D-ACC: Dynamic Adaptive Cruise Control for Highways with Ramps Based on Deep Q-Learning

Lokesh Das and Myounggyu Won
Department of Computer Science, University of Memphis, Memphis, TN, United States
{ldas, mwon}@memphis.edu

Abstract— An Adaptive Cruise Control (ACC) system allows vehicles to maintain a desired headway distance to a preceding vehicle automatically. It is increasingly adopted by commercial vehicles. Recent research demonstrates that the effective use of ACC can improve the traffic flow through the adaptation of the headway distance in response to the current traffic conditions. In this paper, we demonstrate that a state-of-the-art intelligent ACC system performs poorly on highways with ramps due to the limitation of the model-based approaches that do not take into account appropriately the traffic dynamics on ramps in determining the optimal headway distance. We then propose a dynamic adaptive cruise control system (D-ACC) based on deep reinforcement learning that adapts the headway distance effectively according to dynamically changing traffic conditions for both the main road and ramp to optimize the traffic flow. Extensive simulations are performed with a combination of a traffic simulator (SUMO) and vehicle-to-everything communication (V2X) network simulator (Veins) under numerous traffic scenarios. We demonstrate that D-ACC improves the traffic flow by up to 70% compared with a state-of-the-art intelligent ACC system in a highway segment with a ramp.

I. INTRODUCTION

An adaptive cruise control (ACC) is a control system that allows vehicles to keep a desired headway distance to the front vehicle by adjusting acceleration and deceleration using radar, lasers, or cameras [1]. More and more vehicles are equipped with ACC, increasing the market penetration rate every year [2]. It is becoming standard equipment in many recently available commercial vehicles [3]. While ACC is primarily designed for improving driving comfort, numerous research has been performed to demonstrate its benefits on improving traffic flow [4]. A general consensus of these works based on microscopic simulation [5], [6] or macroscopic modelling [7], [8] is that ACC has strong potential to improve traffic flow, and poor configurations of ACC settings may result in degraded traffic flow [7]. However, most of those works focus only on demonstrating the effect of ACC on traffic flow and do not provide an adaptive solution that configures ACC settings to maximize traffic flow in response to dynamically changing traffic conditions.

There are some efforts to optimize ACC settings to maximize traffic flow [9], [10]. However, these works are based on off-line optimization which does not cope well with dynamically changing traffic conditions. There are several methods that address this limitation by adjusting the ACC setting adaptively in real time according to dynamic

traffic conditions [11], [12], [13]. Specifically, Shakel and Arem [11] develop an in-car advisory system that changes the headway distance based on a traffic state prediction model (*i.e.,* the extended generalized Treiber–Helbing filter (EGTF)) [14]. This model determines between free flow and congestion based on the speed data collected from a detector that are aggregated over 1 min. Goni-Ros *et al.* develop an adaptive ACC that adjusts the ACC setting based on the speed of preceding vehicles [13]. Bekiaris-Liberis and Delis propose a similar approach that changes the ACC setting in real time based on the average speed and vehicle density [12]. Spiliopoulou *et al.* design a threshold-based adaptive ACC that adjusts the ACC setting of a connected vehicle based on a threshold of traffic flow measured using the speed of surrounding vehicles [15]. In line with this research, the most recent study at the point of writing this paper is [2]. Essentially, this paper extends [15] by adding acceleration bound changes considering that latest vehicles allow drivers to choose the acceleration strength inspired by Yuan *et al.* [16]. While these model-based approaches enable real-time adaptation of ACC settings according to the current traffic conditions, such static models do not represent effectively the traffic dynamics especially for general highways with ramps (including on/off ramps).

