Received January 24, 2019, accepted February 25, 2019, date of publication March 12, 2019, date of current version April 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2904619

A New Solution for Freeway Congestion: Cooperative Speed Limit Control Using Distributed Reinforcement Learning

CHONG WANG, JIAN ZHANG, (Member, IEEE), LINGHUI XU, LINCHAO LI, AND BIN RAN
Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies and Jiangsu Province Collaborative Innovation Center for Technology and Application of Internet of Things, School of Transportation, Southeast University, Nanjing 210096, China

Corresponding authors: Chong Wang (230149206@seu.edu.cn) and Jian Zhang (jiangzhang@seu.edu.cn)

This work was supported in part by the National Key Research and Development Program in China under Grant 2016YFB0100906 and in part by the Fundamental Research Funds for the Central Universities under Grant 2242018K30015. The work of J. Zhang and C. Wang was supported in part by the Zhishan Young Scholar Support Program of Southeast University, and in part by the Science and Technology Major Project, Transportation of Jiangsu Province.

ABSTRACT This paper presents a novel variable speed limit control system under the vehicle to infrastructure environment to optimize the freeway traffic mobility and safety. The control system is a multiagent system consists of several traffic control agents. The agents work cooperatively using the proposed distributed reinforcement learning approach to maximize the freeway traffic mobility and safety benefits. The traffic mobility objective is to maintain freeway traffic density slightly under the critical point to produce the maximum traffic volume, while the traffic safety objective is to reduce the speed difference between adjacent segments. The merits of distributed reinforcement learning are its model-free nature, and it can improve its performance continually as time goes on. The control system is developed on an open source traffic simulation software. Results revealed that compared with no control cases, the proposed system can noticeably decrease the total travel time and increase the bottleneck outflow. Moreover, the speed difference between freeway segments indicating the potential rear-end collision risk is significantly reduced. We also found that there could be more than one optimal traffic equilibrium according to different control objectives, which inspire us to design more optimal strategies in the future.

INDEX TERMS Cooperative variable speed limit control, vehicle to infrastructure technology, distributed reinforcement learning, traffic simulation.

I. INTRODUCTION

Traffic congestion has become a common transportation problem on freeway throughout the world in the past decades. Congestion often occurs near the freeway bottleneck and spreads to its upstream and downstream. Ergo, traffic control approaches such as variable speed limit (VSL) control and ramp metering (RM) are widely used on freeways. In this study, we mainly discuss the solution of freeway recurrent congestion using VSL control methods. Benefits of the VSL control are summarized by Khondaker *et al.* [1], including: (1) safety improvement; (2) resolving traffic breakdown; (3) Improved throughput and environmental benefits. In addition, compared with the ramp metering that are mostly used near the onramp areas, the VSL control can be implemented nearly anywhere on freeway. Hadiuzzaman and Qiu [2], Li *et al.* [3] pointed out that the VSL control can improve traffic mobility by eliminating the ‘‘capacity drop’’ phenomenon and preventing excessive vehicles into the congested area. According to different studies by Hegyi *et al.* [4], [5] and Carlson *et al.* [6], VSL control can reduce total travel time (TTT) from approximately 10% to 40% in total. On the other hand, most studies [7]–[9] showed that the VSL control also has obvious positive effects on traffic safety, either in work zones, or in preventing potential accidents. Piao and McDonald [10] systematically compared the traffic characteristics of VSL control with no control cases and indicated that VSL control can significantly reduce speed differences between and within lanes and number of small headways, thus safety was improved. Nevertheless, some studies (e.g. Piao and McDonald [10], Zegeye *et al.* [11]) held the opinion that there was a trade-off between traffic mobility and safety,

The associate editor coordinating the review of this manuscript and approving it for publication was Xiwang Dong.

the improvement of traffic safety somehow reduced the traffic efficiency. However, they also claimed that if appropriately implemented, VSL might have positive impacts on very turbulent traffic with frequent shockwaves.

In this paper, we developed a novel VSL control system aimed at improving traffic mobility and safety with different control objectives. The mobility objective is to maintain the bottleneck density slightly below the critical value to maximize the bottleneck outflow. The safety objective is to reduce speed differences between adjacent segments to avoid rear-end conflicts. To fulfill the objectives, we introduced the vehicle to infrastructure (V2I) technology and the distributed reinforcement learning (DRL) approach for the multiagent system (MAS). The V2I technology can obtain freeway traffic state precisely, and ensure the vehicles driving at the given speed limit. The DRL approach guide different traffic controllers to work cooperatively and continuously to improve their performance according to the control objectives.

The rest of this paper is organized as follows: Section II describes the related work. Section III is the top-level design of the proposed VSL control system. Section IV provides details of the modified DRL algorithm. Section V is the case study and gives the simulation results of the control system. Section VI presents the conclusions and future work of the research.

II. RELATED WORK

In this section, we survey and discuss the work that related to our research, and state how our proposed approach advances the state of art. There are two categories of related research. One is VSL control with the connected vehicle (CV) or vehicle to infrastructure (V2I) technologies. The other is the reinforcement learning (RL) approaches in traffic control. They are discussed in the following subsections, respectively.

A. VSL WITH CONNECTED VEHICLE TECHNOLOGY

As far as we know, studies of VSL control with CV technology are still quite limited. In the pioneer research by Grumert *et al.* [12], the authors indicated that with the cooperative VSL system (VSLS), the speed was more harmonized, and the exhaust emission was lowered by approximately 1.5-2.5%. Although the traditional VSLS has higher mean speed compared with the cooperative VSLS, there was scarcely any impact on the travel time. The result showed the potential of cooperative VSLS in increasing traffic efficiency and reducing exhaust emission. Khondaker *et al.* [13] further investigated the relations among safety, mobility and fuel consumption of VSL control under different CV penetration. Their study proposed an interesting conclusion that there is consistency in traffic mobility, safety and fuel consumption under CV environment. Wang *et al.* [14] proposed a bi-level architectural control system. The link level control is to tackle the stop and go waves by using the SPEACILIST algorithm. The vehicle level is to use their own behavior model (Wang *et al.* [15]) for individual optimization. Their study proved a generally better result over uncontrolled driving, but discovered an interesting phenomenon that, as the penetration rate of autonomous vehicles increased, there were even more minor jam waves created after the primary jam wave resolved. They explained that the SPEACILIST algorithm probably leads to this problem because it is originally designed for human drivers. However, we also doubt that the vehicle model also contributed to the traffic instability, since each vehicle concerned only self-interest. Therefore, in this study we pay more attention on the link level control in CV environment, which can be associated with the vehicle level control in the future.

B. REINFORCEMENT LEARNING FOR TRAFFIC CONTROL

Despite reinforcement learning (RL) approaches are widely used in automata field, applications in freeway traffic control are still limited and elementary. To our best knowledge, most applied RL methods are the single agent Q learning (QL) approaches, and there is no multiagent reinforcement learning (MARL) approach for VSL control to date. Li *et al.* [16] proposed a single QL agent VSL control system. The QL agent can maintain the bottleneck density below its critical value to relieve traffic congestion. The results showed that the QL approach is superior to the feedback control methods, both in stabilized and fluctuant demand scenarios. Zhu *et al.* [17] developed VSL control system in a large-scale stochastic traffic environment, using RL agents to optimize the traffic flow. Their control agents worked independently. Nonetheless, their optimization reduced the total travel time (TTT) of the network by 18%. In some other studies incorporating the RL with ramp metering, Rezaee *et al.* [18], [19] discussed the details of RL approach implementation on ramp metering and compared it with the ALINEA algorithm. Results showed that although the RL approach outperformed the ALINEA significantly, the mainline TTT are nearly the same. Fares *et al.* [20] proposed a primary MARL based ramp metering algorithm, which can reduce 6% of the mainline TTT and increase 7.5% of the average speed. However, their systemic value function is the harmonic mean value of independent RL agents, thus it is difficult to tell whether the algorithm can converge to the global optimum value.

