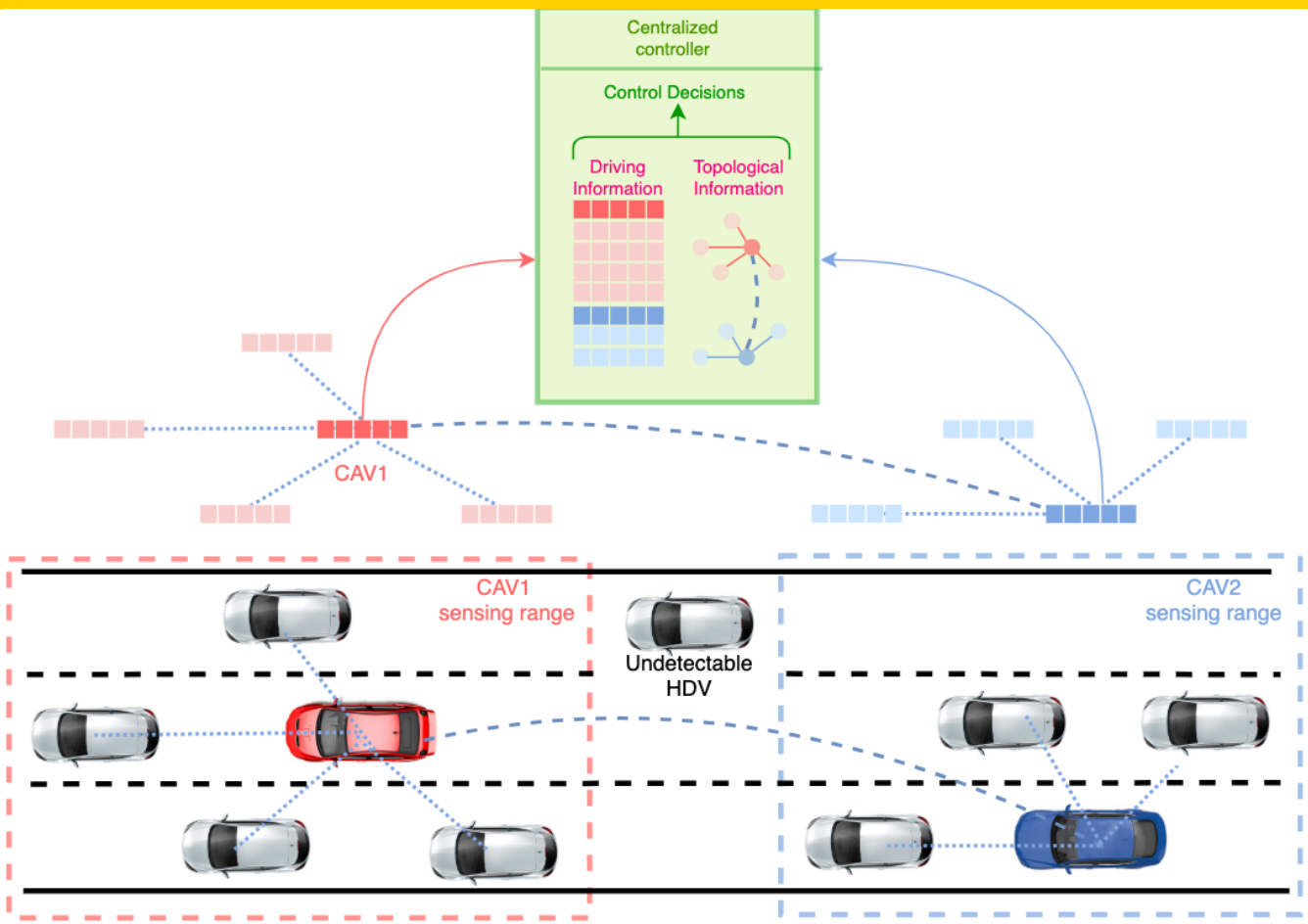
CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Final Report #41
June 2022



Development of AI-based and control-based systems for safe and efficient operations of connected and autonomous vehicles

Jiqian Dong
Runjia Du
Paul (Young Joun) Ha
Sikai Chen
Samuel Labi

PURDUE UNIVERSITY

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

Report No. 41
Project Start Date: January 2019
Project End Date: December 2020

June 2022

Development of AI-based and control-based systems for safe and efficient operations of connected and autonomous vehicles

by

Jiqian Dong
Graduate Researcher

Runjia Du
Graduate Researcher

Paul (Young Joun) Ha
Gradate Researcher

Sikai Chen
Visiting Asst. Professor

Samuel Labi
Professor

Purdue University






ACKNOWLEDGMENTS AND DISCLAIMER

Funding for this research was provided by the Center for Connected and Automated Transportation under Grant No. 69A3551747105 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (OST-R), University Transportation Centers Program. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Dong, J., Du, R., Ha, Y.J., Chen, S., Labi, S. (2022). Development of AI-based and control-based systems for safe and efficient operations of connected and autonomous vehicles, CCAT Report #41, The Center for Connected and Automated Transportation, Purdue University, West Lafayette, IN.

Contacts

For more information

Samuel Labi, Ph.D.
550 Stadium Mall Drive
HAMP G167B
Phone: (765) 494-5926
Email: labi@purdue.edu

CCAT
University of Michigan Transportation Research Institute
2901 Baxter Road
Ann Arbor, MI 48152

uumtri-ccat@umich.edu
(734) 763-2498
www.ccat.umtri.umich.edu

Technical Report Documentation Page

1. Report No. 41	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle: Development of AI-based and control-based systems for safe and efficient operations of connected and autonomous vehicles		**5. Report Date** June 2022 **6. Performing Organization Code** N/A
7. Author(s) Jiqian Dong, Runjia Du, Paul (Young Joun) Ha, Sikai Chen, Samuel Labi		**8. Performing Organization Report No.** N/A
9. Performing Organization Name and Address: Center for Connected and Automated Transportation Purdue University, 550 Stadium Mall Drive, W. Lafayette, IN 47907; and University of Michigan Ann Arbor, 2901 Baxter Road, Ann Arbor, MI 48109		10. Work Unit No. **11. Contract or Grant No.** Contract No. 69A3551747105
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE Washington, DC 20590		**13. Type of Report and Period Covered** Final report, Jan 2019 – Dec 2020 **14. Sponsoring Agency Code** OST-R

15. Supplementary Notes
Conducted under the U.S. DOT Office of the Assistant Secretary for Research and Technology's (OST-R) University Transportation Centers (UTC) program.

16. Abstract
This research is in three parts. The first part, recognizing the range limitations of onboard sensors such as LiDAR and cameras, developed an Artificial Intelligence control system that fuses sensed (local) information and longer-range information to make CAV lane-changing decisions. Then, by integrating both information sources, the CAV is made to fully characterize its surrounding environment to facilitate long-term motion planning and short-term local planning. A Deep Reinforcement Learning based approach is used to provide an end-to-end solution that incorporated a fusion approach and decision processor. The framework can help identify the optimal connectivity range for each domain of prevailing operating traffic density. The second part developed a method to achieve cooperative lane changing at highway ramps. The research investigated the efficacy of roadside units using CAVs in mixed stream traffic. The third part of the research developed a collision avoidance framework for CAVs, to reduce the likelihood of collision with surrounding vehicles, particularly HDVs that drive aggressively or have uncertain or unpredictable behavior. The framework developed in this part of the study is intended to improve operational efficiency without compromising safety unduly, through the use of joint decision-making protocols and sharing of real-time information that is made available via vehicle connectivity.

17. Key Words Autonomous vehicles, Infrastructure, artificial intelligence, control, information fusion, collision avoidance	**18. Distribution Statement** No restrictions.		
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 63	**22. Price**

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

CENTER FOR CONNECTED AND AUTOMATED TRANSPORTATION

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: ARTIFICIAL INTELLIGENCE (AI) CONTROL SYSTEM FOR CAV LANE-CHANGING DECISIONS

1.1 Background

In the transportation domain, connectivity includes vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-cloud (V2C), and other forms of vehicle-to-external (V2X) communication capabilities. It has been postulated that connectivity technology will greatly benefit the safety and efficiency of vehicle operations (Elliott et al., 2019) by promoting greater awareness of the driving environment, and therefore, will facilitate proactive actions to enhance driving performance (FHWA, 2015). We present Figure 1 (below) to illustrate this concept. In the figure, the vehicle of interest (also termed the "ego" vehicle) is denoted in red color. The figure presents a situation where the ego vehicle is presented an opportunity to exploit its connectivity capabilities to make safer rational driving decisions. If the ego vehicle is capable of accessing information only from other vehicles within its immediate vicinity (that is, its sensing range), it will likely decide to stay in its lane (Lane 1) because the (white) vehicle ahead of the ego vehicle is moving at a higher speed compared to the vehicle in lane 2 (blue vehicle) when $v_2 > v_1$, all in the vicinity of the ego vehicle. Assume that further downstream in Lane 1, there exists an imminent hazard associated with different infrastructure settings or traffic situation (for example, a crash site, entry ramp, disabled vehicle, or workzone). If the ego vehicle's sources of information are limited to its local area only, it will be unable to characterize these imminent conditions downstream, and will continue driving until it reaches the threat, whereupon it will need to decelerate sharply or undertake some evasive maneuver. On the other hand, if the ego vehicle's sources of information include connected vehicles sources located further downstream, then it will be able to sense the imminent situation well before it reaches the threat location, and therefore will make an early decision to decelerate while in Lane 1, merge into lane 2, or both.

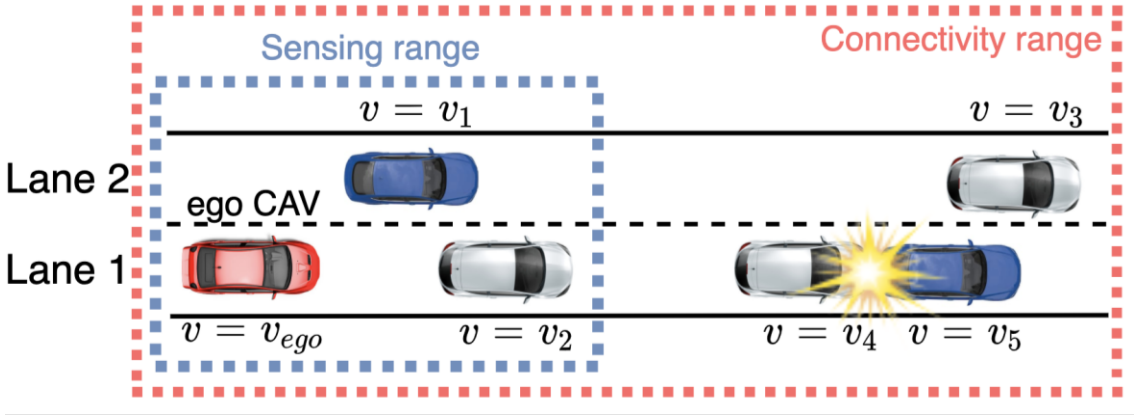


Figure 1.1 Conceptual ranges of sensing and connectivity, in lane-changing situation (the "Ego" vehicle or the CAV of interest, is colored red)

It has been postulated that the combined effect of automation and connectivity will yield benefits that exceed the sum of the individual benefits of these technologies. In this project, we do not investigate this hypothesis or measure the synergistic effect of these two technologies. Nevertheless, we duly recognize that the coupling of connectivity and automation can accentuate vastly the benefits of the latter.

1.2 Motivation

In addressing this issue in the context of CAV operations, this research project makes three main contributions. First, it develops a DRL-based model (using modified a Deep Sets procedure) that integrates information that is locally-obtained and system-wide information collected using connectivity capabilities of the vehicles. Secondly, the work develops an end-to-end framework that uses the fused information to control the CAVs lane-changing decisions in a manner that eliminates the chances of collision. Thirdly, the work assesses the effect of traffic density on the sufficiency of the connectivity range and provides an indication of the connectivity threshold to ensure desirable operational performance (in terms of travel efficiency, safety and comfort) of the CAV.

In this project, we show how these contributions reinforce the justification not only for having connectivity in prospective autonomous vehicles, but also for installing connectivity capabilities in existing human-driven vehicles particularly during the transition period when the traffic stream is shared by CAVs and connected HDVs. It is anticipated that such justification will resonate well in the realms of the state of practice and the state of the art. This is because transportation agencies, as stewards of the public road infrastructure, have a fiduciary stake in ensuring road system efficiency, providing real-time information to road users, and monitoring performance of the taxpayer funded road infrastructure system. To these agencies, these results may provide motivation to establish policies that promote connectivity capabilities in HDVs and ultimately, realize these systemwide benefits. In offering this potential contribution, this work hopefully provides a platform upon which stakeholders can realize the benefits of system connectivity to CAV operations, in terms of the CAV's operational efficiency and the optimal range of connectivity.

The remainder of the report 1 is organized as follows: The study methodology section describes the DRL basics, proposed method and the model architecture. The experiment settings section presents the DRL settings and the details of the implementation on a simulated test track. The results section compares our proposed model with other baseline models and uses a case study to identify the critical connectivity range for a given set of traffic conditions. Also in this section, we demonstrate the practical limitations of the classic Deep Set Q learning method proposed by (Huegle et al., 2019) in terms of model transferability across different scenarios of traffic density. We recognize the potential limitations of reinforcement learning in general including the problem of domain adaptation issue, deviation between the real environment and simulation, uncertainty of guaranteed safety performance, and relatively low transparency.

1.3 Methods

In the standard paradigm of reinforcement learning, an agent can explore the environment and subsequently learn a behavior that promotes desired outcomes and avoids undesired outcomes (Mousavi et al., 2018). In this process, the agent observing the current states, takes action, and receives feedback (a positive or negative reward) from the environment (which is the driving space, in the context of this project). The agent evaluates the feedback signal, and understands the benefits (positive reward) of good actions and the (negative reward) of errant actions. In this project, we use reinforcement learning to facilitate safe and efficient movements of the CAV within in a simulation environment (Figure 2).

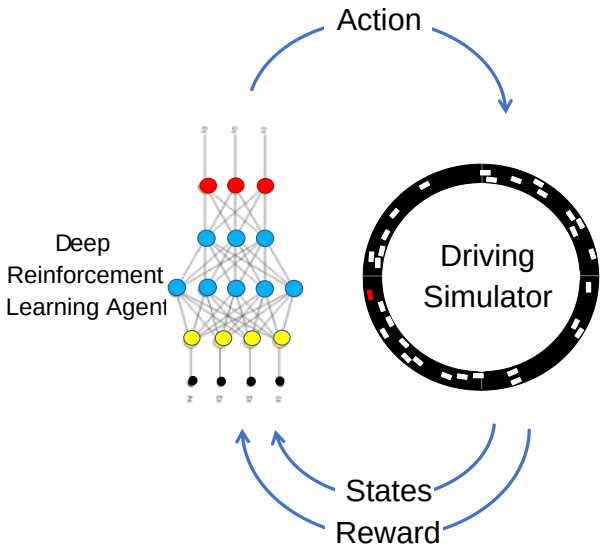


Figure 1.2 Reinforcement learning in the context of the CAV driving simulation

1.3.1 Deep Q learning

Time steps represent an essential feature of reinforcement learning processes in general. In a typical learning process, at each step, *t*, the learning agent undertakes an action, a_t, on the basis of (1) a policy network $\pi_\theta(a_t|s_t)$, which is parametrized as θ , and (2) a current state of "nature", s_t. The agent carries out action a_t and consequently enters a different state s_{t+1} in accordance with the state transition distribution $p(s_{t+1}|s_t, a_t)$, and earns a reward r_t. Reinforcement learning seeks to learn an optimal policy network π_{θ^*} with $\theta^* = argmax_\theta \mathbb{E}[\sum_t r(s_t, a_t)]$. This enables the agent to earn a maximum sum of rewards between the time $t = 0$ to the time at the conclusion of the training episode. In this work, we adopt broadly, the Q learning, a model-free method for purposes of identifying the optimal driving policy. The Q function $Q^\pi(s_t, a_t) = \mathbb{E}_{a_{t'>t} \sim \pi_\theta}[r(s_t)|a_t, s_t]$ is a representation of the total expected reward from time *t* after choosing the action a_t, over the entire trajectory. The Q

function not only provides an easy way to evaluate how "good" the choice of a_t value is, but also gives guidance on the choice of a driving policy that yields a maximum value of the Q function. Recognizing the inherent difficulty of expressing the Q function in an explicit manner, we use a deep neural network technique to yield an approximation of the Q function (this is termed a classical Deep Q Network (DQN) method, which was also applied in (Jianyu Chen et al., 2019; Qi et al., 2019; P. Wang et al., 2018). We also use a replay buffer not only to increase the robustness of the model in all the situations but also to avoid overfitting issues associated with certain problem scenarios. This is a much-needed step where it is sought to generate random experiences for training.

1.3.2. Overview of the model

With regard to the input space of the model, we consider explicitly at each time step t, 3 blocks of state. This includes the information from downstream sources (out of the sensing range but within the connectivity range) X_d; information from proximal or "local" sources (that is, information sources that are within the range of the CAV's sensors), X_l; and the CAV's information, X_{CAV}. In sum, the overall state space can be represented as a triplet (X_d, X_l, X_{CAV}). Information from the farther (downstream) sources are characterized as being of "variable length", that is, it changes when there is a change in the number of vehicles in the CAV's within its connectivity range. To address this variable length input problem, we adopt in this work, a similar Deep Set concept to aggregate the dynamic sized input into a fixed shape but with a superior normalization mechanism. The second information source captures the driving environment within the close neighborhood of CAV, which is incorporated to promote collision-free decisions by the CAV. We use 10-meter as the sensing range for CAV. The inherent large amount of detail is needed to fully describe the movement attributes of vehicles located in the same lane as the CAV, and those located on the lanes left and right of the CAV. In this work, we divide further, the local inputs into "left" lane, "right" lane and the "current" lane (the current lane is that which is occupied by the CAV). Information from the third source (that is, from the CAV itself), which includes its absolute location, speed and lane position, is provided as the final block of inputs to the CAV control system.

In Deep Sets, variable lengths of inputs are first fed into an encoding network to gain proper feature embeddings separately for each input. In this work, we adopted this concept to use fully connected neural networks φ to encode each downstream vehicle input $x_d \in X_d$ within the connectivity range, the input from each sensed lane $x_l \in X_l$ within the vicinity of the CAV, and the CAV's information $x_{CAV} \in X_{CAV}$ into a higher dimension feature space. Then we perform information fusion for the dynamic changing length among the feature space. Here, we simply use the same encoding network for both downstream and local inputs because they have the same meaning and representations. After the encoding network, the downstream embeddings are weighted and summed to obtain a fixed size input for subsequent operation. The total feature embedding obtained from downstream information is:

$$F_d = \sum_{i=1}^{n} w_i \varphi(x_d^i)$$

Where: x_d^i and w_i the raw feature input and weight for i^{th} vehicle that is located downstream of the CAV. The weight values represent the relative importance of information from the various sources, for the CAV driving purposes, and the sum of weights of information from all vehicles in the connectivity range is 1.