In this paper, we propose a novel AI (Artificial Intelligence)-based approach called the Dynamic Adaptive Cruise Control (D-ACC) System for real-time adaptation of ACC settings in response to dynamically changing traffic conditions. To enable vehicles to adapt the ACC setting more effectively in a fine-grained manner according to the current traffic conditions, we adopt deep reinforcement learning (RL) which is known to enable effective decision making in such a complex environment especially for autonomous vehicles [17], [18]. Specifically, we formulate the problem of determining the optimal ACC setting based on the traffic condition information received via vehicle-to-everything communication (V2X) as a Markov Decision Process (MDP) framework [19]. We then solve the problem by designing a deep Q-network [20] that addresses the challenge of dynamic adaptation of the ACC setting in complex environments represented by a large and continuous state space. In particular, considering that merging/exiting traffic is one of the major causes of traffic perturbations on highways [21], [22], we demonstrate the high adaptability of D-ACC under general highway scenarios with on/off ramps.
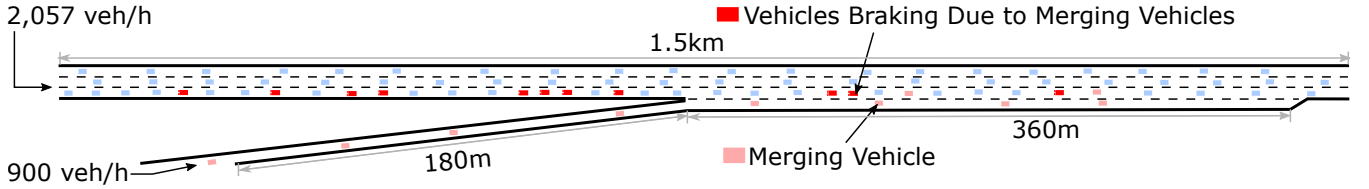
Fig. 1. A snapshot of simulation demonstrating vehicles equipped with a state-of-the-art real-time ACC [2] interfering with merging traffic.
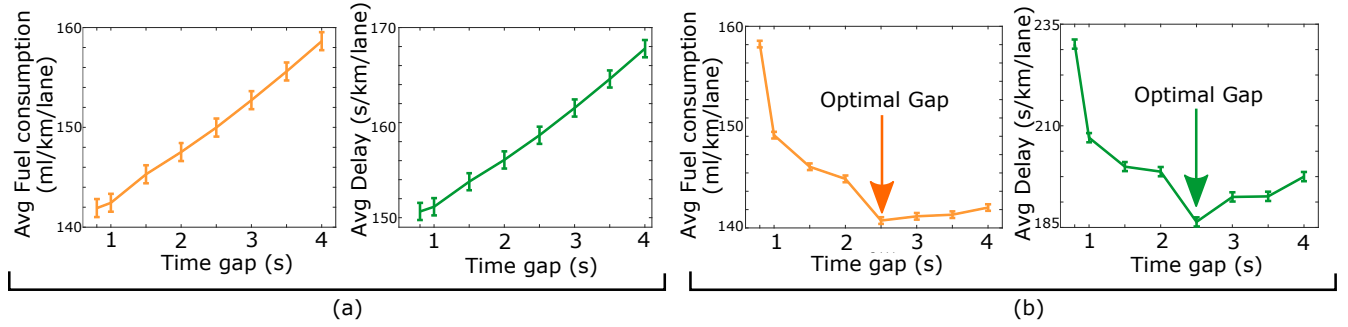

Fig. 2. Average fuel consumption and average delay (a) without merging traffic, (b) with merging traffic.

Extensive simulations are conducted using a combination of a microscopic road traffic simulator SUMO [23] and a V2X network simulator Veins [24] to train the deep Q-network and evaluate the performance of D-ACC. We demonstrate that D-ACC improves traffic flow represented by the average vehicle speed by up to 70% compared to a state-of-the-art real-time ACC system [2]. The contributions of this paper are summarized as follows.

- We propose a novel AI-based ACC system that adaptively configures the ACC settings in response to dynamically changing traffic conditions, overcoming the limitations of the static model-based approaches.
- We formulate the problem of dynamic adaptation of ACC settings as a MDP framework in reference to the transportation literature. A deep-Q network is designed to allow vehicles to make informed decisions on adapting the ACC setting more effectively under complex and dynamically changing traffic conditions.
- Extensive simulations are conducted under general highway scenarios with ramps. We demonstrate that D-ACC outperforms a state-of-the-art intelligent ACC system in various traffic conditions with varying the penetration rates, traffic density of both the ramp and main lane, and length of the acceleration lane, and lane-changing behavior.

This paper is organized as follows. In Section II, we conduct a motivational study to demonstrate that a state-of-the-art real-time ACC system performs poorly due to not appropriately accounting for the dynamics of traffic on a ramp. We then present the details of the proposed D-ACC including the design of a MDP framework and our deep Q-network for optimizing traffic flow in Section III. The simulation settings and results are presented in Section IV. Finally, we conclude in Section V.

II. MOTIVATION

In this section, we demonstrate that a state-of-the-art ACC system does not perform well in determining the optimal ACC setting for highways with ramps. This motivational study is conducted using SUMO [23] and Veins [24]. The Veins, a simulation framework for vehicular network based on OMNeT++ [25], is integrated with SUMO to simulate V2X for vehicles to collect traffic information. Specifically, a highway segment with an on-ramp is considered where vehicles adjust their gaps to the preceding vehicle using a state-of-the-art ACC system [2]. All parameters used for this simulation study are summarized in Table II in Section IV-A. Fig. 1 illustrates a snapshot of simulation where merging vehicles are causing traffic perturbations because of the small headway gap suggested by the state-of-the-art ACC system [2] without considering the traffic on the ramp.

To understand why existing intelligent ACC systems do not perform well for highways with ramps, we vary the headway distance and measure the average fuel consumption and average vehicle delay (metrics selected according to [2]) in both scenarios with and without merging traffic. Fig. 2(a) shows the results. When there is no merging traffic, a small headway distance improves fuel efficiency and delay. This is because a smaller headway gap allows more vehicles to pass the highway segment given the same amount of time. These results agree with the findings of state-of-the-art intelligent ACC systems [15], [2].

