monte Carlo tree are effectively narrowed.

3.2.4. MCTS-AVS Model. Based on the above modules, the framework of MCTS-AVS algorithm was improved over naïve MCTS algorithm (or the direct application of MCTS algorithm, denoted as n-MCTS) as shown in Figure 5.

The model works with the following steps.

- (1) Start from a current state $s(t) = \{Y(t), V(t)\}$, in which $Y(t)$ is the set of all vehicles' distance from the start position, and $V(t)$ is the set of all vehicles' velocity at time t .
- (2) Tree-expansion determination model determines if it is necessary to launch MCTS algorithm via equations (8) and (9). If yes, go to step 4, otherwise go to step 3.
- (3) Move CAV one step ahead, and update the states of CAV and HDV accordingly. Then, go back to step 1.
- (4) Determine if the maximum number of iterations has been reached. If yes, go to step 5, otherwise go to step 6.
- (5) Update the states of CAV and HDV accordingly, then go back to step 1.
- (6) Do *Selection*: determine the optimal action for CAV with the UCT function via equation (6). Update the states of CAV and HDV.
- (7) Do *Expansion*: randomly select a move for CAV to expand the tree.
- (8) Do *Simulation*: update the states of CAV and HDV, and rollout termination module determine if it is a final state via equations (10a)–(10e). If not, go to step 9. Otherwise, go to step 10.
- (9) Select the next random move, and go back to step 8.
- (10) Do *Backpropagation*: calculate the final benefit of X and update the node value. Then, go back to step 4.

4. Case Study

In this section, the proposed MCTS-AVS algorithm was implemented and tested on a typical arterial roadway segment with signal control. Considering that the minimum intersection spacing along an arterial corridor was usually set to be a quarter mile, the test scenario consisted of a 400-meter roadway with a signal-controlled intersection. Considering the typical congestion on the urban roadway network and the queuing process at intersection, a free flow speed of 8.33 m/s (i.e., roughly 20 mph) was used. After decomposition, CAV became the leading vehicle with a platoon of following HDVs. The platoon had six vehicles that are evenly distributed near the roadway entrance. This scenario was shown in Figure 6, and the specific parameters were listed in Table 1. Then, in MCTS-AVS algorithm, the first vehicle in the platoon was assigned as the CAV. The objective function was set to be minimization of CPI.

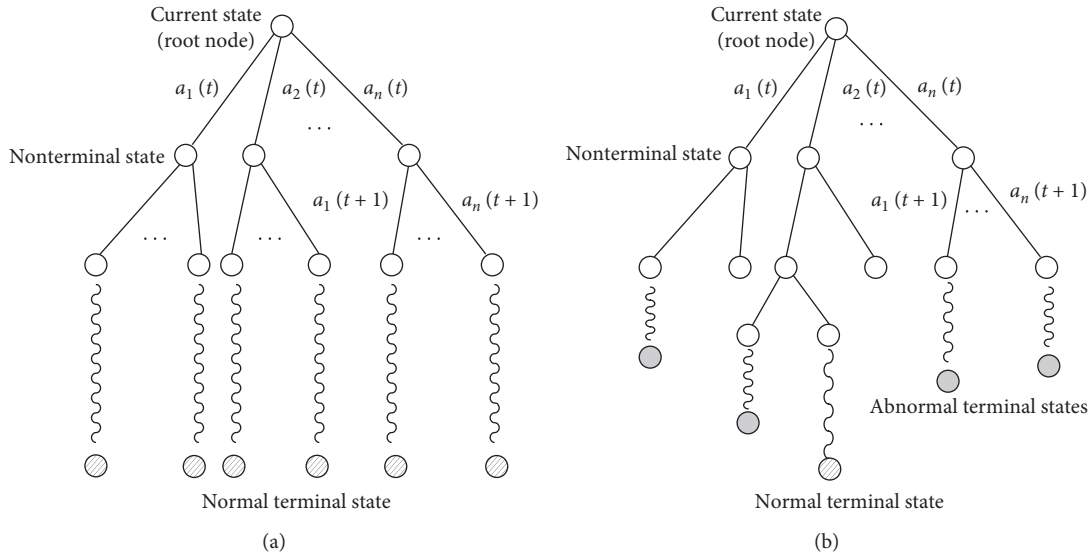


FIGURE 4: Comparison of tree structure: (a) without rollout termination module and (b) with rollout termination module.