On the other hand, there are considerable studies on the MARL urban traffic signal control. El-Tantawy *et al.* [21] thoroughly analyzed the previous related studies and pointed out that in fact most of them are still independent RL approaches. They proposed a MARL algorithm to maximize expected Q-value among neighbor agents in the network. The algorithm can reduce up to 39% intersection delay and save 26% travel time along the busiest roads. Kuyer *et al.* [22] also presented an explicit coordination mechanism between learned agents based on max-plus algorithm, which substantially outperforms the independent RL algorithms. The imperfection is that the method is model dependent and computationally demanding. Furthermore, the agents have to negotiate with each other frequently to report their latest actions, which are not always necessary.

The above-mentioned studies proved that the CV technology and RL approaches have enormous potential in traffic control. However, studies that consider both sides are rare. Besides, there are some shortcomings in current studies need to improve. First, many studies simply partition the continuous traffic state space into the discrete RL control state, which may be negative for convergence. Second, most RL studies in traffic control only discussed traffic mobility improvement, while considerations on safety improvement are seldom mentioned. Third, most existed studies focus on single RL agent traffic control, which are not sufficient for large road network.

This paper proposed a distributed VSL control system with V2I technology to solve the problems presented above. The distributed system can deploy the controllers flexibly along the freeway, and there is no worry about the breakdown of central traffic controller. For each control agent, we programmed a modified distributed Q learning (DQL) algorithm to tackle with the cooperative control problem in continuous traffic state space. Moreover, we proposed a safety objective function derived from the ‘‘Time to Collision’’ (TTC) equation to improve the traffic safety. We also discussed the implementation details of the V2I technology and DQL algorithm on the VSL control system. Since study [21] indicated that the MARL outperforms independent RL significantly, we believe a similar approach also suits large freeway network and it is inevitable trend of the future transportation system.

III. CONTROL SYSTEM

A. OVERALL FRAMEWORK

This section presents the framework of the cooperative VSL control system. As illustrated in Figure 1, the freeway is divided into several equidistance segments. Each segment has one onramp or off-ramp at the most. The length of segment is set longer than the free flow travel distance, i.e. $\mathrm{v_{free}}\Delta t \leq \Delta x$, to meet the Courant-Friedrichs-Lewy (CFL) conditions. For each segment i, state at time t refers to a speed and density pair, i.e. $state_i(t) = [v_i(t), d_i(t)]$.

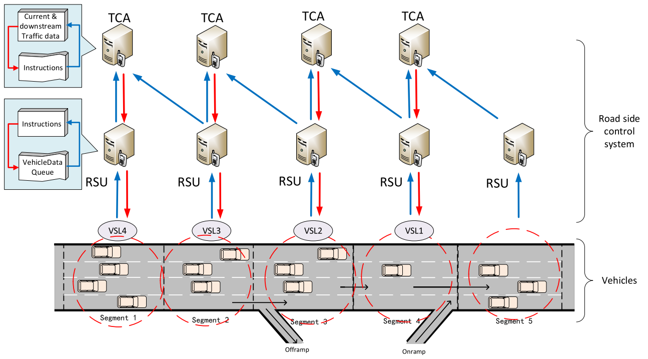


FIGURE 1. The cooperative VSL control system framework.

Each segment has a roadside unit (RSU). The RSU collects the real time traffic state and send it to the traffic control agent (TCA). The TCA extracts the necessary data from the traffic state according to different objective functions, and then calculate the optimal result (i.e. speed limits) using the MARL algorithm. Next, the TCA sends the optimal result back to the segment's RSU. Finally, the RSU sends the speed limit to the vehicles in the segment. As the V2I control system, we assume the vehicles in the study to have basic communication and automatic control ability. They can send their state (position and speed) to the RSU, receive instructions (i.e. speed limits) from the RSU and adjust their speed accordingly. In addition, vehicles can make necessary lane changes automatically.

The meanings of major symbols in this paper are shown in Table 1.

TABLE 1. Meaning of symbols.

Symbol	Meaning
$v_i(t)$	Space average speed of segment i at time t
$d_i(t)$	Density of segment i at time t
$\Delta v_{i,i+1}(t)$	Speed difference between segment i and $i+1$
$N_i(t)$	Vehicle number of segment i at time t
$x_j(t)$ / $v_j(t)$	Position/speed of vehicle j at time t
$v_{j+1}(t)$	speed of vehicle j's front vehicle at time t
L	Length of each segment
$\boldsymbol{s}_t$	Full control state at time t
$\boldsymbol{x}_t$	Full control state of state centers at time t
$x_t^{(m)}$	Control state of state center m at time t
$\boldsymbol{u}_t$	Joint actions of agents
$a_i^{(t)}$	Agent i's action at time t
R_t	Reward at time t
$\Pi(\boldsymbol{s}_t)$	Joint policy for $\boldsymbol{s}_t$
$\pi^{(i)}(\boldsymbol{s}_t)$	Component policy of agent i for $\boldsymbol{s}_t$
$Q_t^{(i)}$	Q value of agent i at time t
$Q_t^{(i,j)}$	Q value of agent i's state center j at time t
w_j	weight associated with state center x_j
p_j	Probability coefficient derived from w_j
ε	Random number of ε-greedy policy
σ	Degree of approximate equality
d_{cr}	Critical density
d_{cg}	Congestion density
$\rho_{s_t,j}$	Distance between $\boldsymbol{s}_t$ and state center j
v_{free}	Free flow speed

B. TRAFFIC STATE COLLECTION

The process of an RSU to collect traffic state and execute the speed limits is shown in Figure 2. When a vehicle is entering a segment, a wireless connection is automatically established and a vehicle data package (VDP) is inserted into the Vehicle Data Queue (VDQ) inside the RSU. The VDP updates the vehicle data every control period. At the same time, the RSU traverses the VDQ and aggregate the segment traffic state from the VDPs. The traffic state of an arbitrary segment i, i.e. the space mean speed $v_i(t)$ and density $d_i(t)$, can be calculated as follows:

$$v_i(t) = \frac{\sum_N v_j(t)}{N(t)}, d_i(k) = \frac{N_i(t)}{L} \tag{1}$$

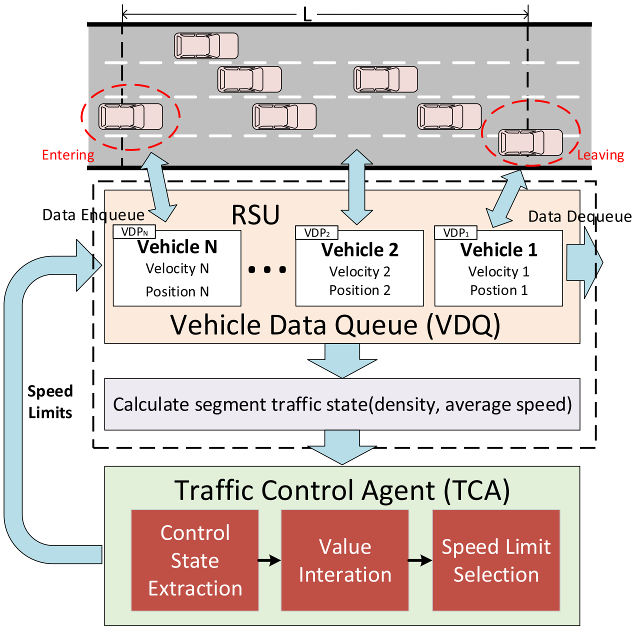


FIGURE 2. Traffic state collection and speed limits executing process.

where $N_i(t)$ is the number of vehicles (equal to the length of VDQ) in segment i at time t, $v_j(t)$ is velocity of each vehicle at time t, L is the length of the segment. After the RSU received the speed limits from the TCA, the RSU will notify the packages in the VDQ. Then the speed limits are transferred to the vehicles through the VDPs. When a vehicle leaves the segment, the connection is closed and its VDP is deleted.