However, these model-based ACC systems perform poorly when there is merging traffic as shown in Fig. 2(b). It is observed that the fuel efficiency and delay actually degrade as the headway gap decreases. The reason is because a small headway gap makes lane changing for merging vehicles very difficult, and a lane change can cause the following vehicles to apply strong braking, leading to higher fuel consumption and delay. A very interesting observation is

that when the headway gap is too large, the fuel efficiency and delay start to degrade, *i.e.,* there is a "sweet spot" for the headway gap that maximizes traffic flow, as indicated by the two arrows in Fig. 2(b). We also observe that such a sweet spot changes when the traffic conditions change over time. This motivational study suggests that a new ACC system should be developed that adjusts the ACC setting adaptively by taking into account the dynamically changing traffic conditions of both the main road and ramp.

III. DYNAMIC ADAPTIVE CRUISE CONTROL (D-ACC)

A. Overview

D-ACC dynamically adjusts the headway distance to maximize traffic flow of a highway with a ramp. It can be used by any vehicle equipped with a standard ACC system and V2X capability. Note that although the V2X technology has yet to be widely adopted, D-ACC can be implemented with a smartphone that can communicate with a roadside unit (or a remote traffic server via a cellular network) to collect real-time traffic information needed to determine the headway distance. The smartphone is connected to the vehicle's default ACC system via the OBD port (*e.g.,* Comma.AI [26]) to control the headway distance. The real-world implementation and experimentation of D-ACC is our future work. In this paper, we focus on the system design and verification of performance via simulation.

As a vehicle equipped with D-ACC approaches a highway segment, it starts to communicate with a RSU such as the one that is already available in latest intelligent transportation systems (ITS) [27] via V2X. The RSU is used to collect real-time traffic information and broadcast it to approaching vehicles. In this work, we utilize possibly all relevant information to make a decision for controlling the headway distance including (1) traffic density of the main lanes, (2) average vehicle speed of the main lanes, (3) traffic density of the ramp, (4) average vehicle speed of the ramp, and (5) the length of the ramp. A vehicle receives such information wirelessly and runs D-ACC to determine the optimal headway distance to maximize traffic flow.

B. Markov Decision Process (MDP) Framework

A Markov decision process (MDP) [19] represents the basis for modeling the decision making behaviors for D-ACC. We formulate the dynamic decision making process of D-ACC as MDP in which D-ACC-equipped vehicles take actions of adjusting the headway distance in response to dynamically changing traffic conditions.

We represent MDP as a 4-tuple $M = < S, A, P_{sa}, R >$, where S is a set of states; A is a set of actions; P_{sa} is the probability for the next state given action $a \in A$ and the current state $s \in S$; and R is the reward function, *i.e.,* $r(s'|s, a)$ is the reward for transitioning from a state s to a new state s' due to an action a. In particular, $\pi_t(a|s)$ is a policy of time t that represents the probability of making an action a given state s. The objective is to find the optimal policy π^* that makes the cumulative sum of the expected reward:

$$Q_\pi(s,a) = \mathbb{E}_\pi[\sum_{k=0}^{\infty} \gamma^k r_{t+k}|s,a] \qquad (1)$$

is maximized over the long run, *i.e.,* $Q_\pi^*(s,a) = \max_\pi Q_\pi(s,a)$, where $\gamma \in (0,1]$ is a discount factor.

1) State Space: The state space S of MDP is designed to include the following traffic parameters being motivated by [28]: (a) traffic density of the main road, (b) traffic density of the ramp, (c) average vehicle speed of the main road, (d) average vehicle speed of the ramp, and (e) length of the ramp. Specifically, Daamen *et al.* note that the traffic density of the main road as well as the ramp influence the vehicle merging behaviors [28]. For example, if the traffic density of the ramp is high, then a larger headway distance would be needed to allow vehicles to join quickly so more vehicles can merge into the main lane. It is also noted in [28] that the vehicle speed, in particular, the speed difference between the vehicles on the ramp and the vehicles on the main road influences the efficiency of merging behavior [28]. For example, if the speed of the vehicles on the ramp is relatively higher, a large gap should be created to ensure safer lane change. Finally, the length of a ramp is incorporated into our state space because a longer ramp would allow vehicles to take more time to complete merging/exiting maneuvers.

2) Action Space: The action space A of MDP is specifically designed for vehicles in the main lane to adjust the headway distance dynamically to allow vehicles to merge/exit more efficiently and consequently improve traffic flow. More precisely, an action performed by a vehicle is defined as setting the headway distance to a certain value selected from a range of possible headway distances. It is worth to mention that D-ACC allows vehicles to perform an action before they reach a highway segment with a ramp. This is motivated by recent research [13] demonstrating that creating a certain inter-vehicle gap before it reaches a congested (or potentially congested) area not only allows vehicles to merge/exit more smoothly but also make the following vehicles decelerate appropriately, reducing the inflow into the congested area and preventing from worsening the traffic congestion.