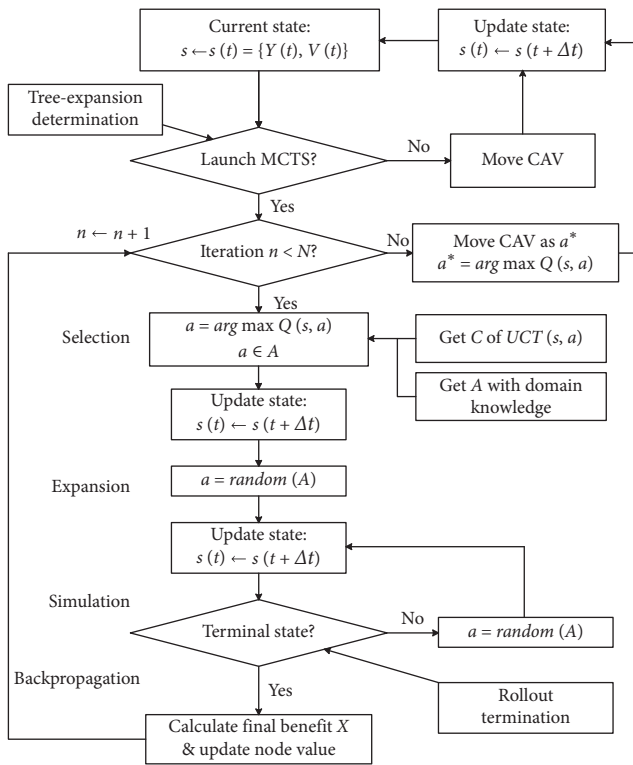


FIGURE 5: The proposed MCTS-AVS algorithm framework.

4.1. Algorithm Result Analysis. For comparison purpose, we defined two benchmark scenarios. The first benchmark scenario had no CAV intelligence, i.e., the CAV drove just like a typical human-driven vehicle. In other words, this first benchmark scenario was equivalent to a pure HDV scenario. The second benchmark scenario used the MCTF-MTF algorithm that was previously developed by the research team [34]. This second benchmark model, however, was developed with the objective of minimize fuel consumption and

travel time of the mixed traffic flow, which makes the comparison with this newly proposed model interesting and demonstrates the safety benefits of this new MCTS-AVS algorithm.

We used the total CPI value minimization as the objective function and found the CPI value dropped from 162.63 in the benchmark scenario (without any CAV intelligence) to 38.12 with the proposed algorithm. In other words, the CPI value was reduced by 76.56%. This benefit was also greater than the previous MCTS-MTF approach, which had a CPI value of 43.36. In other words, when compared with the second benchmark model, a CPI saving of 12.08% was achieved. The capabilities of CPI were also evidenced by the time-space diagram in Figure 7.

In Figure 7, Figure 7(a) represents the benchmark scenario without any CAV intelligence, in which we can see the vehicles firstly drove at a constant high speed to the intersection, then braked and stopped at the intersection due to red light, and finally accelerated and passed intersection when the light turned green. Drastic braking of the lead vehicle caused a series of deceleration of the following HDVs, which significantly increased the crash potential of this traffic flow. On the contrary, a much smoother trajectory was found in Figure 7(b), as this proposed MCTS-AVS algorithm avoided sharp deceleration and acceleration and ensured that CPI value of mixed traffic was kept as low as possible. Figure 7(c) shows a less smooth curve of the previously developed MCTS-MTF method. However, the effect on safety improvement of the previous method was still lower than MCTS-AVS.

4.2. Algorithm Convergence Analysis. Figure 8 below shows the changes in the CPI value at different iterations. The convergence curve shows that CPI value dropped significantly to 38.39 (46.8%) when the number of iterations increased to 25. After that, the results fluctuated with the

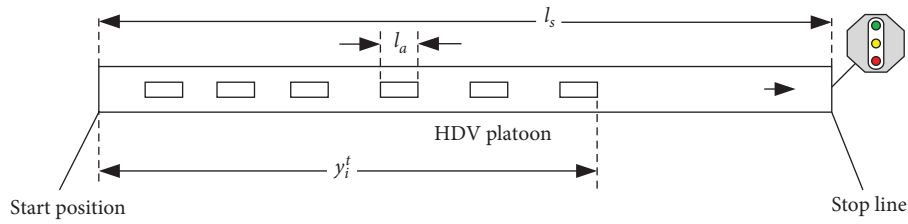


FIGURE 6: The component of case study benchmark.