The local information sources are: "left", "right" and the "current" lanes. A matrix can be used to represent the feature embedding which contains information associated with these 3 lanes, as follows:

$$F_l = \begin{pmatrix} \varphi({x_l}^{left}) \\ \varphi({x_l}^{current}) \\ \varphi({x_l}^{right}) \end{pmatrix}$$

The embeddings of CAV's information has a similar expression as follows:

$$F_{CAV} = \varphi(x_{CAV})$$

The model concatenates the feature embeddings for downstream, local and CAV information to yield a fixed-sized feature map. Then the feature map is flattened and fed into the Q network ρ for Q values. Denoting the overall model that contains the encoding network and Q network as: $\hat{Q}$, with parameters θ, the final Q values can be expressed as:

$$\hat{Q}_\theta(s_t, a_t) = \rho([F_d^t; F_l^t; F_{CAV}^t], a_t)$$

The encoding network and Q network are trained on mini-batches sampled from a replay buffer R, which contains the transitions of (s_t, a_t, r_t, s_{t+1}). For each mini-batch, the objective of the training is to minimize the following loss function:

$$L_\theta = \frac{1}{b} \sum_t y_t - \hat{Q}_\theta(s_t, a_t)$$

Where: b is the batch size and $y_t = r_t + \gamma \max_a \hat{Q}_\theta(s_{t+1}, a)$. In Figure 1.3, we present the layout of the model. Multi-layer perceptron (MLP) with a rectified linear activation function (ReLU) is used for each component with the following architecture:

- Encoding network φ : $Dense(64) + Dense(32)$
- Q network ρ: $3 \times Dense(64) + Dense(32) + Dense(16) + Dense(8)$
- Output layer: $Dense(3)$

It is sought to facilitate full exploration (by the agent) of the environment and to acquire adequate experiences in both categories of driving success and failure (collision). Therefore, we use a Deep-Set Q learning that incorporates an experience reply buffer and a "warming up" phase with total T steps that allows the agent to undertake random actions. From step T+1, we perform training by maximizing the reward and minimizing the losses, as mentioned above. To further reduce the variance for the model, we apply a double Q learning mechanism with a soft

updating for target network as introduced in (Van Hasselt et al., 2016). Algorithm 1 (below) presents the steps for the entire process.

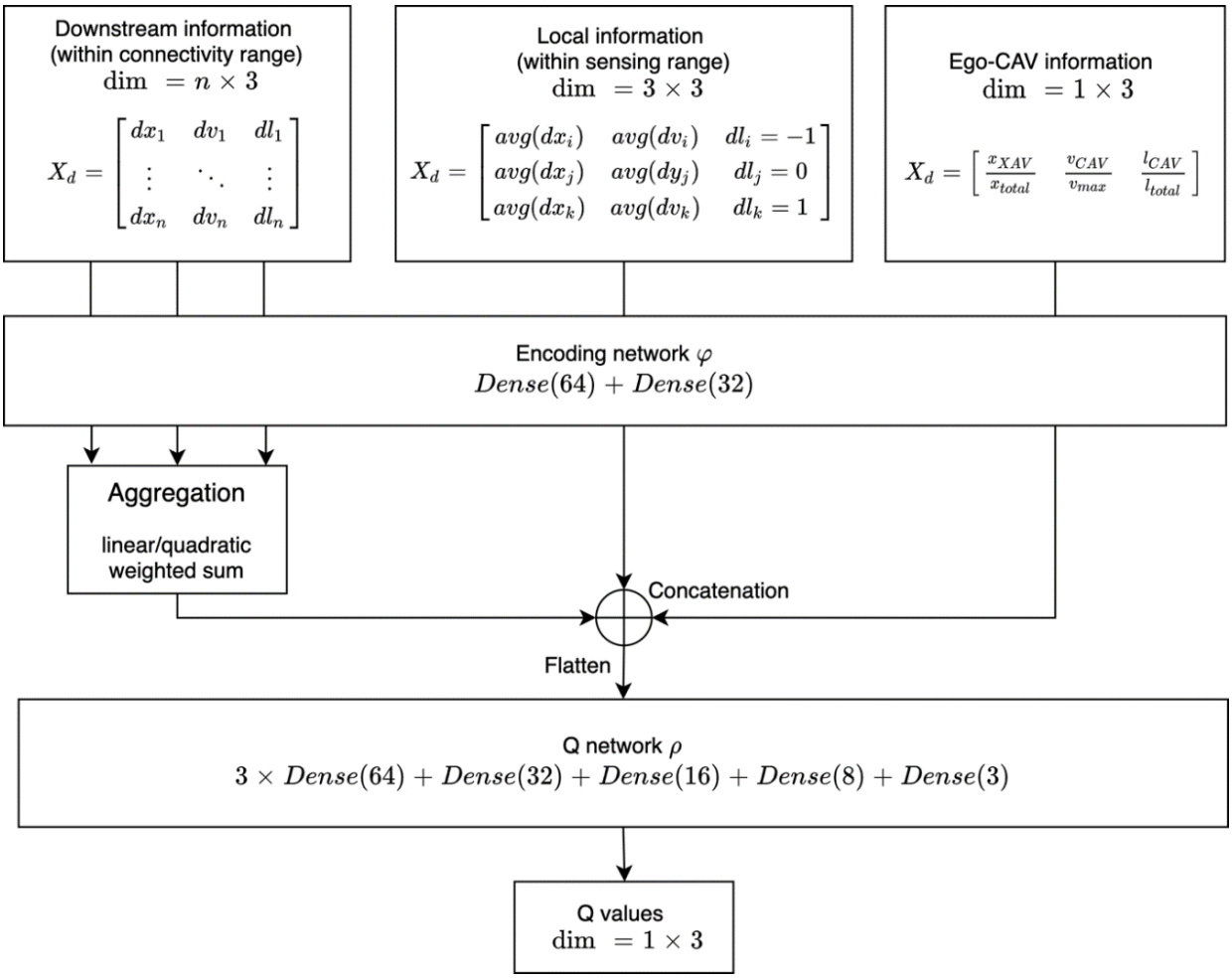


Figure 1.3 Proposed network architecture

Table 1.1 Deep set Q algorithm

Algorithm 1	Spatially Weighted Deep-Set Q Learning with Experience Replay and Target Network

Initialize the reply memory R to capacity N

Initialize the weights for both Encoding network φ and Q network ρ which jointly denoted as Network $\hat{Q}_\theta$ and Target Network $\hat{Q}_t = \hat{Q}_\theta$

Warming up steps

For time step $t = 1$ to T_1 (warming up steps) **do**

Take a random action $a_t = a_r$ and gather the transition (s_t, a_t, r_t, s_{t+1})

Store the transition (s_t, a_t, r_t, s_{t+1}) into the memory buffer R

Main training loop

For time step $t = T_1 + 1$ to T (training steps) **do**

Generate new samples and update memory R

With probability ϵ select a random policy $a_t = a_r$

Otherwise:

Encode the information from the CAV directly, downstream sources and sources in the immediate locality, with φ and weights w_i

$$F_d = \sum_{i=1}^{n} w_i \varphi(x_d^i), \; F_l = \begin{pmatrix} \varphi(x_l{}^{left}) \\ \varphi(x_l{}^{current}) \\ \varphi(x_l{}^{right}) \end{pmatrix}, F_{CAV} = \varphi(x_{CAV})$$

Obtain action $a_t^* = \underset{a_t}{\operatorname{argmax}} \hat{Q}_\theta(s_t, a) = \underset{a}{\operatorname{argmax}} \rho([F_d^t; F_l^t; F_{CAV}^t], a)$

Execute a_t^* and observe reward r_t and next state s_{t+1}

Store transition $(s_t, a_t^*, r_t, s_{t+1})$ into the memory buffer R

Set $s_t = s_{t+1}$

Training the model at each training step

Sample random mini-batch with size b from R

For each training examples with the batch, set the target of Q value

$$y_t = \begin{cases} r_t + \gamma \max_{a_{t+1}} \hat{Q}_\theta(s_{t+1}, a_{t+1}) & if\ s_{t+1}\ is\ not\ done \\ r_t & if\ s_{t+1}\ done \end{cases}$$

Perform a gradient step optimizing loss function in $L_\theta = \frac{1}{b}\sum_t y_t - \hat{Q}_t(s_t, a_t)$

Updating the Target Network

If mod(t, target updating frequency) == 0

Set $\hat{Q}_t = \hat{Q}_\theta$

1.4 Results

1.4.1. Training process

In the training process (Figure 1.4), the first 5×10^5 steps (417 episodes) are "warming up" phase that indicates the reward for making random choices. This phase is intended to equip the agent with a sufficient learning experience that contains both successes and failures. The training commences after 5×10^5 steps and converges in 10^6 steps (833 episodes). Specifically, the "jump" at approximately 420 episodes is a gradual increase which goes up along with the training process. In our case, the model converges fast compare to the "warming up" phase and the convergence phase. After the training, the CAV can perform lane changing maneuvers without collision.

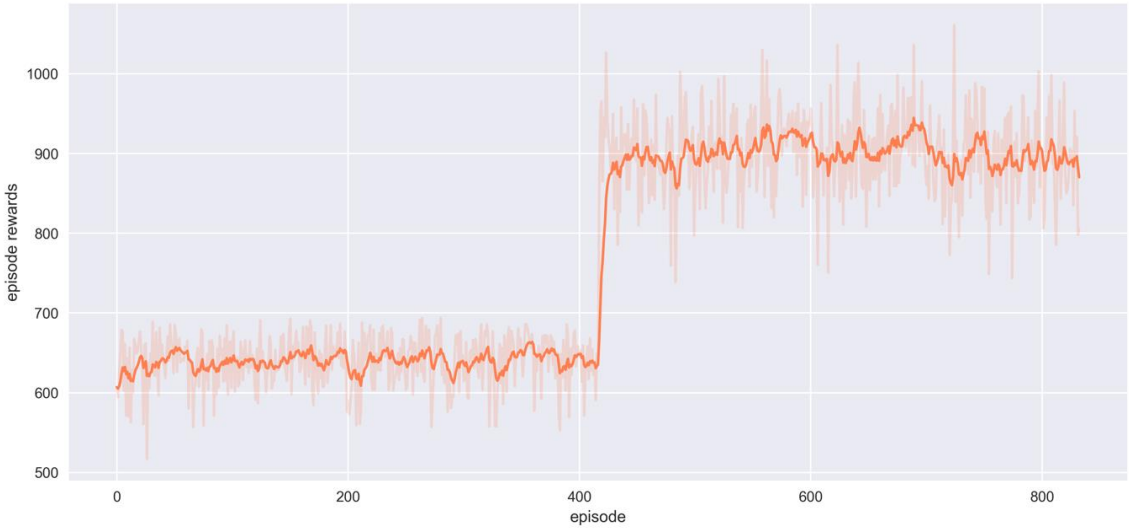
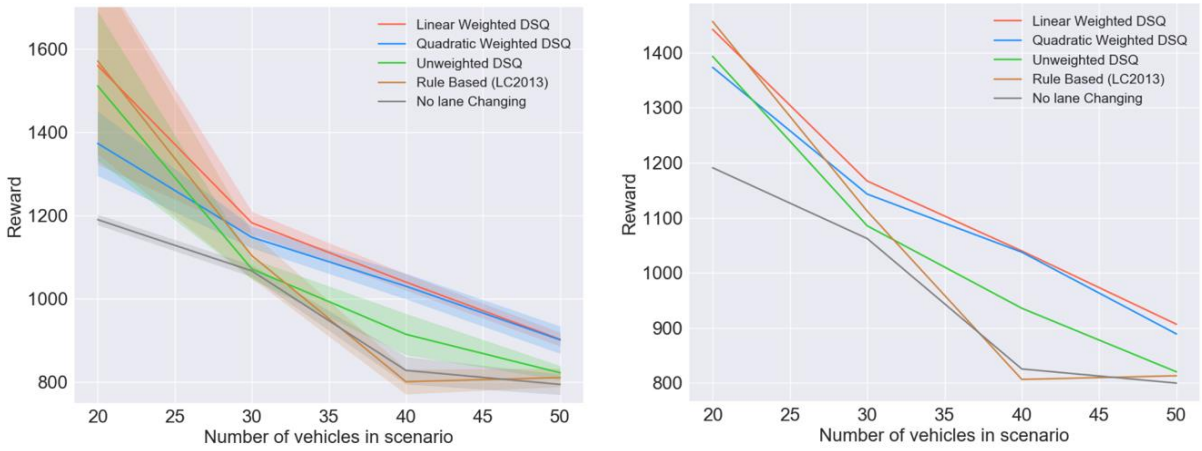


Figure 1.4. Rewards gained vs. the number of training steps

1.4.2. Comparative analysis

We compared the results from our proposed model with the four baseline operation decision models: the unweighted Deep Set Q learning model, the quadratic weighted Deep Set Q learning model, the rule-based lane-change model, and the no-lane-change model. To demonstrate the consistency as well as robustness of the proposed model, we trained our model in 1 specific scenario (1 CAV and 50 HDVs) and test the model in various density scenarios, which is achieved by changing the number of HDVs. The mean and median performance are compared in Figure 1.5 and Table 1.2.



(a) Mean with 95% confidence interval (b) Median

Figure 1.5. Relative performance of the models investigated, using 10 test episodes

Table 1.1 Performance comparison for different models in different scenarios.

Models Scenarios		**No Lane Changing**	**Rule Based (LC 2013)**	**Unweighted DSQ**	**Quadratic Weighted DSQ**	**Linear Weighted DSQ**
20 vehicles	Mean	1189.88	1570.49	1510.71	1372.78	1559.57
	Median	1191.04	1456.63	1393.34	1373.38	1442.28
	S.D.	19.47	359.22	287.22	123.72	382.47
30 vehicles	Mean	1066.95	1103.65	1071.63	1147.56	1182.19
	Median	1062.43	1112.17	1085.52	1143.11	1166.82
	S.D.	21.81	87.45	38.47	40.81	42.26
40 vehicles	Mean	828.18	801.07	914.74	1030.15	1039.95
	Median	825.5	806.57	935.76	1037.64	1039.87
	S.D.	51.99	47.13	79.23	48.78	30.57
50 vehicles	Mean	794.64	810.93	822.45	901.41	902.7
	Median	799.86	813.15	820.27	889.09	906.68
	S.D.	38.73	34.38	26.34	51.94	25.44

The results were tested in various scenarios that differ in terms of their traffic densities. From Figure 1.5, it can be seen that in most scenarios, the linear weighted CAV decision model outperforms the unweighted and the quadratic weighted models, and all these three are superior to the no-lane-change and rule-based baseline decision models. In one scenario with low traffic density (that is, only 21 HDVs on the road track) and where traffic conditions approach free flow, it was found that all the HDVs and the CAV operated at speeds that approach their maximum possible speed under such stable traffic conditions.

Also in this traffic density scenario, for each of the five models, the CAV algorithm was found to make a consistent decision, that is, the CAV keeps in its lane. This is intuitive because under such traffic conditions, there is no incentive for the CAV to change lanes. Further, in this scenario, it is observed that the slower HDVs stay in the rightmost lane and leave the left lane to the other vehicles that have higher speeds. When the rule-based decision model is used in this scenario, we observe that the CAV "captures" the leftmost lane where it maintains a high speed. In that scenario, the rule-based model provides the CAV the highest reward compared to other models.

In another traffic density scenario that involved all 51 vehicles, it is observed that at such high traffic density, the vehicles cannot gain much travel benefit even after making lane changes. In that scenario therefore, all the models were found to yield similar reward level. In a third scenario with traffic density that is in-between the first two scenarios described earlier (that is, 30-40 HDVs on the road track), we observe that the CAV can greatly enhance its operational efficiency by making appropriate lane-changing decisions as and when needed. We find that in this scenario, our proposed 2 "weighted DSQ" models outperform the other 3 baselines while the linear weighted model is slightly superior to the quadratic weighted model. This result may be attributed to the model's capability to obtain and appropriately process (through weighting), the information on traffic conditions further downstream (due to its connectivity capabilities) and traffic conditions in its immediate local environment (due to its sensing capabilities). This capability helps it to identify an optimal driving policy under the given traffic conditions, and to make proactive decisions to avoid travel delay caused to it due to proximal or anticipated imminent delay threats in the traffic environment.

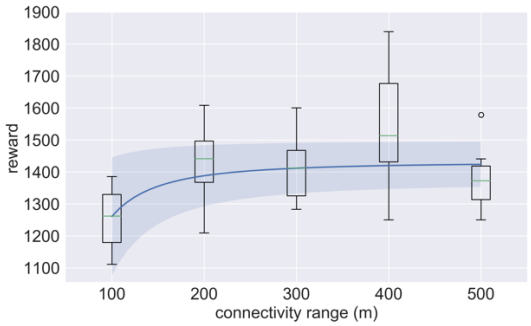
1.4.3. Critical connectivity range

In the context of this research, connectivity range refers to the maximum distance at which connectivity is available. Therefore, the "optimal" connectivity range refers to distance after which any additional benefits of increasing connectivity are negligible. As we stated earlier in this work, the developed model is capable not only of normalizing explicitly, the input scale but also of accounting for the spatial distribution of the inputs. Therefore, it is possible to use the same model under various specified connectivity ranges without the need to retrain the model.

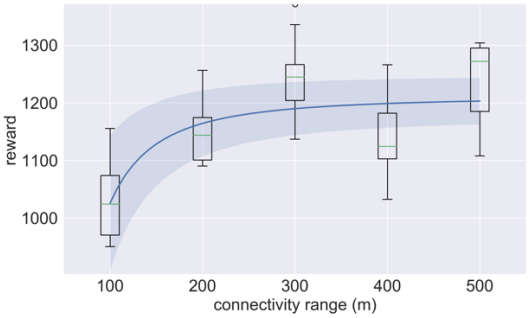
The results of the experiment (Figure 1.6) demonstrate that for a given traffic density, as the connectivity range increases, the model performance increases sharply up to a certain point after which it increases at a reduced rate and almost flattens out. This is because, when connectivity range in low, a unit increase in the level of this attribute causes a proportionately

higher amount of downstream information to be sent to and received by the CAV. To the CAV's decision processor that seeks to make proactive decisions, the incremental benefit of such information is significant. However, when the connectivity range is large, a unit increase in the connectivity range will produce relatively smaller benefits. This is due to the increased variance arising from noise or unrelated information that is received by the CAV, a situation that is exacerbated by the unpredictable and often errant nature of human drivers in HDVs located further away from the CAV. This trend suggests the existence of a critical connectivity range, in other words, a threshold beyond which the marginal benefits of increased range, begin to diminish. In this work, we determine this threshold from the derivative of the trendline, which in general, is an indicator of this marginal benefit. In each scenario (21, 31, 41, 51 vehicles), we evaluate the derivative of the trendline at 100-meter connectivity range $x_0 = 100m$ as the baseline marginal effect g_0. We keep increasing connectivity range x until the derivative of the trendline drops to $0.1g_0$, and then we observe that the marginal benefit drops to only 10% of baseline value, and the corresponding x is the critical connectivity range $x_{critical}$. Based on our experimental settings, for all 4 scenarios, the critical connectivity range $x_{critical}$ is approximately 270m (Figure 1.7 a).

