International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:15, No:1, 2021

An Effective Decision-Making Strategy Based on Multi-Objective Optimization for Commercial Vehicles in Highway Scenarios

Weiming Hu, Xu Li, Xiaonan Li, Zhong Xu, Li Yuan, Xuan Dong

***Abstract*—**Maneuver decision-making plays a critical role in high-performance intelligent driving. This paper proposes a risk assessment-based decision-making network (RADMN) to address the problem of driving strategy for the commercial vehicle. RADMN integrates two networks, aiming at identifying the risk degree of collision and rollover and providing decisions to ensure the effectiveness and reliability of driving strategy. In the risk assessment module, risk degrees of the backward collision, forward collision and rollover are quantified for hazard recognition. In the decision module, a deep reinforcement learning based on multi-objective optimization (DRL-MOO) algorithm is designed, which comprehensively considers the risk degree and motion states of each traffic participant. To evaluate the performance of the proposed framework, Prescan/Simulink joint simulation was conducted in highway scenarios. Experimental results validate the effectiveness and reliability of the proposed RADMN. The output driving strategy can guarantee the safety and provide key technical support for the realization of autonomous driving of commercial vehicles.

***Keywords*—**Decision-making strategy, risk assessment, multi-objective optimization, commercial vehicle.

I. INTRODUCTION

As the main undertaker of road transportation, the safety status of commercial vehicles is the focus of attention at home and abroad. Due to the high center of mass, large outline size, and strong operation intensity, commercial vehicle traffic accidents easily result in large and catastrophic accidents, which seriously threaten the social public security. Relevant statistics confirm that more than 60% of road traffic accidents are derived from improper decision-making [11]. Among all traffic accidents of commercial vehicles, collision and rollover accident rank the first and second respectively. Before a traffic accident occurs, if the pre-processing time can be increased by 0.5s, the number of accidents will be reduced by about 30% to 60%. Therefore, research on accurate and effective driving decision-making methods for the commercial vehicle plays an important role in the enhancement of the operation safety and guarantees capabilities of commercial vehicles and improvement of road traffic safety.

W. Hu and Xu Li are with the School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: hwm_seu@163.com, lixu.mail@163.com).

X. Li is with the China Transport Telecommunications & Information Center, No.1 Anwai Waiguan, Chaoyang District, Beijing 100011, China.

Z. Xu and L. Yuan are with the Shenzhen Youwei Information and Technology Development Co., Ltd, Building F XingHe WORLD No.1 YaBao Road LongGang District, Shenzhen 518129, China.

X. Dong is with the Research Institute of Highway Ministry of Transport, No. 8 Xitucheng Road, Haidian District, Beijing 100011, China.

To deal with the aforementioned issues, researchers have proposed various solutions based on different principles. The existing driving decision-making strategies are mainly categorized into two groups, the rule-based method and the learning algorithm-based method. Among the rule-based methods, a behavior decision system based on Finite State Machine (FSM) was established to implement three modes of decision-making states [1]. In [2] and [3], the behavior decision methods based on Hierarchical State Machine (HSM) were studied. A decision-making framework that contained motion prediction and threat assessment was built for autonomous driving at road intersections [4]. Generally speaking, as a widely used behavioral decision model, the rule-based model has the advantages of simple construction and convenient implementation. However, it still has certain shortcomings in the depth of scene traversal and the accuracy of decision-making. Therefore, the application of the rule-based method is limited in handling complex driving conditions.

Among the learning algorithm-based methods, a decision model based on Partially Observable Markov Decision Processes (POMDP) was established in [5]. However, the targeted scenario in this study is relatively simple. Reference [6] proposed an end-to-end decision-making model based on deep reinforcement learning, and map the driving state to driving action continuously. Nevertheless, the influence of the traffic environment on vehicle behavior decisions is not taken into consideration.