C. THE CONTROL STATE EXTRACTION

The control state refers to the state used by the TCA for optimal control. As mentioned before, the control state is extracted from the traffic state of RSU. Different optimal objectives require different control state. For mobility optimization, we need to know the downstream density. Meanwhile, for safety optimization, the average speed of the current segment and speed difference with downstream segment are required. The control state extraction procedure is shown in Figure 3.

The control state extraction procedure has four major steps:

(1) The RSU aggregates the vehicle data to segment traffic state;
(2) The Traffic Data Center collects traffic state from all RSUs and updates the freeway corridor traffic state;
(3) The TCAs select the control segments and get the segments traffic state from traffic data center;
(4) The TCAs select necessary traffic state element according to the objectives.

IV. CONTROL METHOD

A. THE BASIC DISTRIBUTED Q LEARNING

There are several algorithms for the multiagent system (MAS) problem [23]–[26]. Here we use the DQL algorithm proposed by studies [25] and [26] to solve the cooperative VSL control problem. The DQL algorithm has the following advantages: (1) It has been mathematically proved to converge to the optimal solution. (2) Its control state is fully observed. Fully observation is considered to be more precise than the partly observation algorithms, though it consumes more computing resources. (3) It does not require additional communication between agents. In summary, it is more suitable for middle-sized network, which meets our needs.

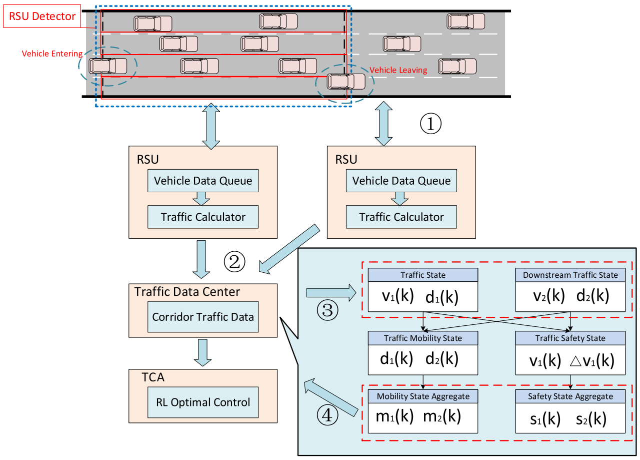


FIGURE 3. Control state extraction process.

We treat the cooperative VSL control with m agents as a deterministic multiagent Markov Decision Process (MDP). The control system can be defined as a tuple (S, A^m, T, R, γ), where:

- S is a finite set denotes the observed state, which is extracted from the freeway traffic state;
- A is a finite set of all elementary actions that can be chosen by an agent. A^m is the set of the combined actions of all agents, which is a set of speed limits;
- $T : S \times U \rightarrow S^{'}$ is the transition function which gives the probability that state S transfer to $S^{'}$ as a joint action $U = \left(a^{(1)}, \ldots, a^{(m)}\right)$ is implemented. For deterministic problem, the probability is one. In the case study, the transition function is implicated in the traffic model;
- $R : S \times U \rightarrow R$ is the joint reward of action U, all the agents share the same reward when learning.
- $0 \leq \gamma < 1$ is the discount factor.

The aim of learning is to find an optimal joint policy according to different state $\Pi{:}S \rightarrow U$ that can maximize the summed discounted reward $\sum_{t=0}^{\infty} \gamma R_t$, where R_t is the reward at time step t. R_t is determined by the action selection policy Π: $R_t = (s_t, \Pi(s_t))$. For the DQL algorithm, the join policy can be split into component policy $\Pi(s_t) = \left(\pi^{(1)}(s_t), \ldots, \pi^{(m)}(s_t)\right)$. The DQL algorithm has two major components: the value iteration and the action selection policy. For each agent, we suppose it can observe the full state, but have to estimate actions of other agents.

The estimation is based on the optimistic assumption that every agent believes others will choose the most benefit action. Therefore, the value function of an agent i is:

$$Q_{t+1}^{(i)}\left(\boldsymbol{s}_t, a_t^{(i)}\right) = \begin{cases} Q_t^{(i)}\left(\boldsymbol{s}_t, a_t^{(i)}\right) = R_{t+1}^{(i)}\left(\boldsymbol{s}_t, \boldsymbol{u}_t\right), \ \textit{if } \boldsymbol{s}_t \textit{ not exist} \\ \max\left\{Q_t^{(i)}\left(\boldsymbol{s}_t, a_t^{(i)}\right), R_{t+1}^{(i)}\left(\boldsymbol{s}_t, \boldsymbol{u}_t\right)\right. \\ \left. +\gamma \max_{a^{(i)}\in A} Q_t^{(i)}\left(\boldsymbol{s}_{t+1}, a^{(i)}\right)\right\}, \textit{otherwise} \end{cases} \tag{2}$$

where $Q_t^{(i)}\left(\boldsymbol{s}_t, a_t\right)$ represents the local Q table of agent i at time step t, $\boldsymbol{s}_t$ refers to the freeway traffic state, $\boldsymbol{u}_t$ refers to the joint action, and a_t refers to the action taken by agent i, respectively. $R_{t+1}^{(i)}\left(\boldsymbol{s}_t, \boldsymbol{u}_t\right)$ is the reward of joint action $\boldsymbol{u}_t$, which is calculated from the objective function. Note that to ensure convergence, the following constraint should meet: $R_{t+1}^{(i)}\left(\boldsymbol{s}_t, \boldsymbol{u}_t\right) \geq 0$ for all $\boldsymbol{s}_t \in S$ and $\boldsymbol{u}_t \in A$.

The central idea behind the value iteration is to update Q value only when a new action results in an improvement over all other actions previously applied in the current state. Similarly, when come to the action selection policy, the agents have to select the optimal action hitherto to ensure its convergence. Hence, the update rule for agent individual policies $\pi_t^{(i)}$ can be written as:

$$\pi_{t+1}^{(i)}\left(\boldsymbol{s}_t\right) = \begin{cases} \pi_t^{(i)}\left(\boldsymbol{s}_t\right), \\ \textit{if } \max_{a^{(i)}\in A} Q_t^{(i)}\left(\boldsymbol{s}_t, a^{(i)}\right) = \max_{a^{(i)}\in A} Q_{t+1}^{(i)}\left(\boldsymbol{s}_t, a^{(i)}\right) \\ a_t^{(i)}, \quad \textit{otherwise} \end{cases} \tag{3}$$

In fact, Eq. (3) is the greedy policy. It only works after the state is explored sufficiently. Here we use the ε-greedy policy $\tilde{\pi}$ to balance the exploration and exploitation. The ε-greedy policy $\tilde{\pi}^{(i)}$ for agent i can be written as:

$$\tilde{\pi}_{t+1}^{(i)}\left(\boldsymbol{s}_t\right) = \begin{cases} \forall a^{(i)} \epsilon A, & \textit{if } \varepsilon < e^{-En}\left(n \in N^{+}\right) \\ \pi_{t+1}^{(i)}\left(\boldsymbol{s}_t\right), & \textit{otherwise} \end{cases} \tag{4}$$

where ε is a random number, E is the coefficient to control the descent speed of the exponential. If $\varepsilon < e^{-En}$, an arbitrary action is selected, this is the exploration. Otherwise, the optimal action fits with Eq. (3) is selected, this is the exploitation. e^{-En} is the probability function that gradually reduces as visits to state $\mathbf{s}_t$ increases, thus the random actions are reduced as exploration becomes more sufficient. In summary, Eq. (2), (3) and (4) constitute the DQL algorithm for VSL control. The state of the algorithm is discussed in the following sections. The action set of the algorithm is given below.