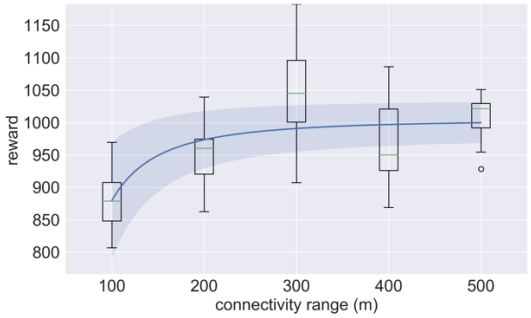
Further, in the scenarios with sparse traffic (21 vehicles in the corridor, that is, 20+1) and very dense traffic (51 vehicles), the convergence of reward is faster due to adding more information can barely improve driving in these scenarios. This can provide CAV manufacturers, the justification for specific critical connectivity range specifications and for them to provide CAV users with flexibility to select appropriate optimal range under a given set of traffic conditions. In other words, to achieve high efficiency in information transmission and usage for its efficient operations, the CAV should be able to automatically identify and adopt a specific connectivity range setting or mode based on the prevailing traffic density it has sensed.



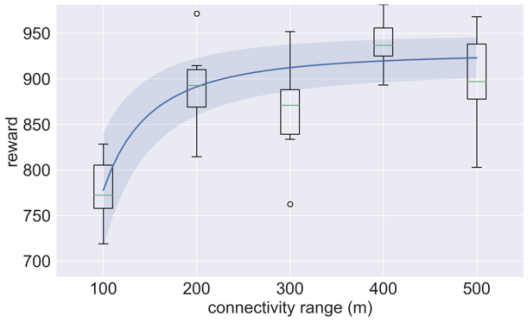
(a) Traffic density = 20 HDVs



(b) Traffic density = 30 HDVs

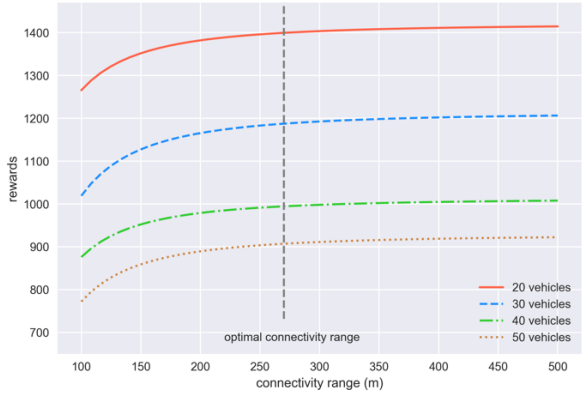


(c) Traffic density = 40 HDVs

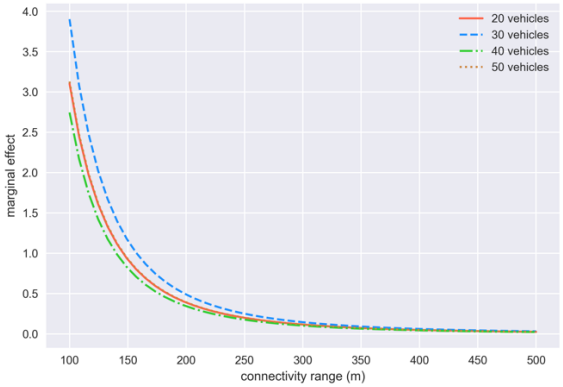


(d) Traffic density = 50 HDVs

Figure 1.6 Reward vs. connectivity range, using the normalization manipulation (linear weighted DSQ) model



(a) Reward vs. connectivity range: The effect of traffic density



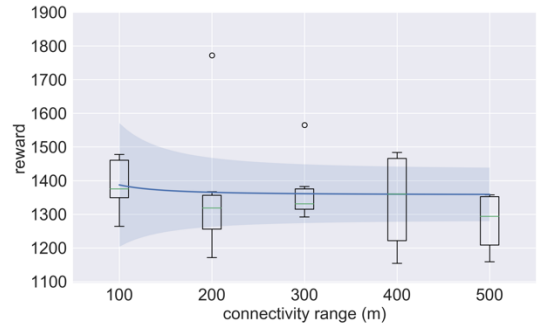
(b) The marginal effect of increasing connectivity range

Figure 1.7. Connectivity range relationship with reward and marginal effect

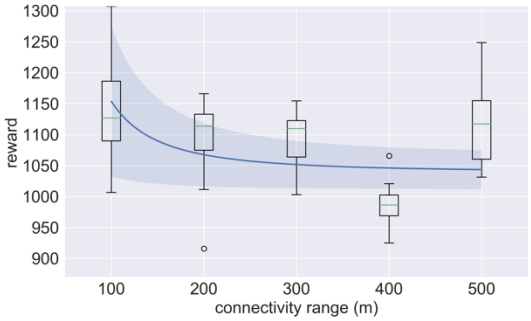
On the basis of the reward that is based on connectivity range, the marginal effect of increasing connectivity range on the reward was plotted (Figure 1.7 b). The figure shows that further increases of the connectivity range is not always beneficial because it exhibits fast diminishing returns in terms of the reward. This is seen for all four scenarios of traffic density, and the convergence of the curves representing the various traffic densities, seems to occur at approximately 270m. The elbow points of the curves seem to be in the range 170-180 m. The results can serve as a guideline for manufacturers of connected vehicle technology regarding not only the default setting of the connectivity range but also the manufacturer's recommended (and subsequently, CAV driver-adjusted) setting of the appropriate connectivity range setting for the prevailing traffic conditions (density). In some cases, a higher connectivity range may come at a higher cost to the driver. In such cases, both the marginal benefits and marginal costs of increased connectivity range will need to be considered in order to establish the most cost-effective level of connectivity, under a prevailing set of traffic conditions.

1.4.4. Analysis involving classic DSQ model

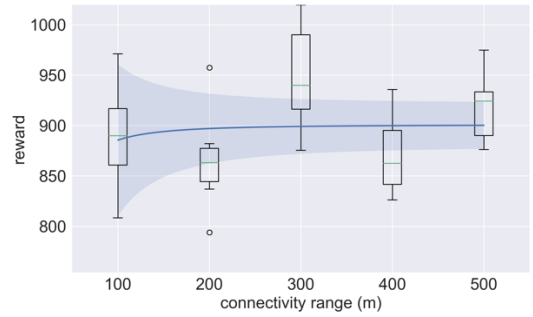
As we discussed earlier in the "Research Gap" section of this work, the unweighted Deep Set Q (DSQ) model proposed in (Huegle et al., 2019) may suffer from the problem of non-transferability across different traffic density conditions, unlike the normalization manipulation (weighted sum operation) model developed in this work. To investigate this hypothesis, we perform the connectivity range experiment using the baseline unweighted DSQ model. The results are presented in Figure 1.8. As shown in the figure, for all the different traffic density scenarios, an increase the connectivity range does not lead to an improvement in CAV's performance, unlike Figure 1.6 (linear weighted DSQ). This is because for the unweighted DSQ, there is no proper normalization mechanism. The embedding scale of downstream information grows linearly with the number of connected vehicles in a fixed space. That is, the scenario with 80 vehicles has larger scale of feature input than that with 40 vehicles. Therefore, increments in the connectivity range will create an unbalance in scale between downstream information, local information and CAV information. When the connectivity range is very large, the unweighted DSQ model causes the downstream information to overwhelm the local information. However, local information is vital for some close-space maneuvers including lane changing. Therefore, in the unweighted DSQ model, such "wiping out" of the local information will lead to a drastic increase in crashes. For this reason, if the DSQ model is used, increases the connectivity range will generally not be seen to improve the CAV's performance, which is counter-intuitive. Therefore, the normalization manipulation (weighted sum operation) model in our proposed framework is more effective in accounting for the benefits of increased connectivity (without sacrificing the local information) and therefore is more appropriate for robust CAV operations.



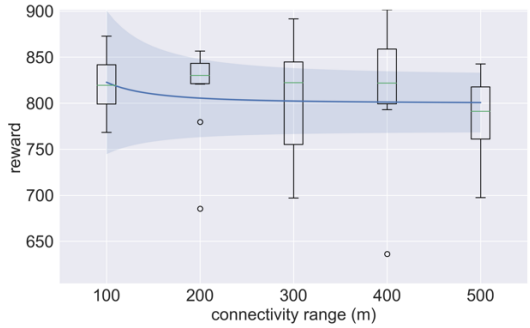
(a) Traffic density = 20 HDVs



(b) Traffic density = 30 HDVs



(c) Traffic density = 40 HDVs



(d) Traffic density = 50 HDVs

Figure 1.8. Reward vs. connectivity range, for different scenarios of traffic density, using the baseline (unweighted DSQ) algorithm

1.5. Concluding remarks

In this research, we present an end-to-end deep reinforcement learning based processor to make high-level decisions in controlling a CAV's lane change operations in complex mixed traffic. In this context of operations, the developed model was observed to achieve its target of helping a CAV increase its travel effectiveness and efficiency in terms of safety and mobility, respectively. As part of efforts to achieve this overarching objective, the research also demonstrates the efficacy of the proposed model in four areas. First, the model adequately fuses the long-range and short-range information based on the spatial importance of information which, in turn, in a function of the spatial distance between the information source and the CAV. Second, the model helps the CAV make safe lane-change decisions even after relaxing the collision-free restriction imposed by the low-level controller in the simulator. Third, the model handles adequately, the highly dynamic length of inputs (that is fed to the CAV). Finally, the model efficacy is demonstrated by applying it to traffic scenarios with different densities without the need to

retrain the model. For a comparative evaluation, we compare the proposed model with four classic baseline methods (unweighted classic DeepSet Q learning method, quadratic weighted DeepSet Q learning method and the multiple-rule-based). The results suggest that the model proposed in this research outperforms the baseline and other models.

With regard to the issue of connectivity range and issues of practical implementation, the research demonstrates how the critical connectivity ranges at a prevailing level of traffic density, could be ascertained. The critical connectivity range we show in our results is based on the specific operational context of CAV lane-changing. It is important to note that other operational, tactical or strategic contexts of CAV operations (acceleration, route choice, and so on) may require different communication ranges. Therefore, the overall practical connectivity range for CAV movements in general, could be established as a function (for example, the maximum) of the connectivity ranges of the individual operations contexts. The model presented in this research can be applied to the other contexts to determine the connectivity ranges for those contexts. Therefore, the research provides CAV manufacturers a justification for specific critical connectivity range specifications. In addition, with the developed model, the CAV can automatically identify and adopt a specific connectivity range setting or mode based on the prevailing traffic density. Therefore, the model also presents to manufacturers, a capability to provide CAV users with flexibility to select appropriate optimal range under a given set of traffic conditions. In general, CAV manufacturers may find this useful in their efforts to develop appropriate vehicle connectivity protocols and architectures.

Moving forward to future work, with the help of connectivity and storage system, research may find it worthwhile to consider temporal information including historical data on the vehicle position, speed, and acceleration accounting for the possibility of longer times (delays) of the CAV's decision process. The incorporation of such historical data in the analysis may help address hypotheses regarding the effect of imminent traffic conditions downstream that often require rerouting or preemptive evasive maneuvers of the CAV. Examples of these downstream conditions include construction sites or workzones, accidents, debris, potholes, and obstacles on the roadway. Therefore, future research could examine the efficacy of trajectory planning in CAV by incorporating both instant (short-term) and long-term information. Also, future research could investigate the efficacy of DRL based method, for purposes of CAV control, in making collaborative decisions that maximize the utility of all agents in the entire corridor rather than the CAV's utility. An example of such research directions is the use of the proposed methodology to promote traffic string stability and cooperative crash avoidance maneuvers in emergency situations. In addition, it is suggested to carry out field experiments using the proposed method, to validate its efficacy and to highlight the capabilities of deployment. Further, in determining the critical connectivity range, future studies may consider not only the marginal benefits as done in this research, but a combination of both marginal benefits and marginal costs of connectivity range increments. The cost aspects could include the initial purchase/installation cost and operations cost of connectivity devices, and the cost of computing power to process the information obtained through connectivity. Finally, while we extol the virtues of DL in helping to fuse space-weighted information for CAV controls, particularly for lane-changing as

demonstrated in this research, we duly recognize the potential limitations of reinforcement learning in general. These include the problem of domain adaptation, the deviation between the real environment and simulation outcomes, uncertainty of guaranteed safety performance, lack of sufficient realistic data for model training and setting calibration, and the relatively low transparency. Future work on enhancing the study framework or applying it in new contexts could incorporate a more advanced algorithm that obviates these limitations.

CHAPTER 2: MULTIAGENT COOPERATIVE CONTROL OF CONNECTED AUTONOMOUS VEHICLES AT HIGHWAY RAMPS

2.1 Introduction

Vehicle automation and connectivity are promising technologies that are expected to completely transform the transportation system (AASHTO, 2018; FHWA, 2018, 2015). The anticipated benefits include improvements in mobility, travel efficiency, productivity, and safety across various categories of transportation facility users and other stakeholders (FHWA, 2019; Li et al., 2020; (FHWA, 2019; Li, Y., S. Chen, P. Ha, J. Dong, A. Steinfeld, 2020; Sinha and Labi, 2007; World Bank, 2005). Connectivity is considered an inseparable sibling of automation (Ha et al., 2020a) and is often discussed within the context of "Internet of Things" (IoT) which enables information sharing between agents in a system. This is consistent with the concept of cooperative awareness within road traffic (that is, road users and roadside infrastructure are informed about each other's position, dynamics and attributes (ETSI, 2019). It is expected that the connectivity technology that is inherently associated with automated or autonomous driving, will facilitate efficient operations of CAVs that operate independently or in a network. In this research, we define a CAV network as a collection of CAVs and HDVs that are detectable (that is, within the CAVs sensing ranges) operating within a specific spatial scope that may be a road network, corridor, or segment. In the CAV network, the nodes are the CAVs and detectable HDVs, and the links are the communication channels between them. The spatial scope constitutes an environment that has a connectivity range where traffic information is shared and instructions are issued for controlling the CAVs movements. This can happen in at least two ways: cooperative sensing and cooperative maneuvering (Hobert et al., 2016). Cooperative sensing leads to an increase in the sensing range and promotes a greater awareness of the driving environment, and cooperative maneuvering promotes collaborative operations of the CAVs on the roadway as their individual movements are planned by a centralized or decentralized decision processor. The CAV control methodology proposed in this work, which uses a graph representation to model the information flow, generating decisions with centralized Q learning, is general in nature and can address a variety of driving decision contexts. We demonstrate the control methodology via a use case involving cooperative lane-changing due to merging maneuvers at the approaches to freeway exit ramps. The methodology can be applied easily to other control decision contexts after changing the action space, reward function and retraining.

2.2 Methodology

2.2.1 CAV networks as a graph

In order to make proactive and safe decisions during the driving task, a CAV needs not only information pertaining to vehicles in its proximity (local information) but also information pertaining to vehicles at downstream or upstream locations (global information). Local information is generally acquired through the use of onboard sensors while global information is obtained through cooperative sensing due to connectivity capabilities of the vehicles.

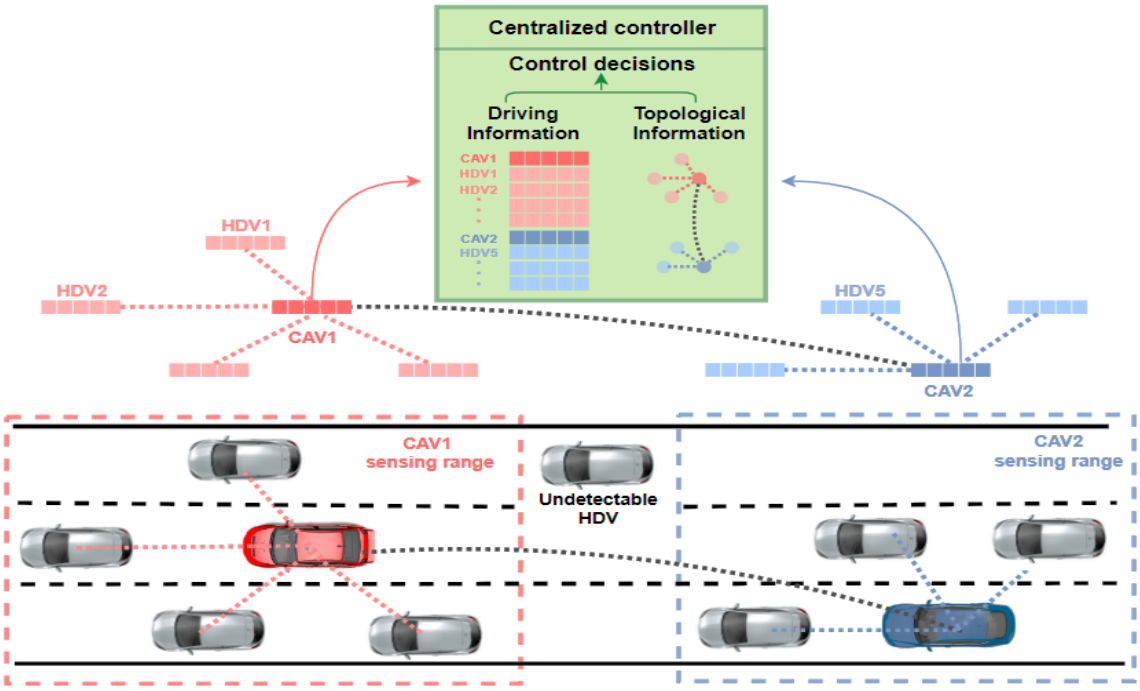


Figure 2.1 Graphic representation of a CAV network

Local information is useful to the CAV's short-term decisions such as immediately making a lane change, while global information enables the CAV to make relatively far-sighted decisions including lane change at a point further downstream. It seems intuitive that consideration of both local and global information is critical to the CAV's task of reconstructing its driving environment and generating safe and effective driving decisions. Also, regarding information dissemination, any information pertaining to human driven vehicles (HDVs) is fed to the CAVs via sensors in the local environment, while CAVs within a certain wider locus (referred to as the "connectivity range") are capable of sharing information with each other. This decision dependency and information flow path can be modeled using a graph. A graph is a data structure with great expressive power to model a set of objects (nodes) and their relationships (edges). Graphical representation has broad applications in a variety of disciplines including social sciences (Barnes, 1969), chemistry (Balaban, 1985), and transportation (Derrible et al., 2011).