Reference [7] studied the lane-keeping decision system based on deep Q network (DQN) and DDAC (Deep Deterministic Actor Critic) algorithm. The results confirmed that DQN can only output discrete actions. Reference [8] designed a smart car lateral control algorithm based on deep reinforcement learning, to keep the vehicle in the center of the lane by controlling the steering wheel angle. However, the presented studies do not fully consider the impact of surrounding traffic participants on decision-making, so that the decision-making results still have certain deficiencies in terms of accuracy and rationality.

Through the comparative analysis of the existing methods and models, it is concluded that the research object of the above decision algorithms is mainly passenger vehicles. Compared with passenger vehicles, commercial vehicles with large mass and high centroid have longer braking distance, poor side-roll stability, and are prone to side-roll when braking or turning to brake. Therefore, the existing decision-making algorithms cannot be directly applied to commercial vehicles. Overall, there are relatively few studies on driving decision-making for operational vehicles. Meanwhile, the impact of the backward collision is ignored. There are still large deficiencies in rationality, effectiveness, accuracy, etc. In particular, there is a vacancy in the research of effective and reliable decision-making of commercial vehicles.

In this paper, we propose a decision-making framework named risk assessment-based decision-making network (RADMN). The RADMN is composed of two networks, aiming at identifying the risk degree and construct a driving decision-making strategy. First, risk degrees of the backward collision, forward collision and rollover are quantified for hazard recognition. Then, the driving decision-making problem which aims at safety is modeled as a Markov decision process. A multi-objective optimized decision-making model in the highway scenario for the commercial vehicle is established. In turn, effective and reliable driving decision-making strategies under different driving conditions can be obtained. The main contributions of this work are as follows:

1) A framework named RADMN comprehensively considers the impact of forward collision, backward collision and rollover on the safety of the commercial vehicle. And it achieves effective and reliable driving decision-making for commercial vehicles in highway scenarios.
2) The proposed framework quantifies driving strategies such as deceleration and steering in the form of numerical values, further improving the effectiveness and reliability of driving decisions. At the same time, the output driving strategy can be adjusted according to different driving conditions.

II. Overview of Proposed Framework

The driving strategy of the proposed framework is realized by RADMN, which is mainly composed of two parts, i.e., risk assessment module and decision-making module, as shown in Fig. 1.

In the risk assessment module, according to multi-sensor information, the risk degree of collision and rollover is calculated in real-time, which is used to provide more state information for the decision-making module. In the decision module, a deep reinforcement learning algorithm optimized for forward collision avoidance, backward collision avoidance and anti-rollover is proposed, which comprehensively considers the potentiality of danger and motion states of surrounding traffic participants.