For the VSL control system, actions are the collection of legal speed limits. For traffic mobility and safety considerations, the limited speed should meet the following constraints:

1. The maximum speed limit must not higher than the free flow speed, i.e. $v_{SL,max} \leq v_{free}$;
2. The minimum speed limit must not lower than an acceptable speed to ensure the minimum flow, i.e. $v_{SL,min} \geq v_{min}$;
3. The change of speed in two consecutive time steps must not exceed a maximum difference value, i.e. $\left|v_{SL}\left(\mathrm{t}+1\right) - v_{SL}\left(\mathrm{t}\right)\right| \leq v_{diff}$;

For this study, the v_{diff} is set to 10km/h, $v_{SL,min}$ is set to 30 km/h, and $v_{SL,max}$ is set to 100km/h. In summary, the action set is {30, 40, 50, 60, 70, 80, 90, 100} km/h.

B. DISTRIBUTED Q LEARNING WITH CONTINUOUS CONTROL SPACE

As pointed out by study [19], using a table to represent the value function limits the practicality of Q learning approach when tackle with continuous state problem. Alternatively, the k-NN-TD algorithm [27] can be used as the general function approximator. Instead of approximate traffic state to the nearest discrete state, k-NN-TD algorithm generates a set of centers in the state space. Extracted state in section 3 can map to these centers. The Q values of the nearest centers are updated according to the distance from the state. Here we integrate the k-NN-TD algorithm with the DQL algorithm. To be more specific, if agent i has k nearest centers with its local state $s_t^{(i)}$, the full state $\boldsymbol{s}_t = \left\{s_t^{(1)}, \ldots, s_t^{(m)}\right\}$ can be rewritten with the state centers: $\boldsymbol{x}_t = \left\{\boldsymbol{x}_t^{(1)}, \ldots, \boldsymbol{x}_t^{(m)}\right\}$, where the $\boldsymbol{x}_t^{(m)} = \{x_{1,t}^{(m)}, \ldots, x_{k,t}^{(m)}\}$. k is the number of the nearest state centers. The weight associated with center x_j is:

$$w_j = \frac{1}{1 + \rho_{s_t,j}^2} \quad \forall j \in [1, \ldots, k] \tag{5}$$

where $\rho_{s_t,j}$ denotes the Euclidean distance between $s_t^{(i)}$ and $x_{j,t}^{(m)}$. Let $Q_t^{(i,j)}\left(\boldsymbol{x}_t, a_t^{(i)}\right)$ denotes the Q value of agent i's center j, relations between $Q_t^{(i,j)}\left(\boldsymbol{x}_t, a_t^{(i)}\right)$ and $Q_t^{(i)}\left(\boldsymbol{s}_t, a_t^{(i)}\right)$ in Eq. (2) are:

$$Q_t^{(i)}\left(\boldsymbol{s}_t, a_t^{(i)}\right) = \sum_{j=1}^{k} p_j Q_t^{(i,j)}\left(\boldsymbol{x}_t, a_t^{(i)}\right) \tag{6}$$

$$Q_t^{(i,j)}\left(\boldsymbol{x}_t, a_t^{(i)}\right) = p_j Q_t^{(i)}\left(\boldsymbol{s}_t, a_t^{(i)}\right) \tag{7}$$

where $p_j = \frac{w_j}{\sum_k w_j}$ is the probability coefficient. Then the Eq. (2) can be solved with $Q_t^{(i,j)}\left(\boldsymbol{x}_t, a_t^{(i)}\right)$. Similarly, the action selection policy (Eq. (3)) can be modified as:

$$\pi_{t+1}^{(i)}\left(\boldsymbol{x}_t\right) = \begin{cases} \pi_t^{(i)}\left(\boldsymbol{x}_t\right), \quad \textit{if} \left| \max_{a^{(i)}\in A} \sum_{j=1}^{k} p_j Q_t^{(i,j)}\left(\boldsymbol{x}_t, a_t^{(i)}\right) \right. \\ \left. - \max_{a(i)\in A} \sum_{j=1}^{k} p_j Q_{t+1}^{(i,j)}\left(\boldsymbol{x}_t, a_t^{(i)}\right) \right| \leq \sigma \end{cases} \tag{8}$$

where σ reflects the degree of approximate equality. If difference between $Q_t^{(i,j)}$ and $Q_{t+1}^{(i,j)}$ is less than σ, they can be considered as the same Q value. k refers to the number of state centers within distance δ from state $s_t^{(i)}$. In this paper, δ is set to the maximum distance between state centers.

Note that though regard with form, the state $s^{(i)} \rightarrow \boldsymbol{x}^{(i)}$ seems enlarged the state space. In fact, the continuous state space $s^{(m)}$ is infinite while $\boldsymbol{x}^{(i)}$ is finite and discrete. The actual state space is reduced.

C. REWARD FUNCTION OF MOBILITY CONTROL

It is obvious from the fundamental diagram that the traffic flow reaches the maximum value when its density is near the critical point, and traffic flow decreases when the density deviates from the critical point. As a result, maintain segment density slightly under the critical density is a widely used control strategy for traffic mobility control [16], [28]. A similar strategy is used here. The reward function can be written as:

$$Rm_i(t) = \begin{cases} \dfrac{d_{i+1}(t)}{d_{cr}} r_{max}, & d_{i+1}(t) \leq d_{cr} \\ \dfrac{d_{i+1}(t)-d_{cg}}{d_{cr}-d_{cg}} r_{max}, & d_{cr} < d_{i+1}(t) \leq d_{cg} \\ 0, & d_{i+1}(t) > d_{cg} \end{cases} \tag{9}$$

where $d_{i+1}(t)$ denotes the downstream density at time step t, r_{max} denotes the maximum reward, d_{cr} denotes the critical density, and d_{cg} denotes the congestion density, respectively. In our paper, d_{cr} is set to 26 veh/km/lane, d_{cg} is set to 45 veh/km/lane, and the r_{max} is one. Eq. (9) is a piecewise function, which may lead to unstable training near the critical density. In practice, we use a fifth-order polynomial reward curve to fit Eq. (9). As shown in Figure 4(a), the original mobility reward is the square marker curve, while the fitting reward is the red dot curve. The red dot curve is smoother than the original mobility reward.

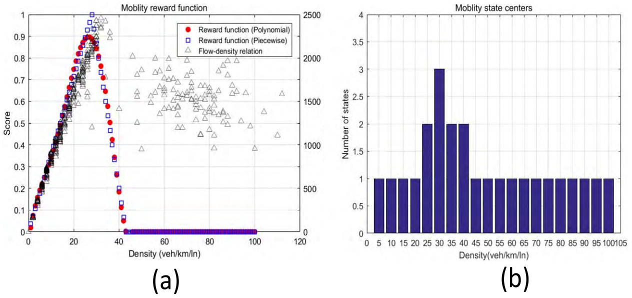


FIGURE 4. The traffic mobility reward (a) and mobility state centers (b).

Study [19] used the bottleneck outflow as the objective. The problem is that the same outflow might occur in both free flow and congestion flow, hence it is hard for the agent to distinguish them. Instead, our density reward function can set the congestion reward limb descent faster than the free flow limb, to encourage the agents to select the actions lead to uncongested traffic. As mentioned above, the control state centers are related to the reward function. The density state centers are set from 0 to d_{jam} (i.e. 100 veh/km/lane) with the interval of 1∼5 veh/km/lane. As pointed out by studies [16] and [19], traffic flow changes more sensitively near the critical density, thus smaller intervals are adopted near the critical point. Figure 4(b) shows the distribution of the density state centers.