3) Reward Function: The design of the reward function of MDP is focused on maximizing traffic flow. More specifically, the reward function is designed to decrease the average vehicle delay required to pass a highway segment. Thus, in computing the reward value R, the traffic congestion speed of the highway segment $v_{congestion}$ and the length of the highway segment $l_{segment}$ received from a RSU via V2X are used in computing the reward value as follows.

$$R = \begin{cases} +1 & \text{if average delay} \leq \frac{l_{segment}}{v_{congestion}} \\ -1 & \text{if average delay} > \frac{l_{segment}}{v_{congestion}} \end{cases} \qquad (2)$$

C. Deep Q-Network

The optimal policy π^* can be found using the Q-function $Q_\pi^*(s,a) = \max_\pi Q_\pi(s,a)$. Specifically, the Q-function follows the Bellman optimality equation:

$$Q^*(s,a) = \mathbb{E}[r + \gamma \max_{a' \in A} Q^*(s',a')|s,a], \qquad (3)$$

which allows us to determine the optimal policy by easily representing the Q-function as a tabular form when the state space is discrete and finite [19]. However, if the state space is large and continuous, it becomes easily untractable [20]. As such, to account for the large and continuous state space for D-ACC to incorporate wide ranges of vehicle density, speed, and length of a ramp, *etc.*, we adopt the deep Q-network (DQN) algorithm to obtain the optimal policy [20]. The idea is to approximate the Q-function using a neural network [29] where the input to the neural network is the states, and the output is the Q-value for each action. Specifically, the approximate value function for deep Q-network is now denoted by $Q(s,a;\theta_i)$ where θ_i are the weights of the Q-network at i-th iteration. The network is trained to update the weights θ_i with the loss function: $L_i(\theta_i) = \mathbb{E}_M[(r + \gamma \max_{a'} Q(s',a';\theta_i^-) - Q(s,a;\theta_i))^2]$, where θ_i^- is the parameter of the network at i-th iteration, which is updated based on the weights of the Q-network θ_i.

Q-Network Parameters	
Neural network architecture	4 hidden layers of size 8, 12, 20, and 16 respectively
Activation functions	Rectified linear units (ReLU)
Replay buffer size	20k samples
$(\gamma, \epsilon_0, \epsilon_{min}, \lambda_{decay})$	(0.95, 1.0, 0.001, 0.9995)
Batch size	32
Loss function	Mean square error (MSE)
Optimization method	Stochastic Gradient Descent (SGD) with learning rate 0.001
Target network update frequency	1k episodes

TABLE I

PARAMETERS USED FOR OUR DEEP Q-NETWORK.

For our deep Q-network, we design a neural network consisting of 4 hidden layers of size 8, 12, 20, and 16, respectively with activation functions of rectified linear units (ReLU). Other specific parameters used for designing our deep Q-network through trial and

Fig. 3.   Rewards of the reward function.

errors are summarized in Table I. In particular, to balance between exploration and exploitation, we adopt an ϵ-greedy policy [30] in training our deep Q-network. Specifically, with the probability of ϵ, an action is randomly selected from the action space; and with the probability of $1 - \epsilon$, the optimal action is selected based on the greedy method. We allow the value of ϵ to decrease gradually as the algorithm iterates, *i.e.,* $\epsilon = \max(\epsilon_0 \cdot \lambda_{decay}, \epsilon_{min})$. The parameters for
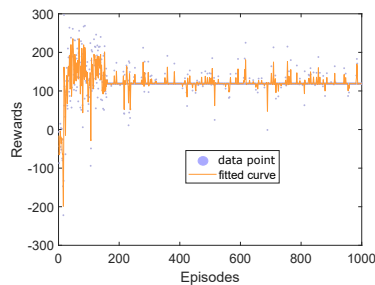
the ϵ-greedy policy are also summarized in Table I. Fig. 3 displays the results of our reward function with the number of episodes up to 1K, demonstrating the fast convergence of the reward function.

IV. SIMULATION RESULTS

A. Simulation Setup

Vehicle Parameters	
Action interval	1 s
Vehicle length	4 m
Min headway	2.5 m
Lane change model	LC2013 [31]
Max acceleration	2.6 m/s^2
Max deceleration	4.5 m/s^2
Traffic Parameters	
Traffic volume (main lane)	900~3700 veh/h (default: 2057 veh/h)
Traffic volume (merging lane)	200~900 veh/h (default: 900 veh/h)
Highway length	1.5 km
Merging lane length	360 m
Acceleration lane length	50~180 m (default: 180m)

TABLE II

PARAMETERS FOR THE SIMULATION.

A simulation study is performed to evaluate the performance of D-ACC using a traffic simulator (SUMO) integrated with a network simulator (Veins). Simulations are conducted using a PC equipped with a 1.4GHz quad-core Intel Core i5 CPU and 8GB of RAM running on MacOS. Deep reinforcement learning for D-ACC is implemented in Python using Keras and Tensorflow [32] which is interfaced with SUMO via Traffic Control Interface (TraCI) [33].