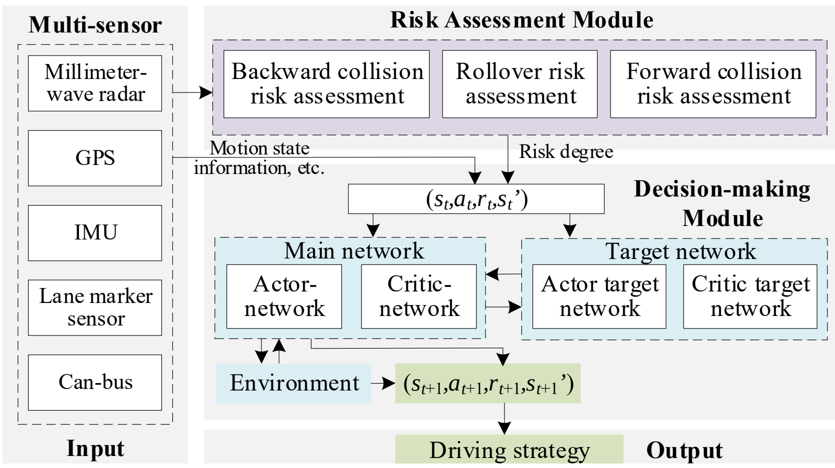

Fig. 1 Overall architecture of the proposed driving decision-making framework

III. Design Details for Driving Decision-Making Strategy

Driving decision-making is utilized to obtain an effective and reliable driving strategy, and then ensures the safety of commercial vehicles by avoiding collision, rollover, and other accidents. For the adaptability improvement to driving conditions, it is necessary to recognize the potentiality of danger. Meanwhile, a few techniques have been studied to promote the effectiveness and reliability of a decision-making strategy. Among them, risk assessment and decision-making algorithms are representative. In this paper, we present the fundamental characteristics of the key technologies, module design, and architecture of DRL-MOO.

A. The Risk Assessment Module

Accurate and real-time risk assessment is important to ensure the safety of the commercial vehicle. To this end, a risk assessment module containing collision risk assessment and rollover risk assessment is established.

1) Backward Collision Risk Assessment

Firstly, the time required for collision between the commercial vehicle and the backward vehicle is described as:

$$RTTC = -\frac{d_{sr}}{v_r} \tag{1}$$

where $RTTC$ denotes the reverse time to collision, d_r denotes the relative distance from the backward vehicle, v_R denotes the resultant velocity of the backward vehicle, v_r denotes the relative velocity which $v_r = v_C - v_R$.

Secondly, the risk degree of backward collision is calculated. According to the national transportation industry-standard named "performance requirements and test procedures for rear collision warning system for commercial vehicles", a backward collision warning is issued when *RTTC* is not less than 2.1 s and not more than 4.4 s, indicating that the backward collision warning system has passed the test. Based on this, the risk degree of backward collision is quantified as

$$\delta_{bc} = \begin{cases} \dfrac{4.4 - RTTC}{4.4 - 2.1} & 0 \le RTTC < 4.4 \\ 0 & RTTC \ge 4.4 \end{cases} \tag{2}$$

where δ_{bc} denotes the quantified value of backward collision risk, when $\delta_{bc} = 0$, it means no backward collision risk, when $0 \le \delta_{bc} \le 0.5$, there is backward collision risk, when $0.5 \le \delta_{bc} \le 1$, it means that the backward collision risk is high.

2) Forward Collision Risk Assessment

Firstly, the time required for collision between the commercial vehicle and the forward vehicle is calculated by

$$ETTC = -\frac{v_F - v_C - \sqrt{(v_F - v_C)^2 - 2d_{sf}\left[a_F - a_C\right]}}{a_F - a_C} \tag{3}$$

where *ETTC* denotes the enhanced time to collision, v_C and v_F denote the resultant velocity of the commercial vehicle and forward vehicle, a_C and a_F denote the acceleration of the commercial vehicle and forward vehicle. d_{sf} denotes the relative distance from the forward vehicle.

Secondly, the risk degree of forward collision is quantified as

$$\delta_{fc} = \begin{cases} \dfrac{\zeta_s - ETTC}{\zeta_s - \omega_s} & 0 \le ETTC < \zeta_s \\ 0 & ETTC \ge \zeta_s \end{cases} \tag{4}$$

where δ_{fc} denotes the risk degree of forward collision, ω_s and ζ_s denote the impact factors of a forward collision. According to the national transportation industry-standard named "performance requirements and test procedures for advanced emergency braking system for operating vehicles", take $\zeta_s = 3$ and $\omega_s = 0.8$.

3) Rollover Risk Assessment

Rollover risk is assessed through the lateral acceleration

$$\delta_r = \begin{cases} \sin\left(\dfrac{\pi}{2} \cdot \dfrac{|a_{lat}|}{a_{thr}}\right) & |a_{lat}| < a_{thr} \\ 1 & |a_{lat}| > a_{thr} \end{cases} \tag{5}$$

where a_{lat} and a_{thr} denote the lateral acceleration and the preset lateral acceleration threshold, respectively. δ_r denotes the quantified value of rollover risk, when $\delta_r = 1$, it means that rollover is about to happen.