As shown in Figure 2.1, each node in the graph represents a vehicle and the edges represent the connection between the vehicles. For example, the CAVs can obtain information pertaining to not only the HDVs in their immediate neighborhood or "sensing range" (via sensors) but also information from other CAVs (via connectivity). Therefore, the edges in the graph represent the information dissemination paths.

Graph Neural Network (GNN) has gained increasing popularity in various domains, including social network analysis (Qiu et al., 2018), knowledge graph (Kipf et al., 2019), recommender system recommendation (Fan et al., 2019), and the life sciences in general (Fout et al., 2017). GNN can extract relational data representations and generate useful node embeddings not only on the node features but also on the features from neighboring nodes. As the generalization of Convolutional Neural Network (CNN), Graphic Convolutional Network (GCN) presented by Kipf & Welling (2019) has great potential to aggregate the information for a clique of nodes when generating the node embeddings. It essential to point out that both local information and global information are essential for CAV to make efficient driving decisions. Therefore, an explicit fusion method is needed to combine the information from both classes of information sources. GCN is considered an ideal technique to do this, due to its inherent capability to aggregate information from different nodes.

A related motivation for applying GCN to fuse data for CAV control purposes, is that, to generate decisions for the CAV node, the information from its neighboring nodes including surrounding HDVs (that provide the local information) and other CAVs (that provide both local and global information) are incorporated contemporaneously. Additionally, the weights of the GCN layers serve as "attention" mechanisms that facilitate the CAV's learning by automatically granting greater focus on information that is deemed more relevant to the CAV's decisions compared to information that is deemed less important. For example, intuitively, when a CAV agent is making a lane-change decision, vehicles in its immediate proximity may be assigned greater weights compared to those farther away. A second example is the assignment of greater weight to vehicles downstream of the ego CAV, compared to those upstream. In GCN, this relative importance is automatically encoded as weights, since the weights are learned automatically, the "importance weights" are not specified by the user but rather intrinsically derived from successful and unsuccessful driving experiences. These weights could be a function of the surrounding vehicles' intentions, lane positions, distances, or other attribute, and generally can help yield high-reward decisions.

The inputs of the GCN block can be the feature matrix containing raw information (speed, location, intention, etc.) of each vehicle, and the adjacency matrix depicts the information flow topology as well as the decision dependency. The output of GCN is a node-level feature embedding map, which contains the information of both locality and global environment by fusing the raw data from two sources. These node embeddings can serve as key knowledge for making informed and collaborative driving decisions for all the CAVs in the CAV network.
A good example of combining GNN and DRL is the Graph Convolutional Reinforcement Learning (DGN) (Jiang et al., 2020). The model uses GNN as the encoder to learn abstract relational representations between agents, and then feeds the representations into a policy network for actions. By jointly training the encoder and policy network, the DGN agents are able to develop cooperative and sophisticated strategies. From Jiang et. al's ablation study, graph convolution greatly enhances the cooperation of agents (Jiang et al., 2020); such cooperation is needed in autonomous driving tasks.

Inspired by DGN, the present study modified the methodology so that it can produce dynamic length outputs, to facilitate adaptation to the nature of autonomous vehicle driving operations. Due to the dynamic number of agents, the use of separate decision processors (i.e., separate Q networks) for each agent will pose difficulties in the joint training process, and it cannot be guaranteed that the agents will jointly collaborate (Zhang et al., 2019). Also, the number of parameters for separate Q networks will grow exponentially with the number of agents, and therefore is not scalable from the perspective of joint training. Further, for a given driving task, all the CAVs should be considered homogeneous and treated equally by using the same control model to control their maneuvers. This prevents some of the models from overfitting specific scenarios at the stage of independent training of the model. Therefore, a more efficient way to achieve the desired outcomes is to use parameter sharing (Gupta et al., 2017), in other words, a shared centralized Q network to output actions for all agents. The overall working flow can be seen from Figure 2.1, the driving information and topological information for detectable vehicles are fed to the centralized controller which generates control commands of all CAVs.

In this setting, the CAVs can be treated as probe sensors to collect data on both local and global driving environments for the centralized controller. When the number of CAVs increase, the overall traffic condition of a road segment can be better understood, which will further enhance the efficacy of the CAV's decisions in terms of safety and systemwide mobility.

Model architecture

At each timestep t, the centralized agent interacts with the environment transition by observing state s_t, taking action a_t, landing in next state s_{t+1} and receiving reward r_t , which can be summarized in a transition quadruplet (s_t, a_t, r_t, s_{t+1}). With regard to the input space of the model, at time step t, there are N vehicles, including all CAVs and detectable HDVs in the proximity of CAVs. So, N is a dynamic number. The state s_t is considered as a tuple of three blocks of information: nodes feature X_t, adjacency matrix A_t, and a CAV mask M_t documenting the index of the CAV: $s_t = (X_t, A_t, M_t)$. With regard to the node feature for any node i (raw information for vehicle i in the network), the following four (4) categories are considered: speed v_i, location p_i, lane position l_i, and intention I_i. At each time step, the CAV in the neighborhood of vehicle i is able to gauge the raw information (except the intention of HDVs) of vehicle i via its onboard sensors, and construct a quadruplet $x_i = (v_i, p_i, l_i, I_i)$ to represent the vehicle i. Since CAVs can directly share their driving information, in practice, there is no need to sense other CAVs in the vicinity of the ego CAV. As the control model (DRL model in this research) is a central control unit implemented on the roadside unit (RSU), the first step is to aggregate the information acquired from CAVs. Therefore, central control unit concatenates all the raw information into overall node features $X_t = [x_i]_{i=1}^{N}$. To preserve the graph structure, an adjacency matrix is constructed during the information aggregation process to indicate the relationship between vehicles. Here, each CAV is connected with its nearby HDVs, and all the CAVs are connected. The information for both CAVs and HDVs are concatenated, and only

node embeddings for CAVs should be fed into the decision processor. Therefore, the indices for filtering out the HDV's node embeddings need to be saved.

At each time step t, node feature matrix X_t is first fed into a Fully Connected Network (FCN) encoder φ to generate node embeddings H_t in d dimensional embedding space $\mathcal{H} \subset \mathbb{R}^{N\times d}$, see Equation (1). Embedding here refers to a high-dimensional feature generated by neural network, and embedding space represents the set of all the embeddings (or, HD features).

$$H_t = \varphi(X_t) \in \mathcal{H} \qquad (1)$$

Then the graphic convolution is performed in the embedding space $\mathcal{H}$ for each vehicle. For each node, the GCN layer computes the node embeddings based on its own node embeddings from the encoder as well as the node embeddings for its neighboring node. In general, the GCN layer computes the nodes embeddings in parallel, for all the nodes in the network, as follows:

$$Z_t = g(H_t, A_t) = \sigma(\widehat{D}_t^{-1/2}\hat{A}_t\widehat{D}_t^{-1/2}H_tW + b) \qquad (2)$$

Where: $\hat{A}_t = A_t + I_N$ is the adjacency matrix with self-loops for each node; $\widehat{D}_t$ is the degree matrix computed from $\hat{A}$; and σ is the nonlinear activation function such as ReLU; W and b represent the weights and bias in the GCN layer. While there could exist multiple GCN layers, the total number of layers should be restricted to avoid "over-smoothing"(D. Chen et al., 2019). After the GCN block, the node embeddings map Z_t (including both CAVs and HDVs) is obtained. Then the node embeddings for CAVs are selected because only CAVs (unlike the detectable HDVs) are controlled. Filtering can be achieved using a simple dot product of mask M_t and Z_t:

$$Z_t^{CAV} = M_t \cdot Z_t \qquad (3)$$

The CAVs node embeddings are finally fed into a Q network ρ to obtain Q values, which indicate the "goodness" of a certain action. All the neural network blocks including FCN, GCN and Q network can be summarized as $\hat{Q}$ network parameterized by θ, where θ is the aggregation of all the weights and a_t represents the actions for all existing CAVs at time t.

$$\hat{Q}_\theta(s_t, a_t) = \rho(Z_t^{CAV}, a_t) \qquad (4)$$

To train the model, the classic Q Learning with Experience Replay and Target Network as proposed in (Van Hasselt et al., 2016; Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, 2016) is applied. In order to stabilize the training, the overall neural network is trained on mini-batches randomly sampled from a replay

buffer R containing transitions of (s_t, a_t, r_t, s_{t+1}). For each mini-batch, the objective of the training is to minimize the loss function (Equation 5):

$$L_\theta = \frac{1}{b}\sum_t y_t - \hat{Q}_\theta(s_t, a_t) \qquad (5)$$

Where: b is the batch size and $y_t = r_t + \gamma \max_a \hat{Q}_\theta(s_{t+1}, a)$ is the target of Q value.

Figure 2.2 presents the model layout. For each component of network, the following architecture is utilized:

- FCN Encoder φ: $Dense(32) + Dense(32)$
- GCN layer g: $GraphConv(32)$
- Q network ρ: $Dense(32) + Dense(32) + Dense(16)$
- Output layer: $Dense(3)$

Additionally, a "warming up" phase with *T* steps is established prior to the training in order to let the agent undertake random actions and fully explore the environment. This setting facilitates the agent's acquisition of adequate experiences in both successful lane changing and unsuccessful lane changing (collision), which further helps guarantee the safe lane-changing decisions. From step T+1, the training is performed by maximizing the reward and minimizing the losses as mentioned above. Algorithm 1 presents the detailed steps.

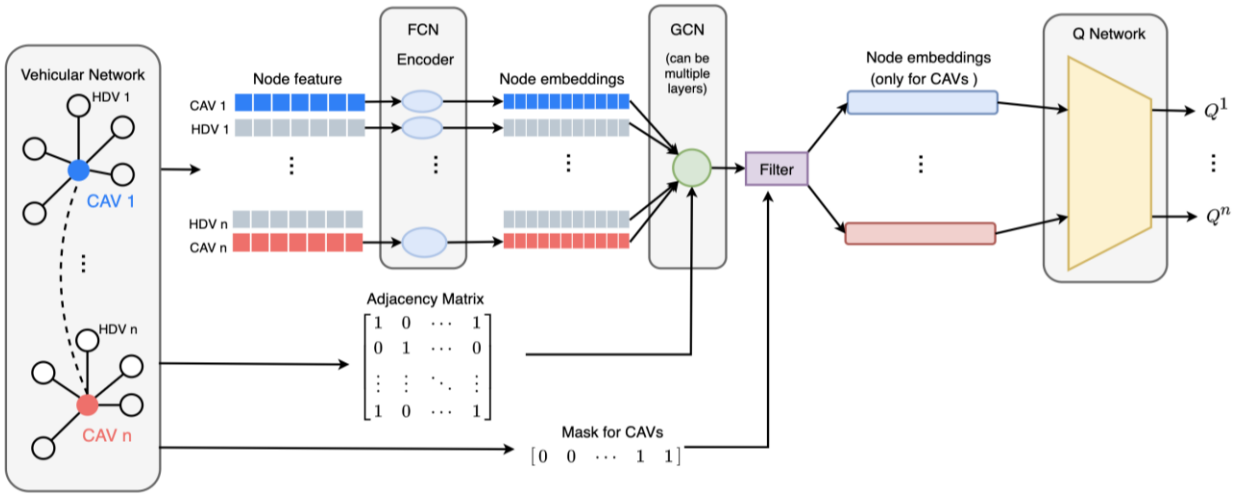


Figure 2.2. Model architecture

Table 2.1 GCQ algorithm

Algorithm 2.1 Graphic Q Learning with Experience Replay and Target Network
Initialize the reply memory R to capacity N
Initialize the weights for both Encoding block φ, graphic convolutional block g, Q network ρ which jointly denoted as Network $\hat{Q}_\theta$ and Target Network $\hat{Q}_t = \hat{Q}_\theta$
Warming up steps
For time step $t = 1$ to T_1 (warming up steps) do
Take random action combination for each agent i: $a_t = \left[a_r^i\right]_{i=1}^n$
Gather the transition (s_t, a_t, r_t, s_{t+1})
Store the transition (s_t, a_t, r_t, s_{t+1}) into the memory buffer R
Main training loop
For time step $t = T_1 + 1$ to T (training steps) do
Generate new samples and update memory R
With probability ϵ select a random policy $a_t = \left[a_r^i\right]_{i=1}^n$
Otherwise do:
$X_t, A_t, M_t = s_t$
Encode the raw node feature into a high dimensional feature map $H_t = \varphi(X_t)$
Perform graphic convolution $Z_t = g(H_t, A_t)$
Filter out the node feature for HDVs $Z_t^{CAV} = M_t \cdot Z_t$
Compute Q values for each action combination a_t $\hat{Q}_\theta(s_t, a_t) = \rho(Z_t^{CAV}, a_t)$
Select the $a_t^* = \underset{a_t}{\operatorname{argmax}}\, \hat{Q}_\theta(s_t, a_t)$
Execute a_t^* and observe reward r_t and next state s_{t+1}
Store transition $(s_t, a_t^*, r_t, s_{t+1})$ into the memory buffer R
Set $s_t = s_{t+1}$
Training the model at each training step
Sample random mini-batch with size b from R
For each training examples with the batch, set the target of Q value
$y_t = \begin{cases} r_t + \gamma \max_a \hat{Q}_\theta(s_{t+1}, a) & if\ s_{t+1}\ is\ not\ done \\ r_t & if\ s_{t+1}\ done \end{cases}$
Perform a gradient step optimizing loss function in $L_\theta = \frac{1}{b}\sum_t y_t - \hat{Q}_t(s_t, a_t)$
Updating the Target Network
If mod(t, target updating frequency) == 0
Set $\hat{Q}_t = \hat{Q}_\theta$

2.3 Results

Figure 2.3 presents the training curve on both loss and episode reward. In the training process, the first 2×10^5 steps (150 episodes) represent the "warming up" phase that CAVs are taking random actions for exploration. After being trained, both LSTM-Q and GCQ model were found to converge within 8×10^5 steps. The GCQ model was observed to exhibit superior performance in terms of convergence rate and reward gained for each episode after convergence. Also, it was observed that both models outperform the average performance of the rule-based model. It was concluded that, after being trained, the designed CAV control algorithm is capable of performing lane-change maneuvers without collision, and no congestion was observed on the road segment.

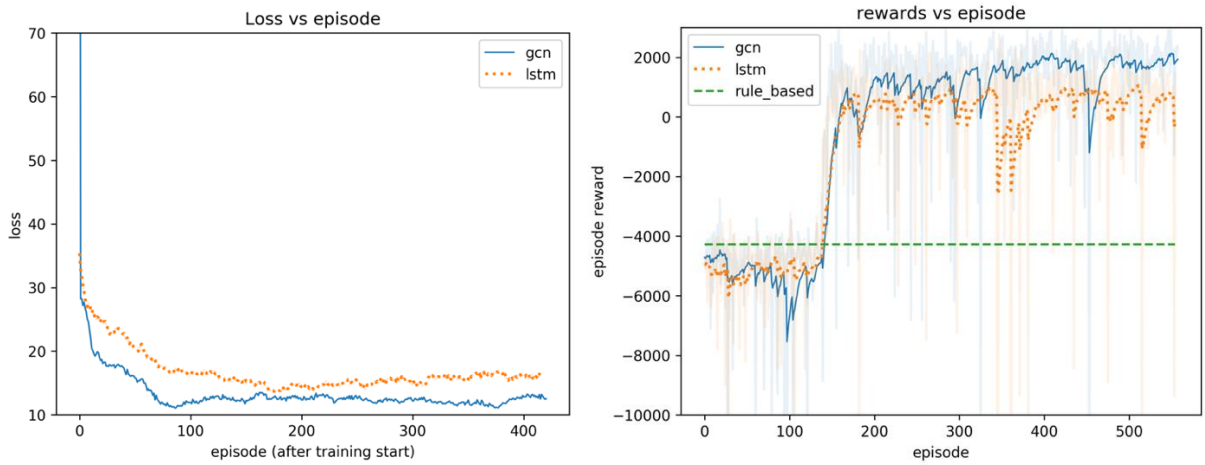


Figure 2.3 Loss and rewards vs. episode

Comparative analysis

To test its robustness, the model is evaluated in mixed traffic with different traffic densities. The training is performed using a density of 0.2 veh/sec inflow rate for HDVs, and 0.1 veh/sec inflow rate for CAVs on both merge_1 and merge_2. For the testing, the HDVs inflow was made to vary from 0.1 to 0.5 veh/sec. For model evaluation, we use the following metrics: the mean, median and standard deviation of the episode reward and total number of simulation steps per episode acquired by running three models separately for ten episodes in different traffic density scenarios. Generally speaking, the episodic reward is a combinative measure of successful merging out (intention reward), efficiency of each CAV(speed reward), safety (crash penalty) and driving comfort (lane change penalty) while the number of simulation steps per episode reflects the overall efficiency of the road segment. As mentioned in an earlier section of the research, the physical meaning of "episodic horizon" is the time cost associated with the sojourn of the twenty CAVs at the road segment (from the moment the first CAV enters the segment until the time the twentieth CAV exits the segment). Figure 2.4 presents the mean and standard

deviation of the model performance statistics for the proposed and baseline methods, across the different scenarios of traffic density in terms of episode reward. Figure 2.5 presents the results for number of simulation steps per episode. After trained with aforementioned parameters, the average successful merging-out rate for CAVs per episode across all the scenarios is consistently around 90% for GCQ model (18/20 CAVs) and 85% for LSTM-Q model (17/20 CAVs).