In this simulation, we consider a highway segment with a ramp. The length of the highway segment, ramp, and acceleration lane are 1.5km, 360m, and (40~180)m, respectively. Vehicles on the main road are generated at a rate of 900 to 3700 veh/h, and that for the ramp is varied between 200 and 900 veh/h to evaluate the effect of the on-ramp traffic density. Each vehicle with a length of 4m performs a lane change according to the lane-changing model [31]. Each vehicle updates the headway distance every second in response to dynamically changing traffic conditions. We ensure that the headway distance is not decreased below 2.5m for ensuring safety. In increasing or decreasing the headway distance, acceleration of up to 2.6m/s^2 and deceleration of up to 4.5 m/s^2 are used. All traffic parameters and their default values used for this simulation study are summarized in Table II.

The main metric measured in this simulation study is the average vehicle speed of all vehicles in the given highway segment on a single run that effectively represents the traffic flow. We vary *all possible parameters for the scenario* including the penetration rate (*i.e.* a portion of vehicles equipped with D-ACC or a state-of-the-art real-time ACC out of all vehicles), traffic density of the main lane, traffic density of the ramp, length of the acceleration lane, and lane-changing behavior.
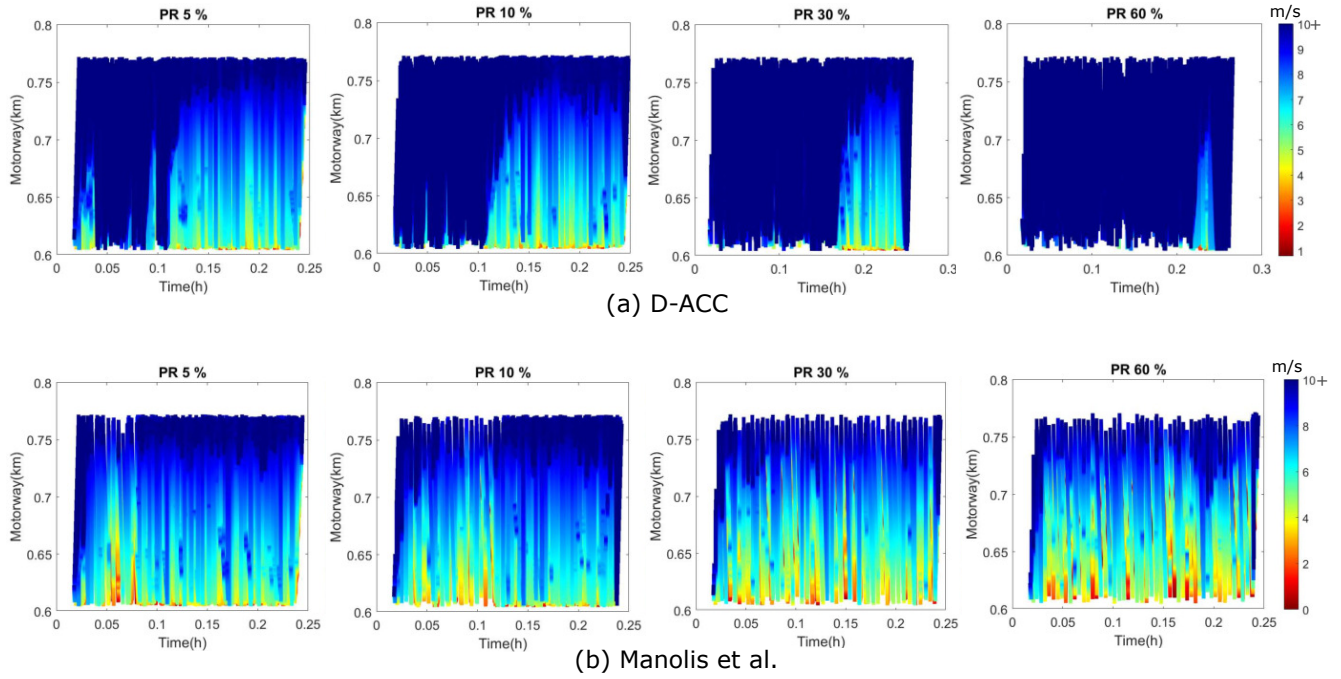
(a) D-ACC

(b) Manolis et al.

Fig. 4. Effect of penetration rate (PR) for (a) D-ACC and (b) Manolis *et al.* [2]. The color represents the vehicle speed in m/s.