D. REWARD FUNCTION OF SAFETY CONTROL

Khondaker *et al.* [13] introduced the "Time To Collision" equation in microscopic model to measure the potential rear-end collision risk between a pair of vehicles, which is written as follows:

$$TTC = \begin{cases} \dfrac{x_{j+1}(t)-x_j(t)}{v_j(t)-v_{j+1}(t)}, & if\ v_j(t) > v_{j+1}(t) \\ \infty, & if\ v_j(t) \leq v_{j+1}(t) \end{cases} \tag{10}$$

where $i+1$ is the leading vehicle and i is the following vehicle, $x_j(t)$ and $v_j(t)$ represents the position and speed of the vehicle at time step t, respectively. Eq. (10) indicates that when a following vehicle is driving faster than the leading vehicle, there is a potential collision risk. Obviously, longer TTC means less collision risk. Therefore, we can derive the traffic safety objective from Eq. (10).

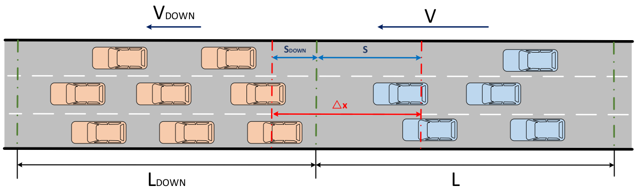


FIGURE 5. Time to collision between two segments.

As shown in Figure 5, the collision is most likely to happen at the rear vehicles in segment $i+1$ and the front vehicles in segment i. The relative distance Δx between them is approximately the sum of the mean space headway of segment i and segment $i+1$. The mean space headway is the reciprocal of segment density. Therefore, Eq. (10) at the macroscopic level can be written as (t is omitted):

$$TTC_{Macro} = \begin{cases} \dfrac{\dfrac{1}{d_i}+\dfrac{1}{d_{i+1}}}{v_i-v_{i+1}} = \dfrac{d_i+d_{i+1}}{d_i d_{i+1}\left(v_i-v_{i+1}\right)}, & if\ v_i > v_{i+1} \\ \infty, \ if\ v_i \leq v_{i+1} \end{cases} \tag{11}$$

where segment $i+1$ is the downstream of segment i, v_i and d_i denotes the average speed and density of segment i, respectively. The safety objective is the product of potential collision vehicles and TTC_{Macro}. In addition, we assume that: (1) the VSL control for safety works only when the traffic is near saturation, at that time the outflow of upstream and downstream are nearly the same, (2) the average upstream speed is higher than the downstream speed, hence the upstream density is lower than downstream in saturated traffic. In this case, the number of potential collision vehicles is approximated to the upstream density $d_i(t)$. With the two assumptions and the relation $q_i = d_i v_i$, the safety function can be simplified as:

$$Rs_i(t) = d_i(t) \cdot TTC_{Macro} = \begin{cases} \dfrac{1 + C_{i,i+1}}{\Delta v_{i,i+1}(t)} - \dfrac{C_{i,i+1}}{v_i(t)} \left(\Delta v_{i,i+1} > 0\right) \\ max\ reward,\ otherwise \end{cases} \quad (12)$$

where $\Delta v_{i,i+1}(t) = v_i(t) - v_{i+1}(t)$ denotes the speed difference between adjacent segments, $\mathrm{C}_{i,i+1} = \frac{q_i}{q_{i+1}}$ is the volume ratio of segment i and $i+1$. $\mathrm{C}_{i,i+1}$ is approximated as one in saturated traffic. The $\Delta v_{i,i+1}(t)$ is relatively small compared with segment speed $v_i(t)$, which restricts the result above zero. Eq. (12) denotes that if the speed difference between two segments is comparatively small, higher following speed (which implies a lower traffic density) can improve traffic safety. This can partly explain that to improve traffic safety will not always hurt traffic mobility. The logic of Eq. (12) is that the safety control agents stabilize the segment speed first, and then increase the segment average speed (traffic safety is improved at the same time). This can also encourage agents to choose higher speed limits when the speed difference is similar. The reward function of traffic safety is depicted in Figure 6(a).

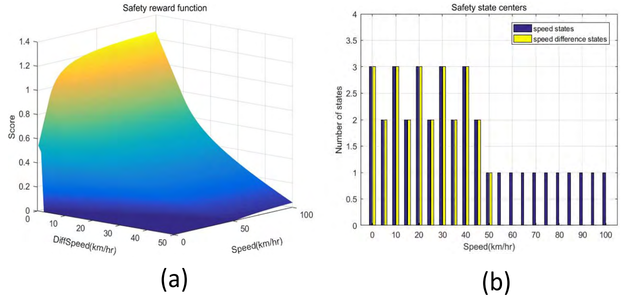


FIGURE 6. The traffic safety reward (a) and safety state centers (b).

According to the reward function, safety state is the Cartesian product of speed difference between adjacent segments $\Delta v_{i,i+1}(t)$ and average upstream speed $v_i(t)$. Similarly, the safety state at time t can be written as: $state_i(t) = [\Delta v_{i,i+1}(t), v_i(t)]$. The range of the speed difference $\Delta v_{i,i+1}(t)$ is from 0 to 50 km/h, and the average speed $v_i(t)$ is from 0 to 100 km/h, both with the state center interval of 10 km/h. Figure 6(b) gives the distribution of safety state centers.

To sum up the above points, the DQL algorithm for a VSL control agent is described with pseudocode in detail as Algorithm 1. Each time step, all the TCAs along the freeway use algorithm 1 to optimize traffic. Note that the total reward R_k is the expectation of all agents' reward.

V. SIMULATION AND RESULT

A. SIMULATION SET UP

We have developed and tested the proposed VSL control system on the open source simulation platform MOTUS [29]. The workflow of the simulation environment is shown in Figure 7. Simulation input includes the road network from google earth (KML) file, RSU (including TCA) info file and traffic demand file. Simulation output includes the speed, the density and the volume data of the freeway segments, which are exported to the csv files.

In the study, the A16 freeway section near Drechttunnel, Netherlands is selected as the testbed, which is shown in Figure 8(a). The freeway section is 2km long with an off-ramp and an onramp. The corresponding simulation network is given in Figure 8(b), which is split to 10 equidistance segments. Each segment has an RSU to collect traffic state. There are three TCAs along the road for VSL control. The traffic demands of mainline and onramp are given in Figure 9. Both demands are increased after an hour and the peak period

Algorithm 1 Learning for Control Agent i using DQL

```
Initialization at time $t = 0$: state space $S$, action
Space $A_i$, discount factor $\gamma$
$Q_0^{(i,j)}(s, a) \leftarrow 0, \forall s \in S, a \in A_i$
Extract initial state $\mathbf{s}_0$ from Traffic Data Center
Foreach time step $t \geq 0$ do:
    Calculate weight $w_j$ of each state center $j$
according to distance $\rho_{s_t,j}$ (Eq. (5))
    Update $Q_t^{(i,j)}\left(s_t, a_{t-1}^{(i)}\right)$ according to $w_j$ (Eq. (7))
    Select an action $a_t^{(i)}$ according to $\varepsilon$-greedy
policy $\tilde{\pi}_t^{(i)}(s_t)$ (Eq. (4) and (8))
    Observe total reward $R_t$:
    For each agent $j \in TCA$, $s_t^{(j)} \in \boldsymbol{s}_t$
        If traffic mobility objective:
            $R_j =$ Eq. (9)
        Else(traffic safety objective):
            $R_j =$ Eq. (12)
        End If
    End for
    $R_t = \sum_{j=1}^{m} R_j / m$
    Extract new state $\boldsymbol{s}_{t+1}$ from Traffic Data Center
    Update $Q_t^{(i)}\left(s_t, a_t^{(i)}\right)$ (Eq. (2) and (6))
    Update greedy policy $\pi_t^{(i)}(s_t)$ (Eq. (6) and (8))
    $\boldsymbol{s}_t \leftarrow \boldsymbol{s}_{t+1}$
End For
```

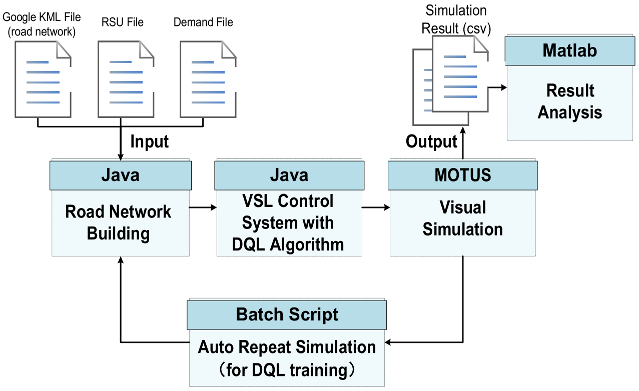


FIGURE 7. Workflow of the simulation environment.