TABLE 1: Environment variables and hyperparameters.

| Variables and hyperparameters | Descriptions and values |
|-------------------------------|--|
| l_{seg} | Roadway length, 400 meters |
| l_a | Average vehicle length, 5 meters |
| v_s | 20 kph |
| Signal cycle | $t_g = 30s, t_y = 5s, t_r = 35s, \text{ and } t_c = 70s$ |
| C | Balance parameter, 0.08 |
| N | Iteration times, 50 |

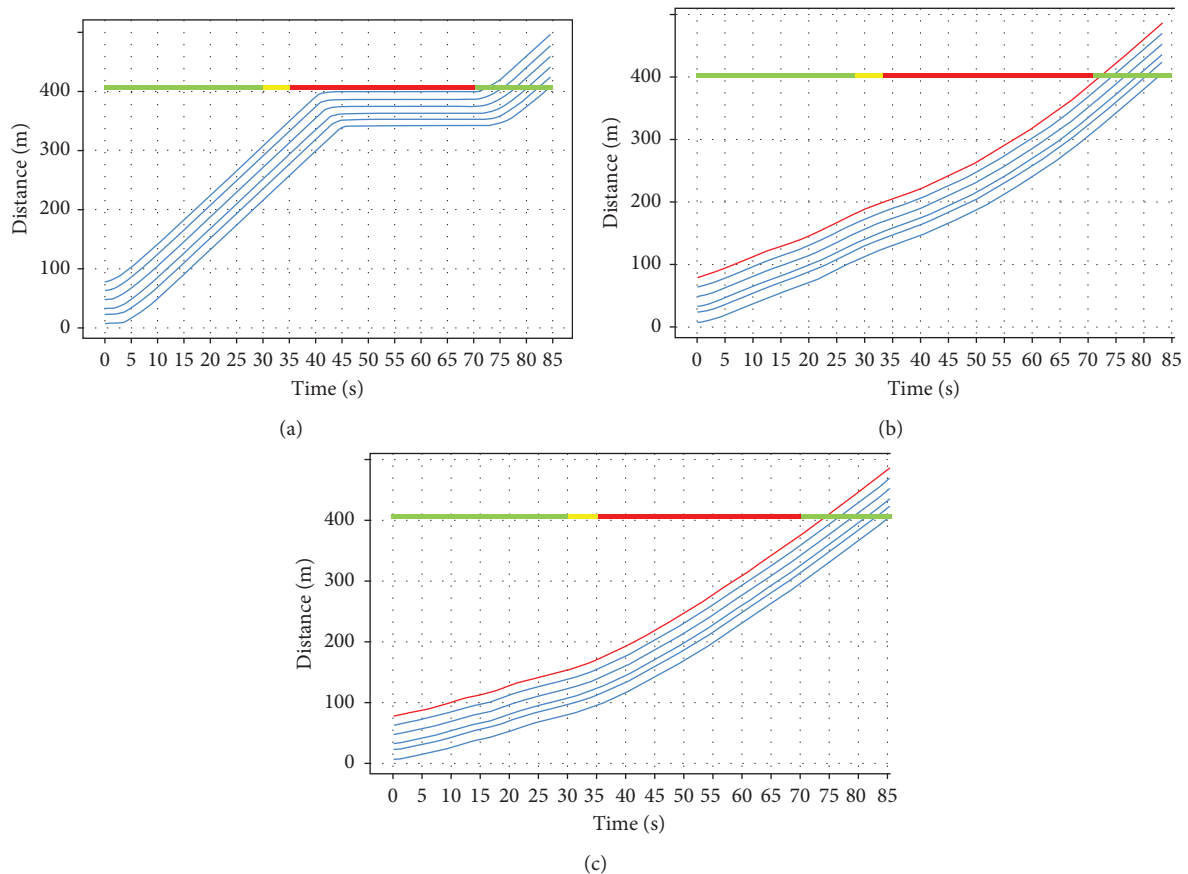


FIGURE 7: Time-space diagram comparison. (a) Benchmark scenario, (b) MCTS-AVS method, and (c) MCTS-MTF method.

increase of iterations. It was also observed that, after 50th iteration, the CPI value actually became very stable, the degree of fluctuation was less than 1, i.e., within $1/73 = 1.37\%$ and can be considered as converged.