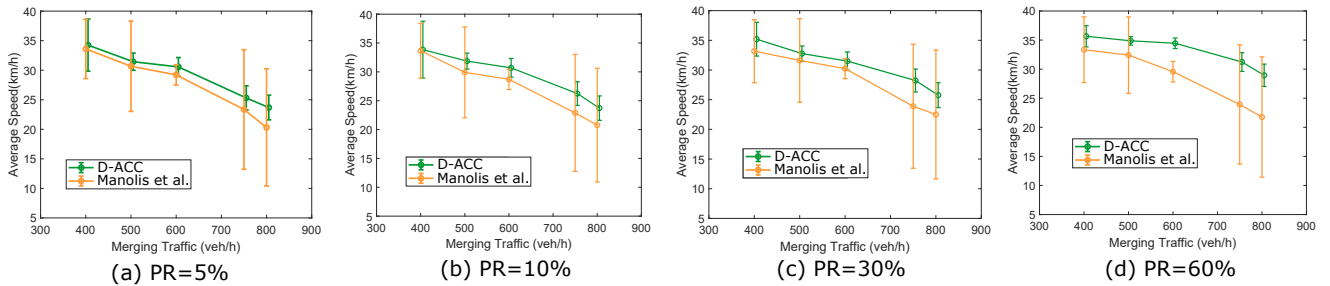(a) PR=5%   (b) PR=10%   (c) PR=30%   (d) PR=60%

Fig. 5. Effect of merging traffic with varying penetration rates.

B. Penetration Rate

We measure the average vehicle speed for both D-ACC and Manolis *et al.* [2] by varying the penetration rate. The results are depicted in Fig. 4 which demonstrate that higher penetrate rates lead to better traffic flow for D-ACC. Interestingly, on the other hand, the opposite results are observed for Manolis *et al.* The reason is that according to Manolis *et al.*, higher penetration rates mean more vehicles trying to maintain smaller headway distance since Manolis *et al.* improve traffic flow by allowing more vehicles to pass the highway segment by using a smaller inter-vehicle gap, which makes merging vehicles difficult to change lanes causing traffic perturbations. Another interesting observation is that D-ACC improves traffic flow by 21% even with a very small penetration rate of 5% compared with Manois *et al.*. More significant improvement in traffic flow is observed with higher penetration rates, *i.e.,* over 30%. Specifically, D-ACC achieves 70% higher average speed compared with Manolis *et al.* when the penetration rate is 60%. In other words, the average speed for D-ACC is increased by 11km/h compared

with Manolis *et al.* when the penetration rate is increased from 5% to 60%. The reason is that with higher penetration rate, more vehicles adjust gaps dynamically to allow merging vehicles to change lanes more smoothly.

C. Traffic Density of Ramp

We evaluate the effect of the traffic density of the ramp on the performance of D-ACC. We measure the average vehicle speed by varying the number of vehicles injected into the merging lane per hour and the penetration rates. The results are depicted in Fig. 5 which demonstrate that as the traffic density of the ramp increases, the average speed for both D-ACC and Manolis *et al.* decreases regardless of the penetration rate. An interesting observation is that D-ACC is more resilient to the increasing traffic density of the ramp than Manolis *et al.* as demonstrated by the increasing gap between the average speed of D-ACC and that of Manolis *et al.* when the traffic density increases for all penetration rates. Most notably, the performance gain increases by 210% as the merging traffic increases from 400 veh/h to 800 veh/h, when the penetration rate is 60%. It is also observed that the

advantages of using D-ACC increase as both the merging traffic and penetration rate increase. Specifically, the gap between the average speed of D-ACC and that of Manolis *et al.* for PR=60% with merging traffic of 800 veh/h is about 140% higher compared with the gap for PR=5% with 800 veh/h. Overall, D-ACC outperforms the state-of-the-art real-time ACC system regardless of the traffic density of the ramp for all penetration rates.

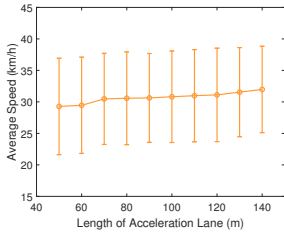### D. Effect of Length of Acceleration Lane

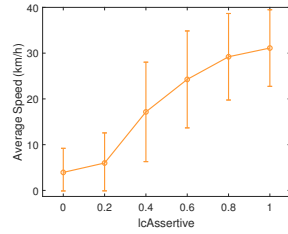Fig. 6.   Effect of the length of the ramp.

Fig. 7.   Effect of the lane-changing behavior.

We evaluate the performance of D-ACC by varying the length of the acceleration lane. We measure the average vehicle speed on the main lane. The results are depicted in Fig. 6. We observe that the average speed increases generally as the length of the acceleration lane increases. The reason is because vehicles on the ramp change lanes more easily with a longer acceleration lane, thereby improving overall traffic flow. Specifically, as we increase the length of the acceleration lane from 50m to 140m, the average speed is increased by 9.2%.

E. Lane Change Behavior

In this section, we analyze the effect of the lane-changing behavior of merging vehicles on the performance of D-ACC. For this simulation, we adopt the default lane-changing model for SUMO [31]. In this lane-changing model, a parameter "lcAssertive" is defined to represent the willingness of a driver to accept lower front and rear gaps on the target lane when executing a lane change [34]. This parameter is selected between 0 and 1 (with the default value of 1 for SUMO), where a greater number represents more aggressive lane-changing behavior.