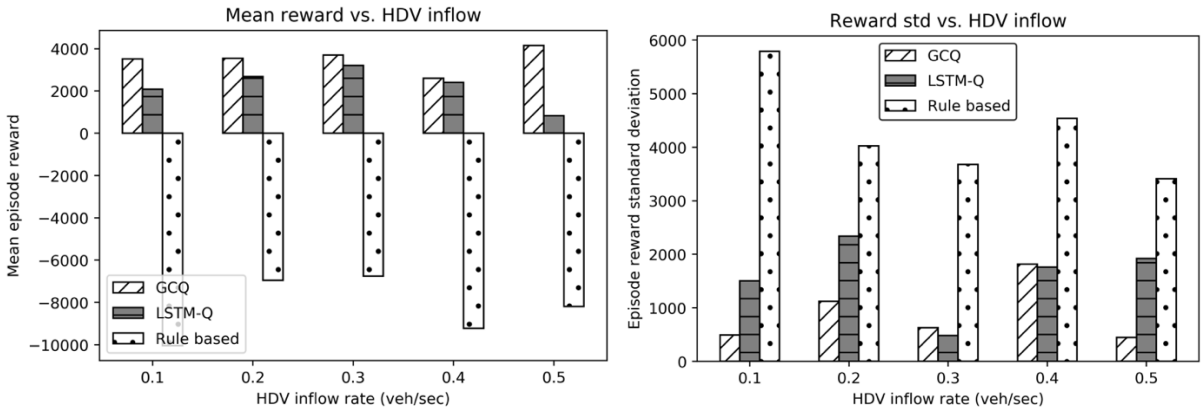


Figure 2.4 Mean and standard deviation of episode rewards across different traffic densities

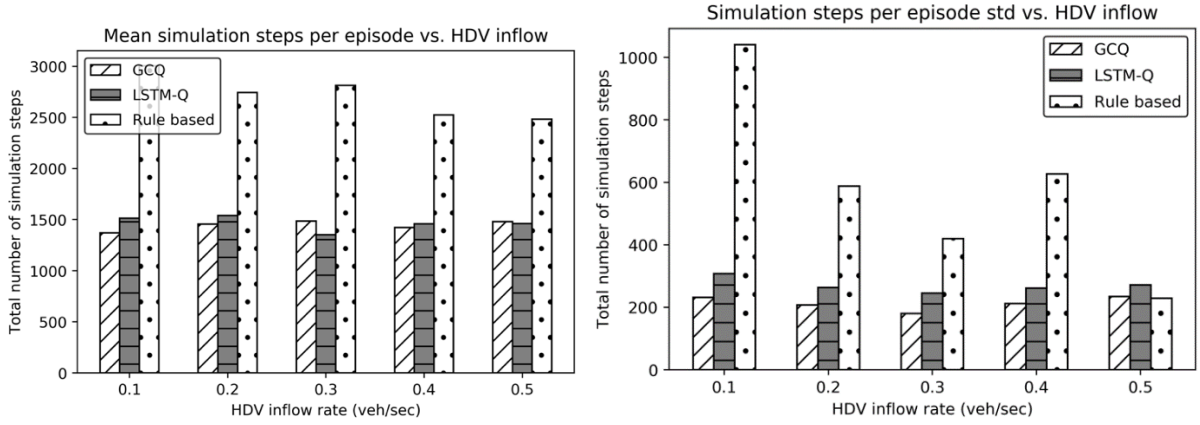


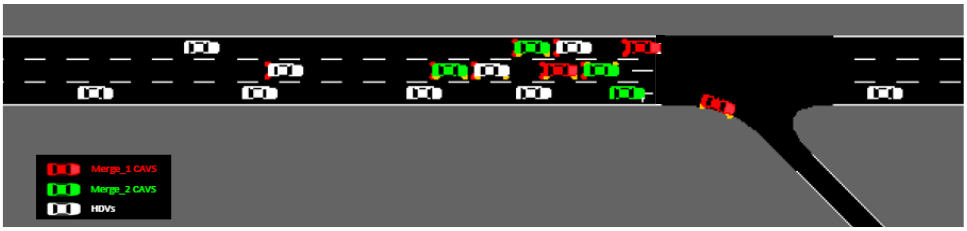
Figure 2.5 Mean and standard deviation of number of simulation steps per episode across different traffic scenarios

As shown in Figure 2.3, the proposed GCQ model outperforms LSTM-Q model in all the experimental traffic density scenarios, and both DRL-based models outperform the Rule-based model (LC-2013) by significant margins. Also, Figure 2.5, the number of steps required for each episode for GCQ model and LSTM-Q model are consistent over the scenarios and are much fewer compared to the rule-based model. This result reflects that when the total number of CAVs in the study area is fixed, both GCQ model and LSTM-Q model can guarantee the systematic efficiency of the road segment across the different traffic-density scenarios. We also examined the graphical simulation of the models and observed that within the HDV inflow rates studied, both the GCQ model and LSTM-Q model exhibited greater efficacy in guiding the CAVs to yield the lane for other vehicles, and in exiting from the intended ramp exit without any congestion or collision. These results may be attributed to the cooperative behavior of the vehicles when they are controlled using the developed model.

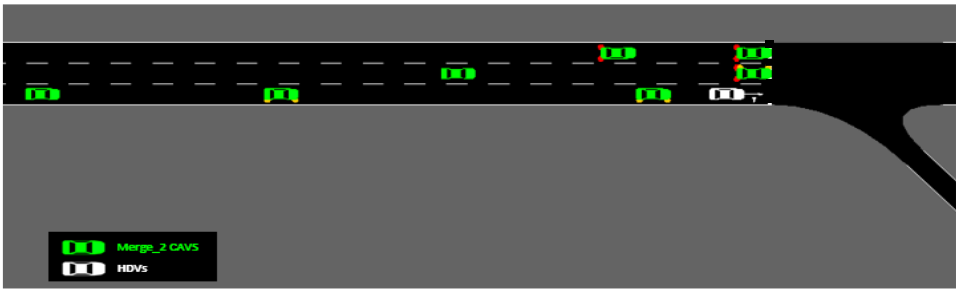
The results also suggest that the GCQ models lead to more consistent decisions with a smaller variance while LSTM-Q models sometimes fails to guide the CAV to make the right decisions. This is because for LSTM models, the order of input sequence matters when generating the context information, which is not always true in driving task because the decisions should depend on the spatial location of surrounding vehicles only instead of the sequence order of the inputs to the model. Therefore, when performing information fusion, a superior strategy, clearly, is to use a “permutation invariant” model such as graphic convolutional neural network.

The rule-based model can guarantee that all the vehicles exit the ramp successfully (100% success rate in merging at the approaches to the ramp), albeit with very low efficiency in majority of cases (as shown in Figure 2.6). Using visualization, an obvious limitation of the rule-based model can be demonstrated. In Figure 2.6 (a) we show that the merge_2 (green) CAVs capture the rightmost lane before the first ramp and block the way of the merge_1(red) CAVs. Under this situation, merge_1 CAVs must wait until the rightmost lane is clear to merge, and therefore this could lead to a traffic jam at the approaches to the ramp exit. This can be avoided if merge_1 CAVs are cooperative and actively yield the rightmost lane to merge_2 (red) CAVs before the first ramp.

Another certain flaw of the rule-based model is shown in Figure 2.6 (b), when the merging CAVs (merge_2, green) fail to reach the rightmost lane before the ramp and therefore need to wait at existing position until the rightmost lane is clear for ramp exit. The situation can be alleviated if CAVs take proactive actions to position themselves in the exiting lane well before they intend to exit. These inefficient outcomes of rule-base model will not only reduce the total efficiency of the system but also may cause severe traffic accidents in the real world.



(a) Target lane captured by a "merge_2" CAVs at Ramp 1



(b) "Merge 2" CAVs fail to merge to rightmost lane before Ramp 2.

Figure 2.6 A demonstration of "flaw" cases for rule-based model

In addition, in the actual highway scenario, the "popularity" of the ramps is generally not the same (i.e., some ramps are having higher probability to be chosen as the destination of vehicles). To further demonstrate the proposed model is robust under such discrepancy in ramp popularity, an extra experiment is conducted by tuning the ratio between number of merge_1 and merge_2 CAVs. In this setting, the total number of CAVs is still set as 20, but the ratios between 2 CAVs are set as 20:0, 15:5, 5:15 and 0:20. The evaluation metrics for this experiment are the mean and standard deviation of episode reward and the successful merging-out rate. From the result, the GCQ model is still robust over all the scenarios with highest episode reward and higher merging-out rate than LSTM-Q model. It is expected that when there exist both CAVs, the performance will drop since the model should distinguish the vehicles' intention and draw decisions accordingly, this can generally complicate the scenario. Another observation is when the ramp 2 becomes more popular, both GCQ models and LSTM-Q models can get higher performance in terms of both episode reward and merging-out rate. It is because in these scenarios, the merging vehicles can have longer operation range to make lane changes. For rule-based model, the successful merging-out rate is always 100%, but it can only get low reward since the operation efficiency is overlooked.

2.3 Concluding remarks

In this research, a DRL-based model combining GCN and deep Q network (GCQ) is proposed to control multiple CAVs within a CAV network to make collaborative lane-change decisions. From the CAVs operations perspective, the experiment results showed that the proposed model enables the CAVs to make successful lane changes to satisfy their individual intention of exiting the freeway from the intended exit ramps in a manner that is both safe and efficient. As part of efforts to achieve this overarching objective, this research also demonstrates the efficacy of the proposed model in: (a) resolving dynamic-number-agents problem (DNAP) specifically for the driving task with high model flexibility; (b) fusing information acquired by cooperative sensing on both local and global information; (c) making safe and collaborative decisions based on the fused information; (d) having enough robustness across scenarios with different traffic density and making consistent decisions without the need of retraining the model.

For a comparative evaluation, the results of the proposed GCQ control model were juxtaposed with those of the two other methods that served as the baseline for comparison: the classic "context extractor" LSTM-Q network and the traditional rule-based model calibrated from the human driving experiences. The results were unequivocal: the proposed model significantly outperforms both baseline methods. Specifically, compared to LSTM-Q model, the GCQ model has much fewer parameters and can be trained much faster, which indicates that the GCN layer can efficiently fuse essential information to generate driving decisions for the CAV. This model can be useful when developing CAV-related centralized control units such as RSUs or cloud computing platforms. In this research, all the CAV driving decisions are made instantly based only on the information at current timestep at the time of the decision.

In future research, as connectivity capabilities and data storage systems increase, it may be worthwhile to consider, as an input to the decision making, temporal information including historical data on the vehicle position, speed, and acceleration at different locations. Such temporal information could serve as an indicator of adverse traffic conditions such as accidents, workzones, potholes etc., that loom ahead and may encourage the CAVs to make longer-term proactive evasive decisions for avoiding trouble slots. Also, historical information can serve as a validation resource to ascertain the correctness of new information received from CAVs, which can further enhance the reliability of entire system.

In addition, future extensions to the methodology described in this research, could incorporate other powerful or newer supervised machine learning and classification algorithms that could reduce the computation time and learning speed, and consequently decrease CAV processing and decision speed, and road safety in the CAV era. These include Enhanced Probabilistic Neural Network, Dynamic Ensemble Learning Algorithm, Finite Element Machine and Neural Dynamic Classification algorithm (Ahmadlou and Adeli, 2010; Rafiei and Adeli, 2017b; Pereira et al., 2020; Rokibul Alam et al., 2020).

Finally, future research could investigate the design and evaluation of CAV operational controls that maximize some defined utility (with different combinations of the criteria types and levels in the reward function) of all vehicles in the entire corridor or overall road network rather than the CAVs only. This would consider the utility of not only CAVs but also detectable and

undetectable HDVs. This can be achieved through collaborative control of the CAVs to achieve systemwide utility, for example, using CAVs to mitigate traffic congestion, promote traffic string stability, or to reduce fuel consumption or emissions. These can be investigated in future research using the proposed GCQ model with other combinations of experiment settings and reward functions.

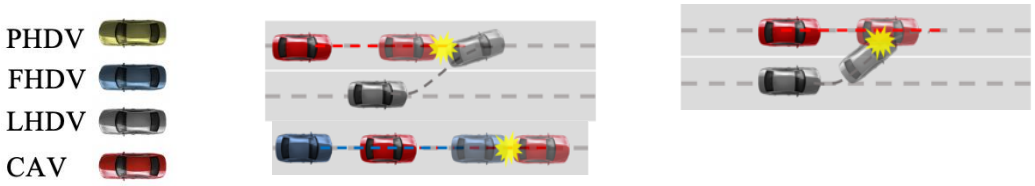
CHAPTER 3 A COOPERATIVE CRASH AVOIDANCE FRAMEWORK FOR AUTONOMOUS VEHICLE UNDER COLLISION-IMMIENT SITUATIONS IN MIXED TRAFFIC STREAM

3.1 Introduction

Traffic-related fatalities and injuries continue to pose a global concern (Sinha et al., 2007b). A recent national report indicated that traffic-related accidents are the second leading cause of death of ages 5~29, and the third leading cause of death of ages 30 ~44 (Anjuman et al., 2007). With increasing global population, travel demand, traffic fatalities and injuries are expected to increase. It has been estimated that approximately 95% of traffic crashes are related to human error. For this reason, vehicle automation which eliminates the human factor in vehicle control, is widely seen as a cure to persistent traffic fatalities and injuries (S. Chen et al., 2021; Dong et al., 2021; Noy et al., 2018; Rahman et al., 2019; Ye et al., 2019). In the mixed traffic era, human error can be separate from two perspectives: Perspective 1) human errors from inside of the vehicle, which can be solved by eliminating the human elements from vehicle control; Perspective 2) human errors from outside of the vehicle, which cannot be easily solved by minimize the human factor from the driver seated in the vehicle.

Most existing studies focus primarily on the first perspective of error, by considering how the AV can operate without compromising the safety of the neighboring HDVs (Jiajia Chen et al., 2013; Kalra et al., 2016; Kim et al., 2019; Koopman et al., 2017). Another problem is that, while AV controllers for motion planning do exist in literature, they are often designed to be conservative. This is because the primary focus has been to design and create resilient autonomous systems that will not lead to safety issues. However, as discussed previously, this approach focuses on eliminating new sources of error but does not effectively address the existing sources of error (human drivers) whose behaviors lead to collision. Thus, this research focuses on the second perspective of error to develop an AV controller to facilitate the safety of vehicles in the vicinity of AVs.

In situations involving hazardous roadway conditions, V2V technology enhances the safety of local system (Ha et al., 2020). First, with its larger range compared with on-board equipment, V2V connectivity allows the driver to receive information much faster, thereby providing greater reaction time during emergencies (Dong, Chen, Li, et al., 2020; Li et al., 2020). Secondly, V2V connectivity, unlike the on-board sensors, does not prone to occlusion or inclement weather. In other words, a connected vehicle still receives the needed information even when it is out of sight from another vehicle or entity (Dong, Chen, Joun Ha, et al., 2020). Existing studies on AV controllers do not recognize the cooperation between the connected HDVs and AVs. A connected autonomous vehicle (CAV) that is connected to its neighboring connected HDVs (CHDVs) can serve as a centralized, local decision maker that control the speed of the neighboring vehicles in a holistic bid to maximize overall safety with the cooperation of the CHDVs. The cooperative framework proposed in this project incorporates V2V technology capabilities between CHDVs and CAV.



Rear-end collision Side-impact collision

Figure 3.1 Crash patterns

Model predictive control (MPC) is an effective approach for solving problems that arise from motion planning. In literature, MPC is frequently used for solving the problem of vehicle path generation to mitigate collision (Babu et al., 2019; Ji et al., 2017; Shen et al., 2017; H. Wang et al., 2019; Werling et al., 2012). However, they mainly focused on the first perspective of human error (from inside the vehicle). Further, MPC does not necessarily result in closed loop stable systems. When the controller structures are too complicated, the stability is subsequently hard to reach using the final state constraint. Thus, researchers have explored various ways of testing and validating the stability of the MPC controller. Di Cairano and Bemporad (di Cairano et al., 2010) used the controller matching techniques to select the MPC weight matrices so that the resulting MPC controller not only behaves similar to the given linear controller but also is globally asymptotically stable.

This project adopts a different approach: sufficient condition for stability of the closed loop system (Simon et al., 2016). By using the Lyapunov function as the cost function, an optimization problem can be formulated as: $min(V_k^* - V_{k+1}^*)$, where V_k^* refers to the objective function at time k and V_{k+1}^* refers to the objective function at time k+1 respectively. For the MPC with different prediction horizon N_p, if there exist a negative $min(V_k^* - V_{k+1}^*)$, then the Lyapunov function V_k is inappropriate to ensure system stability. In summary, there are three main discussions addressed in this research:

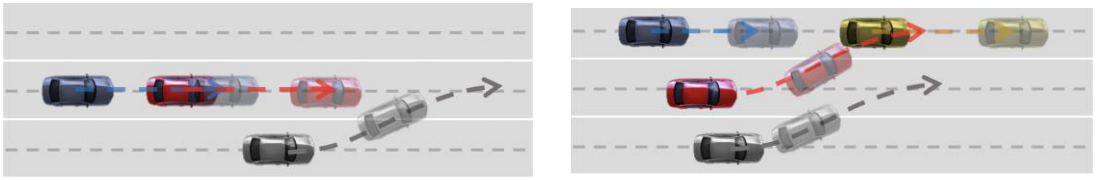
- The coping maneuvers of CAVs to the second perspective of human error.
- The benefits of implementing cooperative framework in the crash imminent situation.
- The choice of parameters in the MPC process to make the controlled system more stable.

3.2 Problem formulation

This project considers the following common types of collisions (Xu et al., 2019): (a) side-impact collision and (b) rear-end collision under a lane-change situation due to HDV driver error. We focus on the enhanced safety benefits by combining the automation and connectivity. Thus, the vehicles included in this research are (Figure 3.1): connected autonomous vehicle (CAV,

colored red); lane-change human-driven vehicle (LHDV, colored gray), which do not have connectivity and act aggressively; connected human-driven vehicles (CHDVs) in the following and preceding positions (FHDV colored blue, and PHDV colored yellow, respectively).

Based on the combination of the collision types and vehicle types two CAV crash avoidance maneuvers (Figure 3.2) are considered. The first scenario is the deceleration maneuver of the CAV to avoid the potential rear-end collision by the lane-change LHDV. When the longitudinal positions of the LHDV and CAV are nearly the same, once the LHDV lane-change process begins, it is difficult for the CAV to avoid collision even when it engages in maximum deceleration. Thus, to avoid the possible side-impact collision (second scenario) the autonomous vehicle has only the choice of lane change to the other lane.



Deceleration maneuver Lane-change maneuver

Figure 3.2. Collision-avoidance Maneuver

Following assumptions are made in this research: 1. Vehicles are all light passenger vehicles that shares same dynamic and static features (e.g., 4-meter length); 2. Each vehicle is represented by a buffer circle, and the initial diameter of buffer circles is 6 meters; lane width is 3.7 meters, and the vehicle speeds on the road are consistent with traveling speeds at the highway class. 3. The lane-change trajectory of the LHDV is assumed to follow a cubic polynomial shape and is predicted reliably (Yang et al., 2018). 4. The LHDV acceleration is assumed to be highly aggressive with little regard to its surroundings.

3.3 Methodology

This section presents the overall framework of the controller. The objective of the mathematical model in this research is to determine the optimal crash avoidance maneuvers (deceleration and lane-change maneuvers) and optimal deceleration/acceleration decisions. To determine the optimal control maneuvers, a vehicle interaction based bi-level optimization problem is formulated. The methodology consists of the following components: A) control framework, B) LHDV motion prediction, C) MPC controller design and bi-level optimization problem considers different vehicle interactions, and D) sufficient stability condition.

3.3.1 Control framework

The proposed control framework deals with multiple vehicles in a crash-imminent situation. Therefore, this project uses MPC controller, which handles multiple constraints. The motion of the

LHDV is regarded as an important reference for the CAV's decision making. Figure 3.3 presents the general structure of the proposed control framework. The CAV controller considers the two maneuvers based on the LHDV's motion (trajectory and speed) in a hierarchical structure. Begin with deceleration maneuver (because it is inherently less disruptive). If the deceleration maneuver is insufficient or is inadequate for a given crash-imminent situation, the controller seeks the other alternative: lane-change maneuver. The MPC is formulated as an optimization problem, the controlled variables are acceleration/deceleration of the controlled vehicles: CAV and CHDVs.