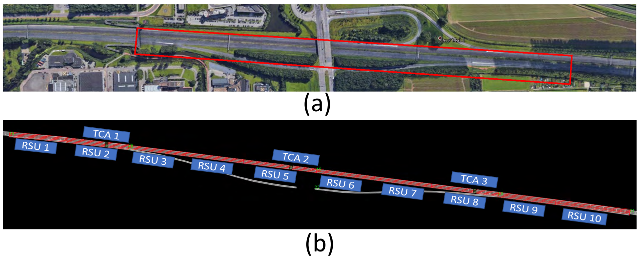


FIGURE 8. Illustration of study freeway section (a) and corresponding simulation network (b).

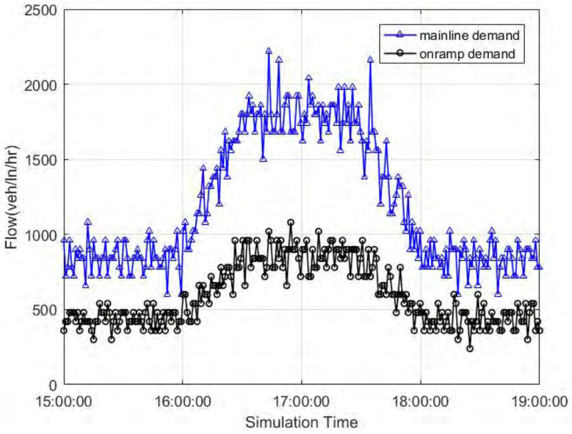


FIGURE 9. Simulation traffic demands of mainline and onramp.

lasts for another hour. To be more realistic, there are small fluctuations in traffic demand. The bottleneck forms at 9th segment due to increasing traffic demand.

MOTUS has developed the Intelligent Driver Model (IDM+) and the Lane-change Model with Relaxation and Synchronization (LMRS) to simulate car-following and lane-changing behaviors. Hence, simulation vehicles can follow the given speed and adjust their speed according to the surroundings automatically. Hence, the simulation can satisfy the assumptions for vehicles in section III. The key parameters includes the free flow speed of 100 km/h, the critical density of 26 veh/km/lane, and the time step of 30 seconds. In addition, the vehicle type is 90% cars and 10% trucks during the simulation. Other model parameters are calibrated using real historical traffic data. Figure 10 is the traffic calibration result. Compared with the real data, the traffic model can satisfy the traffic control requirement. For the details of model calibration, study [30] is recommended.

B. RESULTS ANALYSIS

Two scenarios aimed at traffic mobility and safety improvement are tested. For comparison, we also considered the no control case. The total travel time (TTT) and the total number of vehicle stops (TNVS) are utilized to analyze the simulation results. Their equations are (13) and (14), respectively.

$$TTT = \sum_{t=1}^{t_s} \sum_{i=1}^{n_c} T_c N_i(t) \tag{13}$$

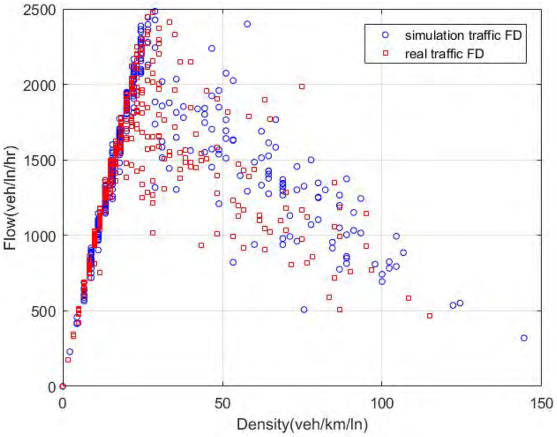


FIGURE 10. Simulation model result compared with the real traffic data.

$$TNVS = \sum_{t=1}^{t_s} \sum_{i=1}^{n_c} N_i^v(t), \tag{14}$$
$$v(\mathrm{t}) < v_{min} \wedge v(\mathrm{t}-1) \geq v_{min}$$

where T_c denotes the control period, $N_i(t)$ denotes the number of vehicles in segment i at time step t, $N_i^v(t)$ denotes the number of vehicles that velocity is lower than the threshold v_{min}, t_s denotes the duration of simulation, and n_c denotes the total segments of the freeway section. v_{min} is set to 5km/h.

Three scenarios are tested: the no control scenario, the traffic mobility control scenario, and the traffic safety control scenario. We made two groups of comparison. One is from the perspective of the entire freeway section, the other one is from the bottleneck segment. We used the TTT of both freeway section and bottleneck, the bottleneck mean speed and the outflow to measure the traffic mobility improvement. We used the total number of stops and the mean speed difference to measure the traffic safety improvement. Both "mean speed difference" and "SD of speed" are used for traffic fluctuation assessment, higher value implies higher collision risk. The "peak hour" (16:30-18:00) simulation results are shown in Table 2.

Table 2 illustrates that, compared with the no control case, DQL based VSL control have an apparent improvement in traffic mobility and safety. The stop and go phenomenon is eliminated and the traffic runs smoother. The bottleneck outflow is also increased. It is expected that the traffic safety control does not have a pronounced traffic outflow increment, since the objective is focused on reducing speed variance. Traffic safety control has smaller mean speed difference than the other two cases, indicates that the traffic under safety control is more stable. Nevertheless, it seems that the VSL control has better performance at the bottleneck since the algorithm is more focus on the bottleneck traffic.

Figure 11 gives a direct comparison of the VSL control effectiveness. It is obvious that the traffic flow is smoother under VSL control (Figure 11(a) and (b)) than the no control case. The average speed is increased and there is less stop