We measure the average speed by varying the value of the parameter. The results are depicted in Fig. 7. As shown, D-ACC performs better with larger "lcAssertive", *i.e.,* more aggressive lane-changing behavior. The reason is that a larger value of "lcAssertive" allows vehicles to change lanes quickly whenever there is a space available to join, thereby giving more time to the following vehicles in the merging lane to make a lane-change decision. On the other hand, a small value of "lcAssertive" occasionally make merging vehicles being unable to quickly change the lanes, leading to traffic perturbations, especially when those merging vehicles fail to merge as they reach almost the end of the ramp.
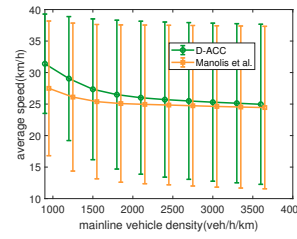

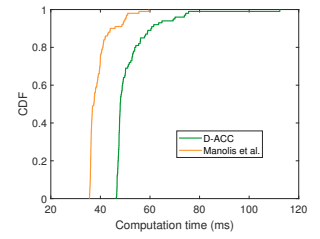Fig. 8.   Effect of the traffic density of the main lane.

Fig. 9.   Computation time.

F. Traffic Density of Main Lane

We evaluate the effect of the traffic density of the main lane on the performance of D-ACC. We measure the average vehicle speed by varying the traffic density. Fig. 8 shows the results. As expected, as the vehicle density increases, the average speed for both D-ACC and Manolis *et al.* decreases. We observe that D-ACC achieves higher average speed than Manolis *et al.* by up to 14.6% especially when the vehicle density is low. An interesting observation is that as the main lane traffic is getting more saturated, the performance gain of D-ACC compared to Manolis *et al.* becomes smaller because merging vehicles cannot change lanes easily when the traffic density is too high.

G. Computation Time

We measure the computation time for obtaining the headway distance for both D-ACC and Manolis *et al.* to demonstrate that D-ACC is capable of adjusting the headway distance quickly in response to dynamically changing traffic conditions. Specifically, the computation time for each vehicle is measured 100 times for both D-ACC and Manolis *et al.* Fig. 9 depicts the cumulative distribution function (CDF) graph for the results. Here we assume that the neural network model for D-ACC has been already established. The average computation time for D-ACC is 64.5% higher than that for Manolis *et al.* partly due to the computation overhead for taking into account various factors for determining the headway distance. However, we note that the average computation time of 51.7ms for D-ACC is sufficient to make a decision at every update interval which is one second.

V. CONCLUSION

We have presented a dynamic adaptive cruise control system (D-ACC) designed to adjust the headway distance adaptively based on reinforcement learning in response to dynamically changing traffic conditions for highways with ramps. We demonstrated that D-ACC improves traffic flow significantly even with a small penetration rate compared with a state-of-the-art real-time ACC system. Our on-going work is to develop a smartphone-based D-ACC, integrate it with an actual vehicle platform and evaluate the performance in real-world traffic environments by recruiting vehicles through crowdsourcing websites such as Amazon Mechanical Turk.

References

[1] A. Vahidi and A. Eskandarian, "Research advances in intelligent collision avoidance and adaptive cruise control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 4, no. 3, pp. 143–153, 2003.

[2] D. Manolis, A. Spiliopoulou, F. Vandorou, and M. Papageorgiou, "Real time adaptive cruise control strategy for motorways," *Transportation Research Part C: Emerging Technologies*, vol. 115, p. 102617, 2020.

[3] M. Makridis, K. Mattas, B. Ciuffo, F. Re, A. Kriston, F. Minarini, and G. Rognelund, "Empirical study on the properties of adaptive cruise control systems and their impact on traffic flow and string stability," *Transportation Research Record*, p. 0361198120911047, 2020.

[4] L. Davis, "Effect of adaptive cruise control systems on traffic flow," *Physical Review E*, vol. 69, no. 6, p. 066110, 2004.

[5] I. A. Ntousakis, I. K. Nikolos, and M. Papageorgiou, "On microscopic modelling of adaptive cruise control systems," *Transportation Research Procedia*, vol. 6, pp. 111–127, 2015.

[6] B. Bayar, S. A. Sajadi-Alamdari, F. Viti, and H. Voos, "Impact of different spacing policies for adaptive cruise control on traffic and energy consumption of electric vehicles," in *Mediterranean Conference on Control and Automation (MED)*, 2016, pp. 1349–1354.

[7] A. I. Delis, I. K. Nikolos, and M. Papageorgiou, "Simulation of the penetration rate effects of acc and cacc on macroscopic traffic dynamics," in *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 336–341.

[8] I. K. Nikolos, A. I. Delis, and M. Papageorgiou, "Macroscopic modelling and simulation of acc and cacc traffic," in *IEEE International Conference on Intelligent Transportation Systems*, 2015, pp. 2129–2134.

[9] C. U. Mba and C. Novara, "Evaluation and optimization of adaptive cruise control policies via numerical simulations." in *International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS)*, 2016, pp. 13–19.