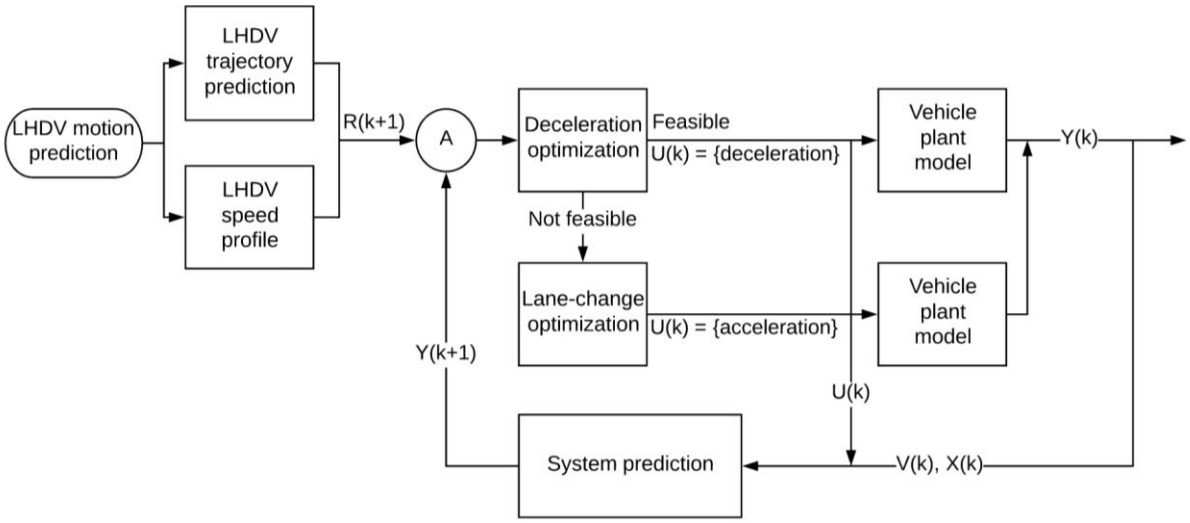


Figure 3.3. Control framework

3.3.2 LHDV motion prediction

LHDV Motion prediction consists of two parts: trajectory prediction and speed profile generation. The trajectory of the aggressive LHDV is incorporated into the CAV controller, and the predicted aggressive trajectory at each time step is assumed by a cubic polynomial curve ($y_t(x_t) = tan\theta_t^i x_t + \frac{3y_t^t - 2x_t^e tan\theta_t^i}{(x_t^e)^2} x_t^2 + \frac{x_t^e tan\theta_t^i - 2y_t^e}{(x_t^e)^3} x_t^3$) (Yang et al., 2018), which has second-order smoothness. The LHDV positions are represented by (x_t, y_t), where x_t and y_t denote the longitudinal and latitudinal positions of LHDV at time step t. θ_t^i represents the initial course angle of the LHDV at time step t, which is the angle between the moving direction and the $x - axis$. The ending position of the lane change trajectory is calculated by implementing rollover-free conditions, which are represented by (x_t^r, y_t^e). The rollover-free condition maintaining the aggressive behavior of the lane-change vehicle while avoid rollover collision, which can be

represented as: $x_t^r = \sqrt{6}\frac{y_t^e u_t^i}{\sqrt{y_t^e a_s^r}}$, with u_t^i represents the initial velocity towards moving direction at time step t.

The speed profile is generated with the purpose of completing the lane change process as fast as possible. Thus, the LHDV accelerates throughout the lane-change process. The aggressive acceleration $a = \frac{2\left(\frac{6y_t^e u_t^i}{\sqrt{6y_t^e a_s^r}} - u_t^i \tau\right)}{\tau^2}$ which is calculated based on the rollover-free conditions as well as the length of each time step: τ.

3.3.3 MPC controller design and bi-level optimization

The controlled vehicles are CAVs and the surrounding CHDVs (FHDV and PHDV). The controller only controls their longitudinal acceleration in both directions. Motion models of PHDV and FHDV are based on acceleration and deceleration respectively. At each time step, the motion model can be represented as discrete-time model:

$$X(k+1) = AX(k) + BU(k) \quad (1)$$

$$Y(k+1) = CX(k+1) \quad (2)$$

$$X(k) = \begin{bmatrix} x(k) \\ v(k) \end{bmatrix}, A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1/2\Delta t \\ \Delta t \end{bmatrix} \quad (3)$$

$x(k), v(k)$ and $a(k)$ are the longitudinal position, velocity, and the acceleration/deceleration of the controlled vehicles. $U(k)$ represents the controlled variable, which is the acceleration on both longitudinal directions ($a(k)$).

In the MPC design, two crucial factors are N_p and N_c. N_p represents the prediction horizon, which is the number of future control intervals that the MPC evaluates. N_c represents the control horizon, which is the number of control actions to be optimized in the control interval. In this project, $N_p = N_c + 1$. Based on the MPC control strategy, the initial value is implemented, and the calculations will be repeated at each time step. With $x(k+1) = Ax(k) + Bu(k)$, the predicted output for control interval N_c, N_p can be represented as:

$$x(k+N_c) = A^{N_c}x(k) + A^{N_c-1}Bu(k) + \cdots + Bu(t+N_c-1) \quad (4)$$

$$x\left(k+N_p\right) = A^{N_p}x(k) + A^{N_p-1}Bu(k) + \cdots + ABu(k+N_c-1) + Bu(t+N_c-1) \quad (5)$$

The system prediction can be rewritten in a more compact form as: $X(k+1) = M_x x(k) + M_u U(k)$, where $X(k+1)$ and $U(k)$ represent the output and input sequence, M_x and M_u represent the parameters in the system of equations:

$$X(k+1) = \begin{matrix} x(k+1) \\ x(k+2) \\ x(k+3) \\ \vdots \\ \lfloor x(k+N_p) \rfloor_{N_p \times 1} \end{matrix} \qquad U(k) = \begin{bmatrix} u(k) \\ u(k+1) \\ \vdots \\ u(k+N_c-1) \end{bmatrix}_{N_c \times 1} \tag{6}$$

$$M_x = \begin{bmatrix} A \\ A^2 \\ \vdots \\ A^{N_p} \end{bmatrix}_{N_p \times 2} \qquad M_u = \begin{matrix} B & 0 & \dots & \dots & 0 \\ AB & B & 0 & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A^{N_c-1}B & A^{N_c-2}B & \dots & \ddots & B \\ \lfloor A^{N_p-1}B & A^{N_p-2}B & \dots & \dots & (A+1)B \rfloor_{2N_p \times N_c} \end{matrix} \tag{7}$$

The MPC controller is designed based on the vehicles' interactions. In both deceleration and lane change maneuvers, two types of vehicle interactions are considered:

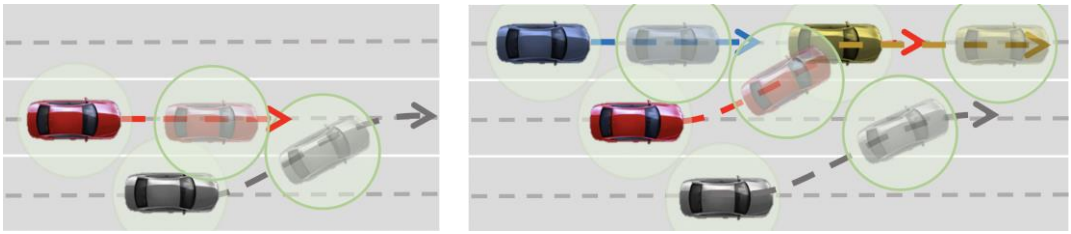
- interaction between the lane-change vehicles and the vehicles on the target lane
- interaction between the vehicles both on the target lane

These two interactions constitute the two levels in the MPC controller. Thus, the optimization problem in the controller can be formulated as a bi-level optimization problem. The controlled variables u_{Vi}, δ_{Vi} are the acceleration/deceleration and speed violation of the controlled vehicles V_i. As shown in equation (8), the lower level of the bi-level optimization problem is focused on the interaction between the lane-change vehicle and vehicle on target lane V_1, the upper level focused on vehicles V_1, V_2 (adjacent to V_1) both on the target lane that affect the control decision mutually.

$$\min_{u_{v1}\delta_{v1},u_{v2},\delta_{v2}} F(u_{V1}, \delta_{V1}, u_{V2}, \delta_{V2}) \tag{8}$$
$$subject.to.$$
$$u_{V1}, \delta_{V1} \in \underset{u_{V1},\delta_{V1}}{\operatorname{argmin}} \{ f(u_{V1}, \delta_{V1}, u_{V2}, \delta_{V2}) : g_j(u_{V1}, \delta_{V1}, u_{V2}, \delta_{V2}) \le 0, j = 1, \dots, J \}$$
$$G_i(u_{V1}, \delta_{V1}, u_{V2}, \delta_{V2}) \le 0, i = 1, \dots, I$$

Lower level: interaction between the lane-change vehicles and target lane vehicles is implemented through vehicle buffer circles' tangent situations. The control inputs need to fulfill the safety requirements of the lane-change vehicle. As shown in Figure 3.4, in the deceleration maneuver, the tangent situation $(x_{LHDV} - x_{AV})^2 + (y_{AV} - y_{LHDV})^2 = 4r^2$ is critical to avoid the crash between the lane-change LHDV and target lane CAV. When the deceleration of CAV is not in the feasible range, the deceleration maneuver will be aborted in favor of the lane-change maneuver. The CAV becomes the lane-change vehicle, and the CHDVs (PHDVs, FHDVs) on the target lane are the target lane vehicles. There are two tangent conditions: PHDV-CAV and FHDV-CAV.

To guarantee lane-change safety, distance and speed constraints are added to the lower level. The longitudinal distances between the lane-change vehicle and target lane vehicles need to be greater than l_1, the target lane vehicles need to have a smaller or equal speed than the lane-change vehicle. However, when the target lane vehicles on the preceding position, the target lane vehicles need to have an equal or larger speed than the lane-change vehicle.



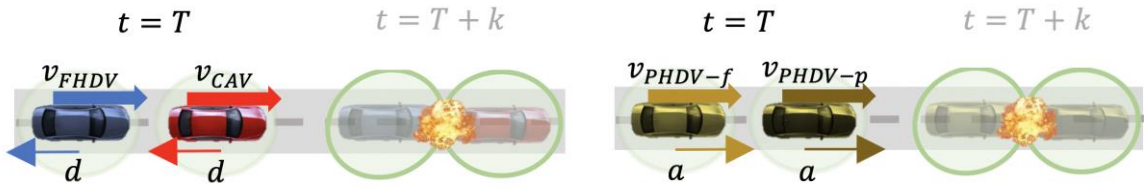
Deceleration maneuver Lane-change maneuver

Figure 3.4 Interaction between lane-change and target lane vehicles

Upper Level: In both maneuvers, the adjacent vehicles both on the target lane will affect each other on their control decisions. When both controlled vehicles are decelerating on the target lane at time step $t = T$, (e.g., in Figure 3.5(1), the CAV-FHDV in the deceleration maneuver), or both accelerating on the target lane at time step $t = T$, (e.g., in Figure 3.5(2), PHDV-p and PHDV-f in the lane-change maneuver), challenges arise as the vehicles might have rear-end collisions in the future time step $t = T + k$. With the maximum acceleration/deceleration threshold being relaxed with connectivity, feasible deceleration for the vehicle in the preceding position is: $max\{f\left(d_{dV_f}^{max}\right), d_{dV_p}^{max}\}$, and feasible acceleration for the vehicle in the following position is: $min\{f\left(a_{aV_p}^{max}\right), a_{aV_f}^{max}\}$ ($dV_j, aV_j \;\; j = \{p, f\}$ represent deceleration/acceleration of vehicle V at position j). To further improve the system safety, the longitudinal distances between the vehicles must maintain a value greater than or equal to l_2, which affects the headway directly (As shown in equation (9-10). Δx is the initial distance between the vehicles, the speed difference is represented by Δv.

$$f\left(d_{dV_f}^{max}\right) = \frac{2l_2 - 2(\Delta x^i + \Delta v^i \tau)}{\tau^2} + d_{dV_f}^{max} \quad (9)$$

$$f\left(a_{aV_p}^{max}\right) = \frac{-2l_2 + 2(\Delta x^i + \Delta v^i \tau)}{\tau^2} + a_{aV_p}^{max} \quad (10)$$



(1) Rear-end of decelerating vehicle (2) Rear-end of accelerating vehicle

Figure 3.5 Interaction between target lane vehicles

The detailed objective function of lower level and upper level can be formulated as follows:

Lower level:

$$\min_{\substack{\delta_{V1} \\ u_{V1}}} \sum_{n=1}^{N_p} \|x_{V1}(k+n) - x_{LV}(k+n)\|_Q^2 + \sum_{n=1}^{N_c} \|u_{V1}(k+n-1)\|_R^2 + \|\delta_{V1}(k+n-1)\|_P^2 \quad (11)$$

The objective function consists of tracking the aggressive lane-change vehicle (LV), the control inputs u_{V1}, velocity soft constraints δ_{V1}, and P, Q, R are the weight parameters. $x_{LV}(.)$ represents the information of the LV, which includes the longitudinal location and velocity. Since the controlled vehicle V_1 can be in preceding or following longitudinal position of the LV, two constraint sets are needed. The constraints for the decelerating vehicles with $n = 1, \dots, N_p$:

$$x_{V1}(k+1) = Ax_{V1}(k) + Bu_{V1}(k) \quad (12)$$

$$x_{V1}(k+n) - r_{LV}(k+n) + \begin{bmatrix} l_1 \\ -\delta_{V1}(n) \end{bmatrix} \leq 0 \; for \; n = 1, \dots, N_c \quad (13)$$

$$x_{V1}(k+n) - r_{LV}(k+n) + \begin{bmatrix} l_1 \\ 0 \end{bmatrix} \leq 0 \; for \; n = N_p \quad (14)$$

$$d_{V1}^{max} \leq u_{AV}(k+n-1) \leq 0 \quad (15)$$

$$f(d_{V2}^{max}) \leq u_{AV}(k+n-1) \leq 0 \quad (16)$$

$$r_{LV}(k+n) - x_{AV}(k+n) + \begin{bmatrix} \sqrt{4r^2 - \left(l_{V1}^y(k+n) - l_{LV}^y(k+n)\right)^2} \\ -\delta_{V1}(n) \end{bmatrix} \leq 0 \; for \; n = 1, .., N_p \quad (17)$$

$$\delta_{V1}(k+n) \geq 0 \quad (18)$$

The constraints for the accelerating vehicles are similar but with opposite symbols.

$$r_{LV}(k+n) - x_{V1}(k+n) + \begin{bmatrix} l_1 \\ -\delta_{V1}(n) \end{bmatrix} \leq 0 \; for \; n = 1, \dots, N_c \quad (19)$$

$$r_{LV}(k+n) - x_{V1}(k+n) + \begin{bmatrix} l_1 \\ 0 \end{bmatrix} \leq 0 \; for \; n = N_p \quad (20)$$

$$0 \leq u_{V1}(k+n-1) \leq a_{V1}^{max} \quad (21)$$

$$0 \le u_{V1}(k+n-1) \le f(a_{V2}^{max}) \quad (22)$$

$$x_{V1}(k+n) - r_{LV}(k+n) + \begin{bmatrix} \sqrt{4r^2 - \left(l_{V1}^{y}(k+n) - l_{LV}^{y}(k+n)\right)^2} \\ \delta_{V1}(n) \end{bmatrix} \le 0 \; for \; n = 1, .., N_p \quad (23)$$

$$\delta_{V1}(k+n) \le 0 \quad (24)$$

Constraints (12) is associated with the vehicle dynamics, (13), (14), (19), (20) represent the distance constraints which are applied to ensure the longitudinal distance safety requirements. Constraints (15), (16), (21), (22) are the controlled variable constraints. (17) and (23) represents the collision avoidance tangent situation. Constraints (20), (24) are the soft constraints regarding the speed.

Upper level:

$$\min_{\substack{u_{V1} \\ \delta_{V1} \\ u_{V2} \\ \delta_{V2}}} \left(\sum_{n=1}^{N_c} \begin{matrix} \sum_{n=1}^{N_p} \|x_{V2}(k+n) - x_{V1}(k+n)\|_Q^2 + \\ \|u_{V1}(k+n-1)\|_R^2 + \|u_{V2}(k+n-1)\|_R^2 \\ + \|\delta_{V1}(k+n-1)\|_P^2 + \|\delta_{V2}(k+n-1)\|_P^2 \end{matrix} \right) \quad (25)$$

On the upper level, $V1$ is the target lane vehicle interact with LV and $V2$ is the vehicle on the target that adjacent to $V1$. The constraints for the decelerating vehicles can be described as follows:

$$u_{V1}, \delta_{V1} \in argmin_{u_{V_1}, \delta_{V1}} \{ f(u_{V1}, \delta_{V1}, u_{V2}, \delta_{V2}) \quad (26)$$

$$: g_j(u_{V1}, \delta_{V1}, u_{V2}, \delta_{V2}) \le 0, j = 1, \dots, J \} \quad (27)$$

$$x_{V2}(k+1) = Ax_{V2}(k) + Bu_{V2}(k) \quad (28)$$

$$x_{V2}(k+n) - x_{V1}(k+n) + \begin{bmatrix} l_2 \\ -\delta_{V2}(n) \end{bmatrix} \le 0 \; for \; n = 1, \dots, N_c \quad (29)$$

$$x_{V2}(k+n) - x_{V1}(k+n) + \begin{bmatrix} l_2 \\ 0 \end{bmatrix} \le 0 \; for \; n = N_p \quad (30)$$

$$d_{V1}^{max} \le u_{V2}(k+n-1) \le 0 \quad (31)$$

$$\delta_{V2}(k+n) \ge 0 \; n = 1, \dots, N_p \quad (32)$$

The accelerating vehicles on the target lane, the constraints are similar with opposite symbols:

$$x_{V2}(k+n) - x_{V1}(k+n) - \begin{bmatrix} l_2 \\ \delta_{V2}(n) \end{bmatrix} \ge 0 \; for \; n = 1, \dots, N_c \quad (33)$$

$$x_{V2}(k+n) - x_{V1}(k+n) - \begin{bmatrix} l_2 \\ 0 \end{bmatrix} \ge 0 \; for \; n = N_p \quad (34)$$

$$0 \le u_{V2}(k+n-1) \le a_{V2}^{max} \quad (35)$$

$$\delta_{V2}(k+n) \le 0 \; k = 1, \dots, N_p \quad (36)$$

Constraint (26) represents the lower optimization problem that considered in the upper level.