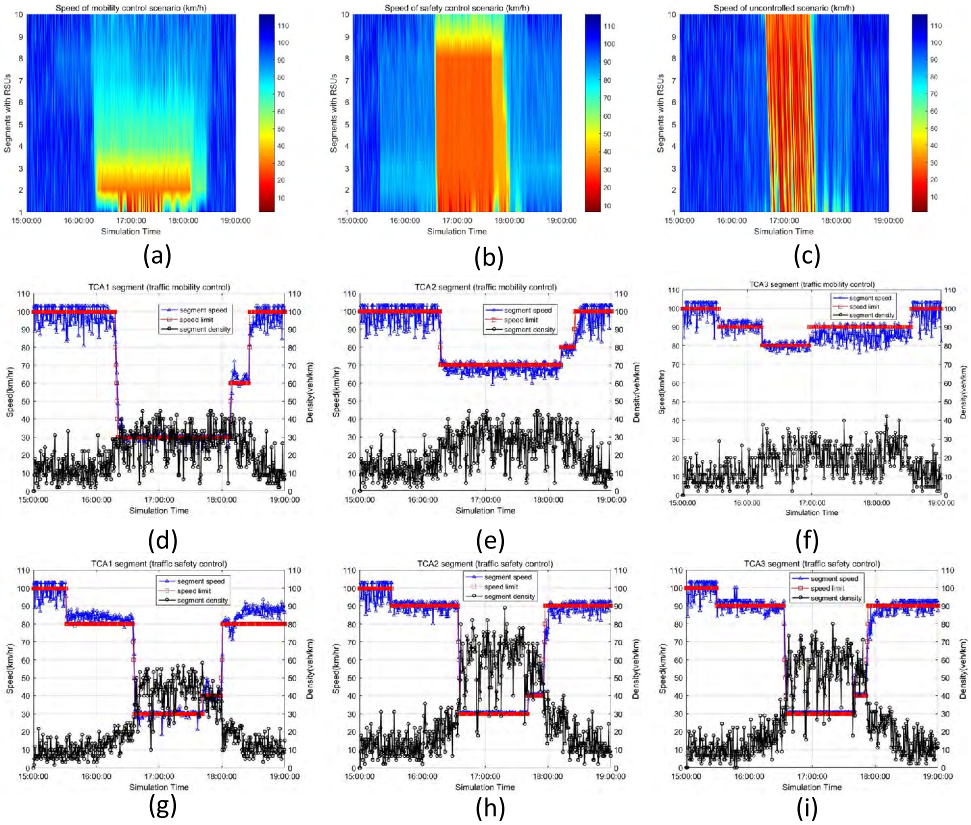


FIGURE 11. The time-space diagram of the freeway corridor with different control strategy (a)-(c), the speed limits of mobility control (d)-(f) and safety control (g)-(i).

TABLE 2. Simulation results of different scenarios.

Scenarios	Freeway section			Freeway bottleneck			
	TTT(h)	Total no. of stops	Mean speed difference (km/h)	TTT(h)	Mean speed (km/h)	SD of speed (km/h)	Average bottleneck outflow (veh/h/ln)
No control	326.9	1020	8.563	37.4	64.936	33.972	1632
Traffic mobility control	229.2 (-29.89%)	0 (-100%)	6.977 (-18.52%)	18.163 (-51.44%)	88.056 (+35.6%)	5.915 (-82.59%)	2036 (+24.77%)
Traffic safety control	305.583 (-6.52%)	0 (-100%)	4.178 (-51.21%)	24.388 (-34.79%)	70.063 (+7.9%)	17.077 (-49.73%)	1795 (+9.98%)

and go phenomenon. Interestingly, result also implied that the traffic under mobility control and safety control is not exactly the same. Traffic under mobility control (Figure 11(d)–(f)) has higher average speed and lower density while the traffic speed under safety control (Figure 11(g)–(i)) is more uniform. Additionally, mobility controllers activated earlier than the safety controllers, indicates that the density reward is more sensitive and can react earlier to the congestion. Moreover, the mobility control agents behave differently to keep density at a lower level, while the safety control agents tend to adopt the similar policy to keep freeway speed with less volatility. Nevertheless, both control policies can adjust speed limits to suppress the congestion and shockwave earlier than their formation, which is the superiority of the system.

Figure 12 further compared the bottleneck control effects. Figure 12(a) shows that without control there is a congestion. Using the mobility control (Figure 12(b)), the congestion is almost eliminated. Using the safety control (Figure 12(c)), the congestion is relieved to an acceptable level. The fundamental diagram (Figure 12(e) is the traffic mobility control and (f) is the traffic safety control) also indicates that the congestion is relieved in both controlled scenarios. However, we cannot simply say that the mobility control is better than the safety control. Figure 11 shows that the vehicles are more concentrated at the upstream of the bottleneck in mobility control while the vehicles distributed more evenly in safety control. The smaller speed difference between adjacent segments (Figure 12(d)) in safety control also implies that the traffic in safety control is more robust toward

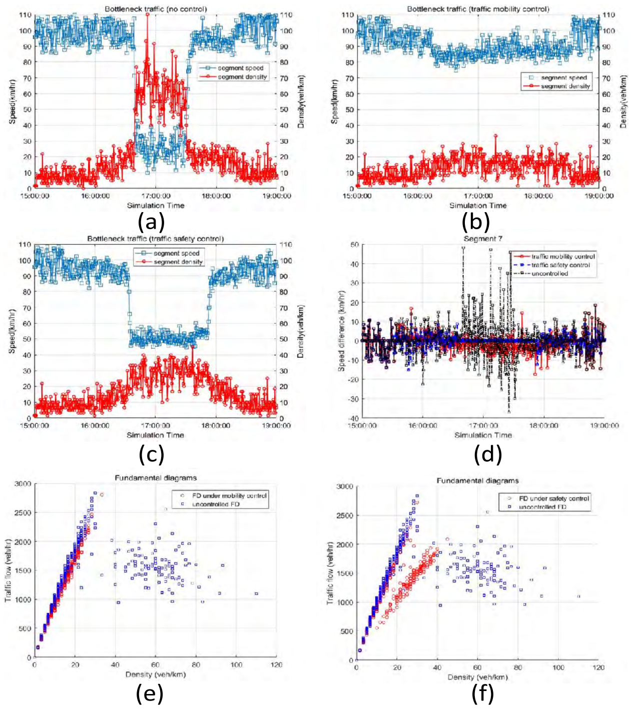


FIGURE 12. **The comparison of the bottleneck traffic with no control and controlled scenarios.**

unexpected disturbances. In summary, it is necessary to choose different control strategies according to different requirements.

VI. CONCLUSION

In this paper, a VSL control system with V2I technology and DQL algorithm is proposed. The aim of the system is to improve traffic mobility and safety. It is a distributed control system, in which agents work synchronously according to their knowledge in a cooperative manner. Agents assume that other agents will choose the best-known actions, thus the calculation complexity is reduced. This paper also presented how to implement V2I technology in the macroscopic traffic control. Besides, we have deduced the safety function that can stabilize freeway traffic, which was seldom discussed in the previous researches. Moreover, we proposed a modified DQL algorithm for multiple traffic controllers' cooperation. Using the DQL algorithm, traffic control on a large network will become more applicable. Simulation results suggested that the proposed control system could effectively relieve traffic congestion on the freeway. Meanwhile, the speed differences between adjacent segments are significantly reduced, which indicated the lower rear-end collision risk. Results also showed that the control system could act proactively before the congestion emerges. An interesting phenomenon found in this study is that there can be different optimized traffic equilibriums toward different control objectives. As a result, more control strategies can be applied to exam their performances.

This work is our first step to study the control effects and traffic characteristics involving both reinforcement learning and V2I technology. The results are promising but more work is necessary for the improvement. First, we will stretch the range of the system to further upstream. By this way, the traffic can response to the congestion earlier and the average speed may increase. Second, we will integrate it with ramp metering control. The advantage of ramp metering is it does not create an ''artificial congestion'' which may degrade upstream traffic condition. However, the waiting queue on ramp is limited thus the trade-off between bottleneck outflow and queue length need to be considered carefully. Third, we have found that deep reinforcement learning approach is another potential technic that can be combined with the V2I environment. It can solve some inherent flaws of traditional RL approaches and can efficiently tackle with continuous state. Future study could also consider the communication latency and sensor faults in the system. In addition, we will integrate the control system with the prevailing vehicle-to-vehicle controllers and test the performance. Moreover, the proposed DQL algorithm will be compared with other elaborate control strategies.

REFERENCES

[1] B. Khondaker and L. Kattan, ''Variable speed limit: An overview,'' *Transp. Lett.*, vol. 7, no. 5, pp. 264–278, Oct. 2017.

[2] M. D. Hadiuzzaman and T. Z. Qiu, ''Cell transmission model based variable speed limit control for freeways,'' *Can. J. Civil Eng.*, vol. 40, no. 1, pp. 46–56, Jan. 2013.

[3] Z. Li, P. Liu, W. Wang, and C. Xu, ''Development of a control strategy of variable speed limits to reduce rear-end collision risks near freeway recurrent bottlenecks,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 866–877, Apr. 2014.

[4] A. Hegyi, B. De Schutter, and J. Hellendoorn, ''Optimal coordination of variable speed limits to suppress shock waves,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 1, pp. 102–112, Mar. 2005.