4.3. Background Traffic Sensitive Analysis. The algorithm’s performance in the reducing CPI value was further tested with varying level of service (LOS, 1~6 corresponds to A~F), and the results were shown in Figure 9 and Table 2. It

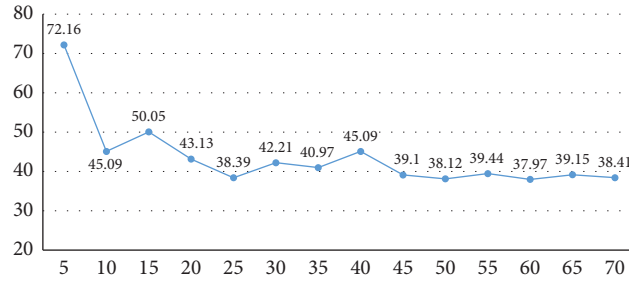


FIGURE 8: MCTS-AVS convergence analysis.

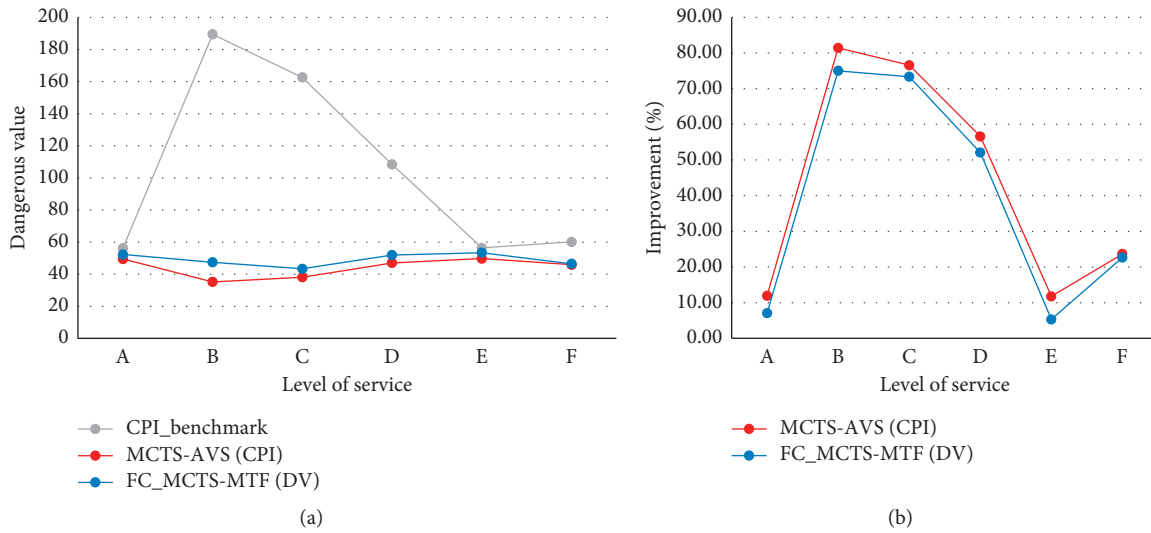


FIGURE 9: Algorithm performance in different LOS. (a) absolute values and (b) saving percentages.

TABLE 2: Algorithm performance comparison in different LOS.

| LOS | Benchmark1 (no intelligence) | Proposed model: MCTS-AVS | Benchmark2: (MCTS-MTF) |
|-----|------------------------------|--------------------------|------------------------|
| A | 56.24 | 49.526 (11.94%) | 52.26 (7.08%) |
| B | 189.54 | 35.22 (81.42%) | 47.41 (74.99%) |
| C | 162.63 | 38.12 (76.56%) | 43.36 (73.34%) |
| D | 108.42 | 47.06 (56.59%) | 51.96 (52.08%) |
| E | 56.4 | 49.76 (11.77%) | 53.41 (5.30%) |
| F | 60.12 | 45.91 (23.64%) | 46.5 (22.65%) |

can be found that, for the CPI value, the maximum saving was observed at LOS B, while near minimum saving was observed at LOS A, E, and F. The guess was that when the traffic was free flowing (e.g., LOS A), not much can be done to reduce the CPI value. On the contrary, there was also a greater risk of collision during a free-flowing traffic (e.g., LOS B) decelerating process due to the change of signal light. Whereas when traffic was congested (i.e., LOS E and F), the percentage of saving was reduced significantly considering slowly moving and a low risk of collision between vehicles.