3.3.4 Sufficient stability condition

As is well-known that MPC controllers do not guarantee internal stability, it is essential to analyze and ensure the stability of the MPC controller in this experiment. To ensure internal stability, it is common to add a final state constraint or final state penalty. However, for the MPC controller in a complex system, it is difficult to show stability using the final state constraints/final state penalty (Simon et al., 2016). All the weights are chosen to guarantee the convexity of the cost function to use the KKT condition to change the bi-level MPC to a single-level optimization problem. The sufficient condition can be proven by showing the value function is decreasing between two consecutive time steps. ($V_k^* - V_{k+1}^* \leq 0$) for any k. If the value: $V_k^* - V_{k+1}^*$ is smaller than or equal to zero for any k, which means V_k is a valid Lyapunov function. Thus, the system can be stabilized by the bi-level MPC controller. However, the test is only sufficient, not necessary. The system might be stable, but the Lyapunov function may not be valid for the closed loop system. However, with the sufficient stability test, there is a greater chance to guarantee the stability of the system.

For both maneuvers, the Lyapunov functions are set to be the cost function of the optimization control problem, which are the higher-level objective functions of the bi-level optimization systems. Different prediction horizon values should be tested based on the sufficient conditions for stability. In this research, the range of the prediction horizon is $N_p = 3, \dots, 7$, the values of the cost functions are tested to check the stability of the system of vehicles after being controlled by the bi-level MPC in deceleration/lane-change maneuver.

3.4 Results

This section illustrates the crash avoidance framework combines the two crash avoidance maneuvers. The experimental simulation is implemented in MATLAB. The simulated time step τ is 0.2s and the weights for the objective functions are chosen as: $[P, Q, R] = [15,10,10]$. The safety requirements for the safety distance between the lane-change vehicles and the vehicles on the target lane is $l_1 = 5m$, the safety distance between the vehicles on the target lane is $l_2 = 10m$. The maximum deceleration as well as acceleration are assumed to be $5.08m/s^2$(Bae et al., 2019).The maximum longitudinal acceleration is assumed to be $3.024m/s^2$(Bokare et al., 2017). We have the y axis origin as the middle of LHDV: $l_{AV}^y = 3.7m$.

3.4.1 Sufficient stability analysis

The initial state of the LHDV in this numerical case is set as follows: location is (5,0) (longitudinal location is 5m), velocity is $17.88m/s$ $(40mph)$. The initial longitudinal bumper-to-bumper distance between the CAV and LHDV is considered as 5~6m. The bumper-to-bumper distance between the CAV and FHDV is considered in the range 5~34m. The CAV's velocity range is set as $17m/s$~$21m/s$, and that of the FHDV is set as $18m/s$~$22m/s$. To choose a proper prediction horizon to make sure the system has greater chance of attaining stability, a sufficient test for stability is implemented.

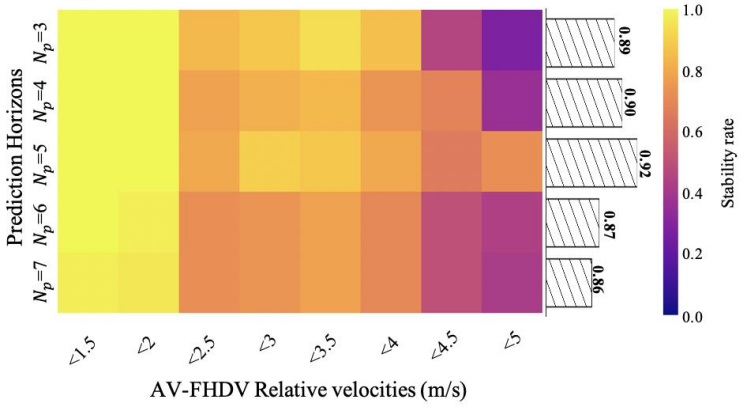


Figure 3.6 Sufficient stability test

As shown in Figure 3.6, the range of the prediction horizon N_p considered in this research is set to be 3 to 7. When the controlled vehicles have different relative velocities, the stability rate will change. The higher the relative velocity, the harder the system to be stable. When N_p=5, the system has the highest probability to be sufficiently stabled, which means the initial states set of the system when N_p=5 is the largest.

3.4.2 Collision Avoidance Framework Success Rate

The CAV velocity in the infeasible cases of the deceleration maneuver is from $17m/s$ to $21m/s$. Multiple lane changes needed to be considered because of the CAV's different initial velocities in the lane-change maneuver. The sufficient stability test is taken based on different CAV lane change motions, and the stabilities vary due to the different CAV lane changes. When the speed is medium ($19m/s$ to $20m/s$), the stability rates are the highest. We choose N_p=5 for the MPC controller of lane-change maneuver because it gives us the highest overall stability among all the prediction horizons.

Figure 3.7 presents the success rates of the deceleration maneuver and the deceleration + lane-change maneuvers. As observed in Figure 3.7(a), the success rate is quite low (<40%) when the bumper-to-bumper distance is smaller than 7m. Also, when the relative velocities are large, the success rates are low as well. However, as shown in Figure 3.7(b), after taking the lane-change maneuver into account, the success rates are greatly enhanced under various situations. The least successful case, which has the largest relative velocity and the smallest bumper-to-bumper distance, has success rate of more than 60%. For the remaining situations, the success rates exceed 90%. For cases with small relative velocity, the success rates are 1, that is, 100%. High success rates are achieved in most cases using the proposed control framework by combining the two maneuvers.

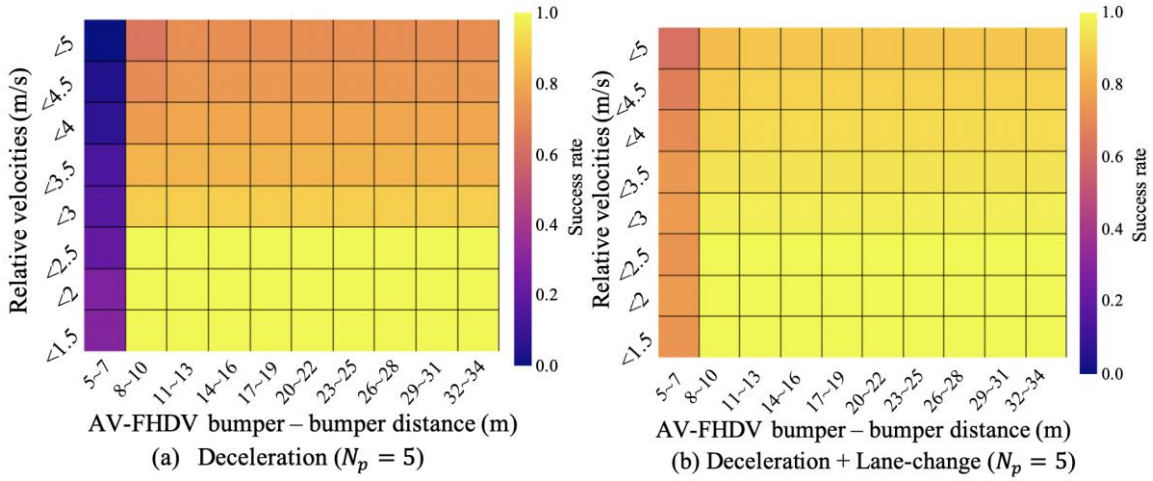


Figure 3.7 Success rates of maneuvers

3.5 Concluding remarks

In recognition of the safety hazards posed by errant behaviors of the human driven vehicles in mixed traffic streams during the prospective AV transition era, this project proposes a Model Predictive Control framework with V2V connectivity. This framework, which was developed to control two maneuvers of the CAV and connected HDVs (deceleration and lane-change), is motivated by the need to address human error associated with the drivers of HDVs. The control framework is demonstrated using a numerical example that considers various traffic situations in terms of the initial velocities and initial locations of the controlled vehicles. The results suggest that the control framework can effectively help the CAV and connected HDVs to avoid collision in crash imminent situations. In this regard, the framework was shown to achieve a minimum of 90% average success rate of collision avoidance throughout the LHDV lane-change process. The success rate was found to reach 100% in certain specific traffic situations.

CHAPTER 4 SYNOPSIS OF PERFORMANCE INDICATORS

4.1 Part I

Two (2) transportation-related courses were offered annually during the study period that was taught by the PI and a teaching assistant who are associated with the research project. One of these was a newly developed course inspired and directly associated with CCAT research. Three graduate students and a post-doctoral researcher (subsequently designated a Visiting Assistant Professor) participated in the research project during the study period. One (1) transportation-related advanced degree (doctoral) program utilized the CCAT grant funds from this research project, during the study period to support graduate students. This research project was leveraged to obtain $210,000 in additional funding from the Indiana DOT titled "Integrating Transformative Technologies in Indiana's Transportation Operations".

4.2 Part II

Research Performance Indicators: 2 journal articles and 2 conference articles were produced from this project. The research from this advanced research project was disseminated to over 1,100 people from industry, government, and academia, through 22 conference presentations. These include the 2020 Purdue Road School, the 2020 Next Generation Transportation Systems Conference (NGTS), the 2019 ITE (Purdue Chapter) Annual Dinner, the 2019 TRB annual meeting, the 2020 TRB annual meeting, the 2021 TRB annual, the 2022 TRB annual meeting, the 2020 INFORMS Annual Meeting, and the 2022 ASCE International Conference on Transportation and Development. One (1) other related research project was funded by a source other than UTC and matching fund sources. At the time of writing, there are no new technologies, procedures/policies, and standards/design practices that were produced by this research project.

Leadership Development Performance Indicators: This research project generated 5 academic engagements and 2 industry engagements. The PI held positions in 2 national organizations that address issues related to this research project. One of the CCAT students who worked on this project holds a leadership position.

Education and Workforce Development Performance Indicators: The methods, data and/or results from this study are being incorporated in the syllabus for the next version (Fall 2022) of the following courses at Purdue University: (a) CE 561: Transportation Systems Evaluation, a mandatory graduate level course at Purdue's transportation engineering M.S. and Ph.D. programs, (b) CE 299: Smart Mobility, an optional undergraduate level course at Purdue' civil engineering B.S. program, and (c) CE 398: Introduction to Civil Engineering Systems, a

mandatory undergraduate level course at Purdue University's civil engineering program. These students will soon be entering the workforce. Thereby, the research helped enlarge the pool of people trained to develop knowledge and utilize the at least a part of the technologies developed in this research, and to put them to use when they enter the workforce.

Collaboration Performance Indicators: There was collaboration with other agencies, and 1 agency provided matching funds.

The outputs, outcomes, and impacts are described in Chapter 5 below.

CHAPTER 5 STUDY OUTCOMES AND OUTPUTS

5.1 Outputs

5.1.1 Publications, conference papers, or presentations

(a) Journal Publications

- Dong, J., Chen, S., Li, Y., Du, R., Steinfeld, A., Labi, S. (2021). Space-weighted Information Fusion Using Deep Reinforcement Learning: The Context of Tactical Control of Lane-Changing Autonomous Vehicles and Connectivity Range Assessment. *Transportation Research Part C: Emerging Technologies, Vol. 128, July 2021, 103192* https://www.sciencedirect.com/science/article/abs/pii/S0968090X21002084

- Chen, S., Dong, J., Ha, P., Li, Y., Labi, S. (2021). Graph Neural Network and Reinforcement Learning for Multiagent Cooperative Control of Connected Autonomous Vehicles. *Computer-Aided Civil and Infrastructure Engineering, 36(7), 838–857.* https://onlinelibrary.wiley.com/doi/abs/10.1111/mice.12702

(b) Conference Publications

- Dong, J., Chen, S., Li, Y., Ha, P. Y. J., Du, R., Steinfeld, A., Labi, S. (2020). Spatio-weighted information fusion and DRL-based control for connected autonomous vehicles. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC) Proceedings, IEEE.* https://ieeexplore.ieee.org/abstract/document/9294550

- Du, R., Chen, S., Li, Y., Ha, P. Y. J., Dong, J., Anastasopoulos, P. C., & Labi, S. (2021, September). A cooperative crash avoidance framework for autonomous vehicle under collision-imminent situations in mixed traffic stream. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC) Proceedings*, pp. 1997-2002. IEEE. https://ieeexplore.ieee.org/abstract/document/9564937

(c) Conference Presentations

- Chen, S., Leng, Y., Labi, S. (2020). Direct Characterization of the Driving Environment using a Deep Learning Algorithm for Purposes of Autonomous Driving Simulation, *Transportation Research Board (TRB) 99th Annual Meeting*. Washington, D.C., USA. January 12-16, 2020.

- Ha, P., Chen, S., Dong, J., Du, R., Li, Y., Labi, S. (2020). Congestion mitigation in physical bottlenecks via deep reinforcement learning. *2020 INFORMS Annual Meeting*, Virtual, November 7-13, 2020.

- Ha, P., Chen, S., Du, R., Dong, J., Li, Y., Labi, S. (2020). Vehicular connectivity and its impacts on next generation transportation. *2020 INFORMS Annual Meeting*. Virtual, November 7-13, 2020.
- Li, Y., Chen, S., Ha, P., Dong, J., Du, R., Steinfeld, A., Labi, S. (2020). Leveraging vehicle connectivity and autonomy to stabilize flow in mixed traffic conditions. *2020 INFORMS Annual Meeting*. Virtual, November 7-13, 2020.
- Li, Y., Chen, S., Du, R., Ha, P., Dong, J., Labi, S. (2020). Leveraging trajectory-data calibrated car-following models and vehicle connectivity and autonomy to stabilize mixed traffic flows. *2020 INFORMS Annual Meeting*. Virtual, November 7-13, 2020.
- Dong, J., Chen, S., Ha, P., Li, Y., Du, R., Steinfeld, A., Labi, S. (2020). A deep reinforcement learning based multi-agent control system for vehicular networks. *2020 INFORMS Annual Meeting*. Virtual, November 7-13, 2020.
- Dong, J., Chen, S., Ha, P., Du, R., Li, Y., Labi, S. (2020). A deep reinforcement learning based framework for information fusion and control for connected and autonomous vehicle. *2020 INFORMS Annual Meeting*. Virtual, November 7-13, 2020.
- Du, R., Chen, S., Li, Y., Ha, P., Dong, J., Labi, S. (2020). Control framework for autonomous vehicles collision avoidance based on model predictive control under mixed traffic flow. *2020 INFORMS Annual Meeting*. Virtual, November 7-13, 2020.
- Du, R., Chen, S., Li, Y., Dong, J., Ha, P., Labi, S. (2020). A cooperative control framework for lane changing AV considers the mixed traffic flow. *2020 INFORMS Annual Meeting*. Virtual, November 7-13, 2020.
- Du, R., Chen, S., Li, Y., Ha, P., Dong, J., Labi, S. (2020). Model predictive control based AV collision-avoidance framework in mixed traffic conditions. *2020 Conference on Next Generation Transport System (NGTS 2020),* Virtual, December 28-31, 2020.
- Ha, P., Chen, S., Dong, J., Du, R., Li, Y., Labi, S. (2020). Using deep reinforcement learning techniques to mitigate traffic bottlenecks. *2020 Conference on Next Generation Transport System (NGTS 2020),* Virtual, December 28-31, 2020.
- Dong, J., Chen, S., Ha, P., Li, Y., Du, R., Steinfeld, A., Labi, S. (2020). Information integration using deep reinforcement learning for CAV operations and control. *2020 Conference on Next Generation Transport System (NGTS 2020),* Virtual, December 28-31, 2020.
- Li, Y., Chen, S., Du, R., Ha, P., Dong, J., Labi, S. (2020). Smoothening mixed traffic flow using vehicle connectivity and autonomy features. *2020 Conference on Next Generation Transport System (NGTS 2020),* Virtual, December 28-31, 2020.

- Ha, P., Chen, S., Dong, J., Du, R., Labi, S. (2021). Leveraging the capabilities of connected and autonomous vehicles and multi-agent reinforcement learning to mitigate highway bottleneck congestion. *Transportation Research Board 100th Annual Meeting*, Washington, DC.
- Du, R., Chen, S., Li, Y., Ha, P., Dong, J., Labi, S. (2021). Collision avoidance framework for autonomous vehicles under crash imminent situations. *Transportation Research Board 100th Annual Meeting*, Washington, DC.
- Du, R., Chen, S., Li, Y., Dong, J., Ha, P., Labi, S. (2021). A cooperative control framework for CAV lane change in a mixed traffic environment. *Transportation Research Board 100th Annual Meeting*, Washington, DC.
- Li, Y., Chen, S., Du, R., Ha, P., Dong, J., Labi, S. (2021). Using empirical trajectory data to design connected autonomous vehicle controllers for traffic stabilization. *Transportation Research Board 100th Annual Meeting*, Washington, DC.
- Dong, J., Chen, S., Ha, P., Li, Y., Du, R., Steinfeld, A., Labi, S. (2021). A framework for controlling connected autonomous vehicle movements using deep reinforcement learning and spatially-weighted data integration. *Transp. Research Board 100th Annual Meeting*, Washington, DC.
- Dong, J., Chen, S., Ha, P., Li, Y., Du, R., Labi, S. (2022). Reason induced visual attention for explainable autonomous driving. *Transp. Research Board 101st Annual Meeting*, Washington, DC.
- Dong, J., Chen, S., Ha, P., Li, Y., Du, R., Labi, S. (2022). Development and testing of an image transformer for explainable autonomous driving systems, *Transp. Research Board 101st Annual Meeting*, Washington, DC.
- Dong, J., Chen, S., Ha, P., Li, Y., Du, R., Labi, S. (2022). A cooperative multi lane-change framework for connected autonomous vehicles in a mixed-traffic environment, *2022 ASCE International Conference on Transportation and Development*, Seattle, Washington, May 31–June 3, 2022.

5.1.2 Other products

Other products of this research are as follows:

- A set of analytical models that describe AI-based and control-based systems for safe and efficient operations of connected and autonomous vehicles.
- Material for the Purdue Graduate course "CE 597 – Artificial intelligence and machine learning for autonomous vehicle operations."
- Research material and datasets to support future research related to the subjects of multi-level control for safe and efficient operations of connected and autonomous vehicles.

5.2 Outcomes

The outcomes of this project are the prospective changes that can be made to the transportation system, or its regulatory, legislative, or policy framework, resulting from research and development outputs. These are:

- Increased understanding and awareness of the need for widespread vehicle connectivity among CAVs and HDVs
- Strong justification to both CAV company and DOT's investment in installing connectivity facilities; and for CAV manufacturers, technology companies, and the road agencies to invest in connectivity equipment and facilities

5.3 Impacts

The impacts of this project are the effects of outcomes on the transportation system, or society in general, such as reduced fatalities, decreased capital or operating costs, community impacts, or environmental benefits. This includes how the research outcomes can potentially improves the operation and safety of the transportation system, increase the body of knowledge and technologies, enlarges the pool of people trained to develop knowledge and utilize new technologies and put them to use, and improve the physical, institutional, and information resources that enable people to have access to training and new technologies. A list of specific impacts from this research project, are as follows:

- The impacts of the part 1 of this research will hopefully give a strong justification to both CAV company and DOT's investment in installing connectivity facilities, and that investments in connectivity facilities can greatly benefit the entire transportation system by enhancing mobility and safety. We expect that the development of an innovative AI for CAV controls will yield positive effects on the transport system and society in general. These includes reduced crashes, travel efficiency (reduced travel time) which

translate into lower vehicle operating costs, higher economic productivity and more free time for social activities).