[5] A. Hegyi, M. Burger, B. De Schutter, J. Hellendoorn, and T. J. J. van den Boom, ''Towards a practical application of model predictive control to suppress shock waves on freeways,'' in *Proc. Eur. Control Conf.*, Jul. 2007, pp. 1764–1771.

[6] R. C. Carlson, I. Papamichail, M. Papageorgiou, and A. Messmer, ''Optimal mainstream traffic flow control of large-scale motorway networks,'' *Transp. Res. C, Emerg. Technol.*, vol. 18, no. 2, pp. 193–212, Apr. 2010.

[7] M. Abdel-Aty, J. Dilmore, and A. Dhindsa, ''Evaluation of variable speed limits for real-time freeway safety improvement,'' *Accident Anal. Prevention*, vol. 38, no. 2, pp. 335–345, Mar. 2006.

[8] N. J. Fudala and M. D. Fontaine, ''Interaction between system design and operations of variable speed limit systems in work zones,'' *Transp. Res. Rec.*, vol. 2169, no. 1, pp. 1–10, 2010.

[9] P. Edara, C. Sun, and Y. Hou, ''Evaluation of variable advisory speed limits in congested work zones,'' *J. Transp. Saf. Secur.*, vol. 9, no. 2, pp. 123–145, Apr. 2017.

[10] J. Piao and M. McDonald, ''Safety impacts of variable speed limits—A simulation study,'' in *Proc. 11th IEEE Int. Conf. Intell. Transp. Syst.*, Oct. 2008, pp. 833–837.

[11] S. K. Zegeye, B. De Schutter, J. Hellendoorn, and E. A. Breunesse, ''Variable speed limits for green mobility,'' in *Proc. 14st Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2011, pp. 2174–2179.

[12] E. Grumert and A. Tapani, ''Impacts of a cooperative variable speed limit system,'' *Procedia-Social Behav. Sci.*, vol. 43, pp. 595–606, Jan. 2012.

[13] B. Khondaker and L. Kattan, ''Variable speed limit: A microscopic analysis in a connected vehicle environment,'' *Transp. Res. C, Emerg. Technol.*, vol. 58, pp. 146–159, Sep. 2015.

[14] M. Wang, W. Daamen, S. P. Hoogendoorn, and B. V. Arem, "Connected variable speed limits control and car-following control with vehicle-infrastructure communication to resolve stop-and-go waves," *J. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 559–572, Nov. 2016.

[15] M. Wang, W. Daamen, S. P. Hoogendoorn, and B. V. Arem, "Rolling horizon control framework for driver assistance systems. Part II: Cooperative sensing and cooperative control," *Transp. Res. C, Emerg. Technol.* vol. 40, pp. 290–311, Mar. 2014.

[16] Z. Li, P. Liu, C. Xu, H. Duan, W. Wang, "Reinforcement learning-based variable speed limit control strategy to reduce traffic congestion at freeway recurrent bottlenecks," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3204–3217, Nov. 2017.

[17] F. Zhu and S. V. Ukkusuri, "Accounting for dynamic speed limit control in a stochastic traffic environment: A reinforcement learning approach," *Transp. Res. C, Emerg. Technol.*, vol. 41, pp. 30–47, Apr. 2014.

[18] K. Rezaee, B. Abdulhai, and H. Abdelgawad, "Application of reinforcement learning with continuous state space to ramp metering in real-world conditions," in *Proc. 15th Int. IEEE Conf. Intell. Transp. Syst.*, Sep. 2012, pp. 1590–1595.

[19] K. Rezaee, B. Abdulhai, and H. Abdelgawad, "Self-learning adaptive ramp metering: Analysis of design parameters on a test case in Toronto, Canada," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2396, pp. 10–18, 2013.

[20] A. Fares and W. Gomaa, "Multi-agent reinforcement learning control for ramp metering," *Progress in Systems Engineering*. Cham, Switzerland: Springer, 2015, pp. 167–173.

[21] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, "Multiagent reinforcement learning for integrated network of adaptive traffic signal controllers (MARLIN-ATSC): Methodology and large-scale application on downtown Toronto," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1140–1150, Sep. 2013.

[22] L. Kuyer, S. Whiteson, B. Bakker, and N. Vlassis, "Multiagent reinforcement learning for urban traffic control using coordination graphs," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Berlin, Germany: Springer, 2008, pp. 656–671.

[23] M. L. Littman, "Friend-or-foe Q-learning in general-sum games," in *Proc. ICML*, vol. 1, Jun. 2001, pp. 322–328.

[24] L. Buşoniu, R. Babuška, B. De Schutter, "Multi-agent reinforcement learning: An overview," in *Innovations in Multi-Agent Systems and Applications—1*. Berlin, Germany: Springer, 2010, pp. 183–221.

[25] M. Lauer and M. Riedmiller, "An algorithm for distributed reinforcement learning in cooperative multi-agent systems," in *Proc. 17th Int. Conf. Mach. Learn.*, 2000, pp. 535–542.

[26] F. L. D. Silva, R. Glatt, and A. H. R. Costa, "MOO-MDP: An object-oriented representation for cooperative multiagent reinforcement learning," *IEEE Trans. Cybern.*, vol. 49, no. 2, pp. 567–579, Feb. 2017.

[27] J. de Lope, and D. Maravall, "Robust high performance reinforcement learning through weighted k-nearest neighbors," *Neurocomputing* vol. 74, no. 8, pp. 1251–1259, Mar. 2011.

[28] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 4, pp. 271–281, Dec. 2002.

[29] (19. Jan, 2019). *MOTUS. 2019. MOTUS: Microscopic Open Traffic Simulation*. [Online]. Available: http://homepage.tudelft.nl/05a3n/

[30] W. J. Schakel, V. L. Knoop, and B. V. Arem, "Integrated lane change model with relaxation and synchronization," *Transp. Res. Rec.*, vol. 2316, no. 1, pp. 47–57, 2012.

CHONG WANG received the B.S. degree from the Nanjing University of Information Science and Technology, Nanjing, China, in 2009, and the M.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, in 2012. He is currently pursuing the Ph.D. degree in transportation engineering with the Research Center for Internet of Mobility, Southeast University, Nanjing.

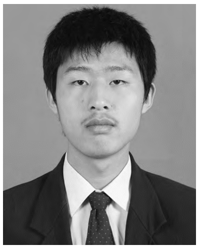
From 2012 to 2014, he was a Software Engineer with the Nanjing Fujitsu Nanda Software Technology Co., Ltd. His research interests include freeway active traffic management, traffic control with connected vehicles, and the freeway traffic simulation.

JIAN ZHANG (M'13) received the Ph.D. degree from Southeast University, Nanjing, China, in 2011.

He is currently the Vice Director of the Research Center for Internet of Mobility, Southeast University. His research interests include transportation application of mobile phone data, connected vehicles and public transportation systems. He is also a member of the American Society of Civil Engineers (ASCE).

LINGHUI XU received the B.S. degree from Chang'an University, Xi'an, China, in 2014. She is currently pursuing the Ph.D. degree with the Research Center for Internet of Mobility, Southeast University.

Her research interests include intelligent transportation systems, connected and automated vehicle, and traffic state prediction.



LINCHAO LI received the M.S. degree from Chang'an University, Xi'an, China, in 2013. He is currently pursuing the Ph.D. degree with the Research Center for Internet of Mobility, Southeast University.

His research interests include the use of machine learning in applications of transportation and predict the traffic state.

BIN RAN received the Ph.D. degree from the University of Illinois, Chicago, IL, USA, in 1993.

He is currently a Professor with the Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI, and the Director of the Research Center for Internet of Mobility, Southeast University, Nanjing, China. He is one of the co-founders of the Chinese Overseas Transportation Association (COTA) and he was the first chairman. He has authored or coauthored more than 90 articles on international journals, including *Transportation Science*, *Transportation Research Part B*, and *Transportation Research Part C*.