[10] H. Liu, H. Wei, T. Zuo, Z. Li, and Y. J. Yang, "Fine-tuning adas algorithm parameters for optimizing traffic safety and mobility in connected vehicle environment," *Transportation Research part C: Emerging Technologies*, vol. 76, pp. 132–149, 2017.

[11] W. J. Schakel and B. Van Arem, "Improving traffic flow efficiency by in-car advice on lane, speed, and headway," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 4, pp. 1597–1606, 2014.

[12] N. Bekiaris-Liberis and A. Delis, "Feedback control of freeway traffic flow via time-gap manipulation of acc-equipped vehicles: A pde-based approach," *IFAC-PapersOnLine*, vol. 52, no. 6, pp. 1–6, 2019.

[13] B. Goñi-Ros, W. J. Schakel, A. E. Papacharalampous, M. Wang, V. L. Knoop, I. Sakata, B. van Arem, and S. P. Hoogendoorn, "Using advanced adaptive cruise control systems to reduce congestion at sags: An evaluation based on microscopic traffic simulation," *Transportation Research Part C: Emerging Technologies*, vol. 102, pp. 411–426, 2019.

[14] J. Van Lint and S. P. Hoogendoorn, "A robust and efficient method for fusing heterogeneous data from traffic sensors on freeways," *Computer-Aided Civil and Infrastructure Engineering*, vol. 25, no. 8, pp. 596–612, 2010.

[15] A. Spiliopoulou, D. Manolis, F. Vandorou, and M. Papageorgiou, "Adaptive cruise control operation for improved motorway traffic flow," *Transportation Research Record*, vol. 2672, no. 22, pp. 24–35, 2018.

[16] K. Yuan, "Capacity drop on freeways: Traffic dynamics, theory and modeling," *Ph.D. Thesis, Technical University of Delft*, 2016.

[17] Z. Chen and X. Huang, "End-to-end learning for lane keeping of self-driving cars," in *IEEE Intelligent Vehicles Symposium (IV)*, 2017, pp. 1856–1860.

[18] S. Triest, A. Villaflor, and J. M. Dolan, "Learning highway ramp merging via reinforcement learning with temporally-extended actions," in *IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 1595–1600.

[19] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[20] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

[21] A. K. Gupta and V. Katiyar, "Phase transition of traffic states with on-ramp," *Physica A: Statistical Mechanics and its Applications*, vol. 371, no. 2, pp. 674–682, 2006.

[22] S. van Maarseveen, "Impacts of truck platooning at motorway on-ramps," *Technical Report, TuDelft*, 2017.

[23] P. Fernandes and U. Nunes, "Platooning of autonomous vehicles with intervehicle communications in sumo traffic simulator," in *International IEEE Conference on Intelligent Transportation Systems*, 2010, pp. 1313–1318.

[24] C. Sommer, R. German, and F. Dressler, "Bidirectionally coupled network and road traffic simulation for improved ivc analysis," *IEEE Transactions on Mobile Computing*, vol. 10, no. 1, pp. 3–15, 2010.

[25] A. Varga and R. Hornig, "An overview of the OMNeT++ simulation environment," in *International Conference on Simulation Tools and Techniques for Communications, Networks and Systems & Workshops*, 2008, p. 60.

[26] E. Santana and G. Hotz, "Learning a driving simulator," *arXiv preprint arXiv:1608.01230*, 2016.

[27] A. Duret, M. Wang, and A. Ladino, "A hierarchical approach for splitting truck platoons near network discontinuities," *Transportation Research Part B: Methodological*, vol. 132, pp. 285–302, 2020.

[28] W. Daamen, M. Loot, and S. P. Hoogendoorn, "Empirical analysis of merging behavior at freeway on-ramp," *Transportation Research Record*, vol. 2188, no. 1, pp. 108–118, 2010.

[29] A. O. Al-Abbasi, A. Ghosh, and V. Aggarwal, "Deeppool: Distributed model-free algorithm for ride-sharing using deep reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4714–4727, 2019.

[30] M. Wunder, M. L. Littman, and M. Babes, "Classes of multiagent Q-learning dynamics with epsilon-greedy exploration," in *Proceedings of the International Conference on Machine Learning (ICML-10)*, 2010, pp. 1167–1174.

[31] J. Erdmann, "Sumo's lane-changing model," in *Modeling Mobility with Open Data*, 2015, pp. 105–123.

[32] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous distributed systems," *arXiv preprint arXiv:1603.04467*, 2016.

[33] A. Wegener, M. Piórkowski, M. Raya, H. Hellbrück, S. Fischer, and J.-P. Hubaux, "TraCI: an interface for coupling road traffic and network simulators," in *Proceedings of the Communications and Networking Simulation Symposium*, 2008, pp. 155–163.

[34] M. Berrazouane, K. Tong, S. Solmaz, M. Kiers, and J. Erhart, "Analysis and initial observations on varying penetration rates of automated vehicles in mixed traffic flow utilizing sumo," in *International Conference on Connected Vehicles and Expo*, 2019.