5. Conclusion and Future Research

This manuscript presents a reinforcement learning modeling approach, named Monte Carlo tree search-based autonomous vehicle safety algorithm, or MCTS-AVS, to optimize the safety of mixed traffic flow, on a one-lane roadway with signalized intersection control. Crash potential index is defined to quantitatively measure the safety performance of the traffic flow. The CAV trajectory planning problem is formulated as an optimization model, and the solution

procedure is proposed. The tree-expansion determination module and rollout termination module are developed to identify and reduce the unnecessary tree expansion, so as to train the model more efficiently towards the desired direction. The case study results found that the proposed algorithm was able to reduce the CPI by 76.56%, when compared with a benchmark model without any intelligence, and 12.08% when compared with another benchmark model which the team developed earlier. These results demonstrated the satisfactory performance of the proposed algorithm in enhancing the safety of the traffic flow.

In order to expand the research scenario from one-lane traffic to a general roadway with multiple lanes, future research may be focused on the following topics. First, how to decompose this mixed traffic to satisfy the proposed algorithm or become a cornerstone of algorithm improvement is a topic worth investigation. Furthermore, with the increase of the number of lanes, there is not only car-following behavior but also lane-changing movements with greater randomness of this scenario. From the algorithm itself, how to improve the simulation efficiency and identify the unnecessary tree expansion node under the complex conditions can also be investigated.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] D. J. Fagnant and K. Kockelman, "Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 167–181, 2015.
- [2] H. Qi, R. Dai, Q. Tang, and X. Hu, "Coordinated intersection signal design for mixed traffic flow of human-driven and connected and autonomous vehicles," *IEEE Access*, vol. 8, pp. 26067–26084, 2020.
- [3] H. Qi, R. Dai, Q. Tang, and X. Hu, "Quasi-real time estimation of turning movement spillover events based on partial connected vehicle data," *Transportation Research Part C: Emerging Technologies*, vol. 120, Article ID 102824, 2020.
- [4] Q. Tang, Y. Cheng, X. Hu, C. Chen, Y. Song, and R. Qin, "Evaluation methodology of leader-follower autonomous vehicle system for work zone maintenance," *Transportation Research Record: Journal of the Transportation Research Board*, 2020.
- [5] A. Eskandarian, C. Wu, and C. Sun, "Research advances and challenges of autonomous and connected ground vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 683–711, 2021.
- [6] N. Chen, M. Wang, T. Alkim, and B. Van Arem, "A robust longitudinal control strategy of platoons under model uncertainties and time delays," *Journal of Advanced Transportation*, vol. 201813 pages. In press(PT.2), Article ID 9852721, 2018.
- [7] Z. Wu, Z. Gao, W. Hao, and J. Ma, "An optimal longitudinal control strategy of platoons using improved particle swarm optimization," *Journal of Advanced Transportation*, vol. 2020, Article ID 8822117, 2020.
- [8] A. A. Malikopoulos, L. Beaver, and I. V. Chremos, "Optimal time trajectory and coordination for connected and automated vehicles," *Automatica*, vol. 125, no. 6, p. 109469, 2021.
- [9] D. Chen, S. Ahn, M. Chitturi, and D. A. Noyce, "Towards vehicle automation: roadway capacity formulation for traffic mixed with regular and automated vehicles," *Transportation Research Part B: Methodological*, vol. 100, pp. 196–221, 2017.
- [10] A. Ghiasi, X. Li, and J. Ma, "A mixed traffic speed harmonization model with connected autonomous vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 104, pp. 210–233, 2019.
- [11] H. Jiang, J. Hu, S. An, M. Wang, and B. B. Park, "Eco approaching at an isolated signalized intersection under partially connected and automated vehicles environment," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 290–307, 2017.
- [12] M. Pourmehrab, L. Elefteriadou, S. Ranka, and M. Martin-Gasulla, "Optimizing signalized intersections performance under conventional and automated vehicles traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 7, 2017.
- [13] M. Wang, W. Daamen, S. P. Hoogendoorn, and B. van Arem, "Rolling horizon control framework for driver assistance systems. Part I: mathematical formulation and non-cooperative systems," *Transportation Research Part C: Emerging Technologies*, vol. 40, pp. 271–289, 2014.
- [14] K. Yang, S. I. Guler, and M. Menendez, "Isolated intersection control for various levels of vehicle technology: conventional, connected, and automated vehicles," *Transportation Research Part C: Emerging Technologies*, vol. 72, pp. 109–129, 2016.
- [15] M. Zhou, Y. Yu, and X. Qu, "Development of an efficient driving strategy for connected and automated vehicles at signalized intersections: a reinforcement learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 433–443, 2020.
- [16] X. He, H. X. Liu, and X. Liu, "Optimal vehicle speed trajectory on a signalized arterial with consideration of queue," *Transportation Research Part C: Emerging Technologies*, vol. 61, pp. 106–120, 2015.
- [17] X. Wu, X. He, G. Yu, A. Harmandayan, and Y. Wang, "Energy-optimal speed control for electric vehicles on signalized arterials," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2786–2796, 2015.
- [18] X. Liang, S. I. Guler, and V. V. Gayah, "Joint optimization of signal phasing and timing and vehicle speed guidance in a connected and autonomous vehicle environment," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 4, pp. 70–83, 2019.
- [19] H. Yao, J. Cui, X. Li, Y. Wang, and S. An, "A trajectory smoothing method at signalized intersection based on individualized variable speed limits with location optimization," *Transportation Research Part D: Transport and Environment*, vol. 62, pp. 456–473, 2018.
- [20] J. Ma, X. Li, F. Zhou, J. Hu, and B. B. Park, "Parsimonious shooting heuristic for trajectory design of connected automated traffic part II: computational issues and optimization," *Transportation Research Part B: Methodological*, vol. 95, no. 2341, pp. 421–441, 2017.
- [21] F. Zhou, X. Li, and J. Ma, "Parsimonious shooting heuristic for trajectory design of connected automated traffic part I:

- theoretical analysis with generalized time geography,” *Transportation Research Part B: Methodological*, vol. 95, no. 5, pp. 394–420, 2017.
- [22] J. Lee and B. Park, “Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 81–90, 2012.
- [23] K. Ahn, H. A. Rakha, and S. Park, “Ecodrive application: algorithmic development and preliminary testing,” *Transportation Research Record Journal of the Transportation Research Board*, vol. 2341, no. 1, pp. 1–10, 2013.
- [24] Z. Cao, S. Jiang, J. Zhang, and H. Guo, “A unified framework for vehicle rerouting and traffic light control to reduce traffic congestion,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 7, pp. 1958–1973, 2017.
- [25] X. Liang, S. I. Guler, and V. V. Gayah, “An equitable traffic signal control scheme at isolated signalized intersections using connected vehicle technology,” *Transportation Research Part C: Emerging Technologies*, vol. 110, pp. 81–97, 2020.
- [26] X. Liang, S. I. Guler, and V. V. Gayah, “Traffic signal control optimization in a connected vehicle environment considering pedestrians,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 10, pp. 499–511, 2020.
- [27] A. Mirheli, L. Hajibabai, and A. Hajbabaie, “Development of a signal-head-free intersection control logic in a fully connected and autonomous vehicle environment,” *Transportation Research Part C: Emerging Technologies*, vol. 92, pp. 412–425, 2018.
- [28] A. Mirheli, M. Tajalli, L. Hajibabai, and A. Hajbabaie, “A consensus-based distributed trajectory control in a signal-free intersection,” *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 161–176, 2019.
- [29] E. Walraven, M. T. J. Spaan, and B. Bakker, “Traffic flow optimization: a reinforcement learning approach,” *Engineering Applications of Artificial Intelligence*, vol. 52, pp. 203–212, 2016.
- [30] M. Treiber, A. Hennecke, and D. Helbing, “Congested traffic states in empirical observations and microscopic simulations,” *Physical Review E*, vol. 62, no. 2, pp. 1805–1824, 2000.
- [31] D. C. Gazis, R. Herman, and R. W. Rothery, “Non-linear follow-the-leader models of traffic flow,” *Operations Research*, vol. 9, no. 4, pp. 545–567, 1961.
- [32] C. B. Browne, E. Powley, D. Whitehouse et al., “A survey of Monte Carlo tree search methods,” *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 4, no. 1, pp. 1–43, 2012.
- [33] H. Qi and X. Hu, “Monte Carlo tree search-based intersection signal optimization model with channelized section spillover,” *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 281–302, 2019.
- [34] Y. Cheng, X. Hu, Q. Tang, H. Qi, and H. Yang, “Monte Carlo tree search-based mixed traffic flow control algorithm for arterial intersections,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 8, pp. 167–178, 2020.