- In part 2 of the research, it is anticipated that the proposed research will provide strong justification for CAV manufacturers, technology companies, and the road agencies to invest in connectivity equipment and facilities, and therefore, will have a higher stake in CAV deployment. We expect that the research will provide proof that connectivity-equipped AVs and connectivity investments for HDVs can greatly benefit the entire traffic stream in the sense that it will enhance operational efficiency and mobility.
- In part 3 of the research, the study product is expected to be impactful to three key stakeholders: the public (who will be provided greater confidence in the safety of AV operations); and the AV and HDV manufacturers (who will be motivated or mandated to install connectivity in their vehicles); and road agencies (who will be incentivized or mandated to installing connectivity facilities along the roadways). It is anticipated that the research product will offer proof that investments in connectivity equipment and facilities can have profound safety benefits for road users overall. After such safety benefits associated with CHDVs are demonstrated, it is expected that CHDVs ownership will be promoted.

REFERENCES

AASHTO. (2018). *Infrastructure Needs for Autonomous Vehicles*.

Anjuman, T., Hasanat-E-Rabbi, S., Siddiqui, C. K. A., & Hoque, M. M. (2007). Road Traffic Accident: a Leading Cause of the Global Burden of Public Health Injuries and Fatalities. *International Conference on Mechanical Engineering*, *2007*(December).

Babu, M., Theerthala, R. R., Singh, A. K., Baladhurgesh, B. P., Gopalakrishnan, B., Krishna, K. M., & Medasani, S. (2019). Model predictive control for autonomous driving considering actuator dynamics. *Proceedings of the American Control Conference*, *2019-July*. doi: 10.23919/acc.2019.8814940

Bae, I., Moon, J., & Seo, J. (2019). Toward a comfortable driving experience for a self-driving shuttle bus. *Electronics (Switzerland)*, *8*(9). doi: 10.3390/electronics8090943

Balaban, A. T. (1985). Applications of Graph Theory in Chemistry. *Journal of Chemical Information and Computer Sciences*. doi: 10.1021/ci00047a033

Barnes, J. A. (1969). Graph theory and social networks: A technical comment on connectedness and connectivity. *Sociology*. doi: 10.1177/003803856900300205

Bokare, P. S., & Maurya, A. K. (2017). Acceleration-Deceleration Behaviour of Various Vehicle Types. *Transportation Research Procedia*, *25*. doi: 10.1016/j.trpro.2017.05.486

Chen, D., Lin, Y., Li, W., Li, P., Zhou, J., & Sun, X. (2019). *Measuring and Relieving the Over-smoothing Problem for Graph Neural Networks from the Topological View*.

Chen, Jiajia, Zhao, P., Mei, T., & Liang, H. (2013). Lane change path planning based on piecewise bezier curve for autonomous vehicle. *Proceedings of 2013 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2013*. doi: 10.1109/ICVES.2013.6619595

Chen, Jianyu, Yuan, B., & Tomizuka, M. (2019). Model-free Deep Reinforcement Learning for Urban Autonomous Driving. *2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019*. doi: 10.1109/ITSC.2019.8917306

Chen, S., Dong, J., Ha, P., Li, Y., & Labi, S. (2021). Graph neural network and reinforcement learning for multi-agent cooperative control of connected autonomous vehicles. *Computer-Aided Civil and Infrastructure Engineering*, *36*(7). doi: 10.1111/mice.12702

Derrible, S., & Kennedy, C. (2011). Applications of graph theory and network science to transit network design. *Transport Reviews*. doi: 10.1080/01441647.2010.543709

di Cairano, S., & Bemporad, A. (2010). Model predictive control tuning by controller matching. *IEEE Transactions on Automatic Control*, *55*(1). doi: 10.1109/TAC.2009.2033838

Dong, J., Chen, S., Joun Ha, P. Y., Li, Y., & Labi, S. (2020). A DRL-based multiagent cooperative control framework for CAV networks: A graphic convolution Q network. In arXiv.

Dong, J., Chen, S., Li, Y., Du, R., Steinfeld, A., & Labi, S. (2021). Space-weighted information fusion using deep reinforcement learning: The context of tactical control of lane-changing autonomous vehicles and connectivity range assessment. *Transportation Research Part C: Emerging Technologies*, *128*. doi: 10.1016/j.trc.2021.103192

Dong, J., Chen, S., Li, Y., Ha, P. Y. J., Du, R., Steinfeld, A., & Labi, S. (2020). Spatio-weighted information fusion and DRL-based control for connected autonomous vehicles. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*. doi: 10.1109/ITSC45102.2020.9294550

Elliott, D., Keen, W., & Miao, L. (2019). Recent advances in connected and automated vehicles. In Journal of Traffic and Transportation Engineering (English Edition). doi: 10.1016/j.jtte.2018.09.005

Fan, W., Ma, Y., Li, Q., He, Y., Zhao, E., Tang, J., & Yin, D. (2019). Graph neural networks for social recommendation. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, 417–426. doi: 10.1145/3308558.3313488

FHWA. (2015). Estimated Benefits of Connected Vehicle Applications: Dynamic Mobility Applications , AERIS , V2I Safety, and Road Weather Management Applications. In Tech. Rep. Nr. FHWA-JPO-15-255, Washington, DC.

FHWA. (2018). *FHWA National Dialogue on Highway Automation. Federal Highway Administration*.

FHWA. (2019). *Evaluation Methods and Techniques: Advanced Transportation and Congestion Management Technologies Deployment Program, Tech. Rep. Nr. FHWA-HOP-19-053, prepared by the Volpe National Transportation Syst; Washington D.C.*

Fout, A., Byrd, J., Shariat, B., & Ben-Hur, A. (2017). Protein interface prediction using graph convolutional networks. *Advances in Neural Information Processing Systems*.

Gupta, J. K., Egorov, M., & Kochenderfer, M. (2017). Cooperative Multi-agent Control Using Deep Reinforcement Learning. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi: 10.1007/978-3-319-71682-4_5

Ha, P., Chen, S., Du, R., Dong, J., Li, Y., & Labi, S. (2020). Vehicle Connectivity and Automation: A Sibling Relationship. In Frontiers in Built Environment (Vol. 6). doi: 10.3389/fbuil.2020.590036

Hobert, L., Festag, A., Llatser, I., Altomare, L., Visintainer, F., & Kovacs, A. (2016). Enhancements of V2X communication in support of cooperative autonomous driving. *Infocommunications Journal*.

Huegle, M., Kalweit, G., Mirchevska, B., Werling, M., & Boedecker, J. (2019). *Dynamic Input for Deep Reinforcement Learning in Autonomous Driving*.

Ji, J., Khajepour, A., Melek, W. W., & Huang, Y. (2017). Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints. *IEEE Transactions on Vehicular Technology*, *66*(2). doi: 10.1109/TVT.2016.2555853

Jiang, J., Dun, C., Huang, T., & Lu, Z. (2020). Graph Convolutional Reinforcement Learning. *ArXiv Abs/1810.09202 (2018)*.

Kalra, N., & Paddock, S. M. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice*, *94*. doi: 10.1016/j.tra.2016.09.010

Kim, J., Emersic, Z., & Han, D. S. (2019). Vehicle Path Prediction based on Radar and Vision Sensor Fusion for Safe Lane Changing. *1st International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2019*. doi: 10.1109/ICAIIC.2019.8669081

Kipf, T. N., & Welling, M. (2019). Semi-supervised classification with graph convolutional networks. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*.

Koopman, P., & Wagner, M. (2017). Autonomous Vehicle Safety: An Interdisciplinary Challenge. *IEEE Intelligent Transportation Systems Magazine*, *9*(1). doi: 10.1109/MITS.2016.2583491

Li, Y., Chen, S., Joun Ha, P. Y., Dong, J., Steinfeld, A., & Labi, S. (2020). Leveraging vehicle connectivity and autonomy to stabilize flow in mixed traffic conditions: Accounting for human-driven vehicle driver behavioral heterogeneity and perception-reaction time delay. In arXiv.

Li, Y., S. Chen, P. Ha, J. Dong, A. Steinfeld, and S. L. (2020). Leveraging Vehicle Connectivity and Autonomy to Stabilize Flow in Mixed Traffic Conditions: Accounting for Human-Driven Vehicle Driver Behavioral Heterogeneity and Perception-Reaction Time Delay. *Under Review:Transportation Research Part C: Emerging Technologies*.

Mousavi, S. S., Schukat, M., & Howley, E. (2018). Deep Reinforcement Learning: An Overview. In Lecture Notes in Networks and Systems. doi: 10.1007/978-3-319-56991-8_32

Noy, I. Y., Shinar, D., & Horrey, W. J. (2018). Automated driving: Safety blind spots. In Safety Science (Vol. 102). doi: 10.1016/j.ssci.2017.07.018

Qi, X., Luo, Y., Wu, G., Boriboonsomsin, K., & Barth, M. (2019). Deep reinforcement learning enabled self-learning control for energy efficient driving. *Transportation Research Part C: Emerging Technologies*. doi: 10.1016/j.trc.2018.12.018

Qiu, J., Tang, J., Ma, H., Dong, Y., Wang, K., & Tang, J. (2018). DeepInf: Social influence prediction with deep learning. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. doi: 10.1145/3219819.3220077

Rahman, M. S., Abdel-Aty, M., Lee, J., & Rahman, M. H. (2019). Safety benefits of arterials' crash risk under connected and automated vehicles. *Transportation Research Part C: Emerging Technologies*, *100*. doi: 10.1016/j.trc.2019.01.029

Shen, C., Guo, H., Liu, F., & Chen, H. (2017). MPC-based path tracking controller design for autonomous ground vehicles. *Chinese Control Conference, CCC*. doi: 10.23919/ChiCC.2017.8028887

Simon, D., & Lofberg, J. (2016). Stability analysis of model predictive controllers using Mixed Integer Linear Programming. *2016 IEEE 55th Conference on Decision and Control, CDC 2016*. doi: 10.1109/CDC.2016.7799391

Sinha, K. C., & Labi, S. (2007a). Transportation Decision Making: Principles of Project Evaluation and Programming. In Transportation Decision Making: Principles of Project Evaluation and Programming. doi: 10.1002/9780470168073

Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double Q-Learning. *30th AAAI Conference on Artificial Intelligence, AAAI 2016*.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, M. R. (2016). Playing Atari with Deep Reinforcement Learning. In IJCAI International Joint Conference on Artificial Intelligence. doi: 10.1038/nature14236

Wang, H., Huang, Y., Khajepour, A., Zhang, Y., Rasekhipour, Y., & Cao, D. (2019). Crash Mitigation in Motion Planning for Autonomous Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *20*(9). doi: 10.1109/TITS.2018.2873921

Wang, P., Chan, C. Y., & De La Fortelle, A. (2018). A Reinforcement Learning Based Approach for Automated Lane Change Maneuvers. *IEEE Intelligent Vehicles Symposium, Proceedings*. doi: 10.1109/IVS.2018.8500556

Werling, M., & Liccardo, D. (2012). Automatic collision avoidance using model-predictive online optimization. *Proceedings of the IEEE Conference on Decision and Control*. doi: 10.1109/CDC.2012.6426612

World Bank. (2005). A Framework for the Economic Evaluation of Transport Projects. In Transport Notes.

Xu, C., Ding, Z., Wang, C., & Li, Z. (2019). Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes. *Journal of Safety Research*, *71*. doi: 10.1016/j.jsr.2019.09.001

Yang, D., Zheng, S., Wen, C., Jin, P. J., & Ran, B. (2018). A dynamic lane-changing trajectory planning model for automated vehicles. *Transportation Research Part C: Emerging Technologies*, *95*. doi: 10.1016/j.trc.2018.06.007

Ye, L., & Yamamoto, T. (2019). Evaluating the impact of connected and autonomous vehicles on traffic safety. *Physica A: Statistical Mechanics and Its Applications*, *526*. doi: 10.1016/j.physa.2019.04.245

Zhang, K., Yang, Z., & Başar, T. (2019). *Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms*. 1–72.

APPENDIX

CCAT Project: Development of AI-based and control-based systems for safe and efficient operations of connected and autonomous vehicles

Published Related Work

Paper 1: Dong, J., Chen, S., Li, Y., Du, R., Steinfeld, A., Labi, S. (2021). Space-weighted information fusion using deep reinforcement learning: The context of tactical control of lane-changing autonomous vehicles and connectivity range assessment. *Transportation Research Part C: Emerging Technologies*. *Vol. 128, July 2021, 103192*

Abstract
The connectivity aspect of connected autonomous vehicles (CAV) is beneficial because it facilitates dissemination of traffic-related information to vehicles through Vehicle-to-External (V2X) communication. Onboard sensing equipment including LiDAR and camera can reasonably characterize the traffic environment in the immediate locality of the CAV. However, their performance is limited by their sensor range (SR). On the other hand, longer-range information is helpful for characterizing imminent conditions downstream. By contemporaneously coalescing the short- and long-range information, the CAV can construct comprehensively its surrounding environment and thereby facilitate informed, safe, and effective movement planning in the short-term (local decisions including lane change) and long-term (route choice). Current literature provides useful information on CAV control approaches that use only local information sensed from the proximate traffic environment but relatively little guidance on how to fuse this information with that obtained from downstream sources and from different time stamps, and how to use the fused information to enhance CAV movements. In this paper, we describe a Deep Reinforcement Learning based approach that integrates the data collected through sensing and connectivity capabilities from other vehicles located in the proximity of the CAV and from those located further downstream, and we use the fused data to guide lane changing, a specific context of CAV operations. In addition, recognizing the importance of the connectivity range (CR) to the performance of not only the algorithm but also of the vehicle in the actual driving environment, the study carried out a case study. The case study demonstrates the application of the proposed algorithm and duly identifies the appropriate CR for each level of prevailing traffic density. It is expected that implementation of the algorithm in CAVs can enhance the safety and mobility associated with CAV driving operations. From a general perspective, its implementation can provide guidance to connectivity equipment manufacturers and CAV operators, regarding the default CR settings for CAVs or the recommended CR setting in a given traffic environment.

Paper 2: Chen, S., Dong, J., Ha, P., Li, Y., Labi, S. (2021). Graph neural network and reinforcement learning for multiagent cooperative control of connected autonomous vehicles. *Computer-Aided Civil and Infrastructure Engineering 36(7), 838–857.*

Abstract
A connected autonomous vehicle (CAV) network can be defined as a set of connected vehicles including CAVs that operate on a specific spatial scope that may be a road network, corridor, or segment. The spatial scope constitutes an environment where traffic information is shared and instructions are issued for controlling the CAVs movements. Within such a spatial scope, high-level cooperation among CAVs fostered by joint planning and control of their movements can greatly enhance the safety and mobility performance of their operations. Unfortunately, the highly combinatory and volatile nature of CAV networks due to the dynamic number of agents (vehicles) and the fast-growing joint action space associated with multi-agent driving tasks pose difficultly in achieving cooperative control. The problem is NP-hard and cannot be efficiently resolved using rule-based control techniques. Also, there is a great deal of information in the literature regarding sensing technologies and control logic in CAV operations but relatively little information on the integration of information from collaborative sensing and connectivity sources. Therefore, we present a novel deep reinforcement learning-based algorithm that combines graphic convolution neural network with deep Q-network to form an innovative graphic convolution Q network that serves as the information fusion module and decision processor. In this study, the spatial scope we consider for the CAV network is a multi-lane road corridor. We demonstrate the proposed control algorithm using the application context of freeway lane-changing at the approaches to an exit ramp. For purposes of comparison, the proposed model is evaluated vis-à-vis traditional rule-based and long short-term memory-based fusion models. The results suggest that the proposed model is capable of aggregating information received from sensing and connectivity sources and prescribing efficient operative lane-change decisions for multiple CAVs, in a manner that enhances safety and mobility. That way, the operational intentions of individual CAVs can be fulfilled even in partially observed and highly dynamic mixed traffic streams. The paper presents experimental evidence to demonstrate that the proposed algorithm can significantly enhance CAV operations. The proposed algorithm can be deployed at roadside units or cloud platforms or other centralized control facilities.

Paper 3: Du, R., Chen, S., Li, Y., Ha, P. Y. J., Dong, J., Anastasopoulos, P. C., & Labi, S. (2021, September). A cooperative crash avoidance framework for autonomous vehicle under collision-imminent situations in mixed traffic stream. *2021 IEEE International Intelligent Transportation Systems Conference Proceedings (ITSC)*, 1997-2002, IEEE.

Abstract
Autonomous vehicles (AVs) are expected to increase the safety of transportation systems because automation minimizes human error in driving tasks. It is likely that such benefits will be fully manifested only when AV market penetration reaches 100%. However, the transition from a system of human-driven vehicles (HDVs) dominant to AVs dominant is expected to be time consuming. Thus, the safety benefits of AVs will be curtailed by the human error persisting through the human-driven vehicles (HDVs) during mixed traffic flow comprised of both AVs and HDVs. Such heterogeneity causes unsafe traffic operations maneuvers due particularly to the errant nature of human driving, especially in high-velocity lane-change maneuver. In this study, two perspectives of human error under the mixed traffic environment are proposed: 1) human error from inside of the vehicles; 2) human error from outside. This paper focuses on the second perspective, in the context of aggressive lane-change HDV. By formulating a Model Predictive Control (MPC) and V2V based cooperative framework, the AVs in such situations will be able to avoid side-impact and rear-end collision with the aggressive HDV. The framework is tested under different traffic conditions in terms of the vehicle bumper-to-bumper distance and relative velocities. The crash avoidance success rate averages at 90%, even reaches 100% when the relative velocity was low.

Paper 4: Dong, J., Chen, S., Li, Y., Ha, P. Y. J., Du, R., Steinfeld, A., Labi, S. (2020). Spatio-weighted information fusion and DRL-based control for connected autonomous vehicles. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC) Proceedings*, IEEE.

Abstract:
While on-board sensing equipment of CAVs can reasonably characterize the surrounding traffic environment, their performance is limited by the range of the sensors. By integrating short- and long-range information, a CAV can comprehensively construct its surrounding environment, thereby allowing it to plan both short and long-term maneuvers. Coalescing local information and downstream information is critical for the CAV to make safe and effective driving decisions. While literature is replete with CAV control approaches that use information sensed from the local traffic environment, studies that fuse information from various temporal-spatial instances to facilitate CAV movements is limited. In this paper, we propose a Deep Reinforcement Learning (DRL) based approach that fuses information obtained (via sensing and connectivity) on the local downstream environment for CAV lane changing decisions. We adopt learning-based

techniques to provide an integrated solution that incorporates the information fusion and movement-decision processor. We also determine the optimal connectivity range for each operating traffic density. We anticipate that deployment of the proposed algorithm in a CAV will facilitate reliable proactive driving decisions and ultimately enhance the overall operational efficiency of CAVs in terms of safety and mobility.