B. The Decision Module

To reduce the traffic accidents caused by collisions or rollovers and improve the safety of commercial vehicles, a driving decision module is configured.

The complexity and uncertainty of traffic conditions and road state are important affecting factors for decision-making. Considering that deep reinforcement learning applies to fully mine and characterizing the high-dimensional features of traffic conditions, a driving decision-making model is established based on a deep reinforcement learning algorithm.

Generally, there are three decision-making strategies, namely the value-based method, policy search-based method, and actor-critic based method. The value-based method cannot deal with the issue of continuous output and cannot meet the requirements for continuous output driving strategy. Compared with the method based on policy search, the actor-critic method combines the value function estimation and policy search and possesses a faster update velocity. The deep deterministic policy gradient (DDPG) algorithm draws lessons from the DQN experience playback method and shows a better performance in the output of continuous action space. Therefore, after considering the influence of vehicle operation state, forward and backward obstacle types, and potentiality of danger on vehicle operation safety, the DDPG algorithm is adopted for establishing the decision-making model [9].

1) Defining the Basic Parameters of the Driving Decision Model

Considering that the future motion state of the commercial vehicle is affected by both the current motion state and current actor, the decision-making is modeled as a Markov decision process (MDP). The basic parameters of the model include the state-space S_t, action-space A_t, and the corresponding return value R_t of action-space.

a) State-Space

The collision risk and roll stability of a commercial vehicle are related to not only the vehicle's motion state but also the traffic conditions. Therefore, state-space (as illustrated in Table I) is defined based on vehicle motion state parameters and interactive information of traffic participants, which is described as

$$S_t = \left[p_x, p_y, v_{lon}, v_{lat}, a_{lon}, a_{lat}, \omega_{yaw}, \theta_h, \theta_{roll}, \theta_{swa}, \delta_{thr}, \delta_{brake}, d_l, d_{sr}, d_{sf}, \delta_{bc}, \delta_{fc}, \delta_r\right] \tag{6}$$

b) Action-Space

In the actual driving, the throttle and brake control quantities will not be applied to the commercial vehicle at the same time. Simultaneously, considering the effect of both lateral and longitudinal movements on vehicle running state, steering wheel and brake/throttle opening are used as control quantities to define the driving strategy, that is, the action-space represents

$$A_t = \left[\theta_{str\_out}, p_{out}\right] \tag{7}$$

where θ_{str_out} denotes the steering wheel angle control which is normalized in the range [-1,1]. p_{out} denotes the brake/throttle control which is normalized in the range [-1,1]. When $p_{out} > 0$, it represents that the throttle control is exerted to the commercial vehicle for acceleration. When $p_{out} < 0$, it represents that the brake pedal control is exerted to the commercial vehicle for deceleration.

TABLE I
DESCRIPTION OF THE STATE-SPACE S_t

Symbol	Unit	Description
p_x	/	Local lateral position of the commercial vehicle
p_y	/	Local longitudinal position of the commercial vehicle
v_{lon}	m/s	Longitudinal speed of the commercial vehicle
v_{lat}	m/s	Lateral speed of the commercial vehicle
a_{lon}	m/s^2	Longitudinal acceleration of the commercial vehicle
a_{lat}	m/s^2	Lateral acceleration of the commercial vehicle
ω_{yaw}	rad/s	Yaw velocity of the commercial vehicle
θ_h	°	Heading angle of the commercial vehicle
θ_{roll}	°	Roll angle of the commercial vehicle
θ_{swa}	°	Steering wheel angle of the commercial vehicle
δ_{thr}	/	Throttle opening of the commercial vehicle
δ_{brake}	/	Brake pedal opening of the commercial vehicle
d_l	m	Distance between vehicle and lane centerline
d_{sr}	m	Clearance from the backward vehicle
d_{sf}	m	Clearance from the forward vehicle
δ_{bc}	/	Quantitative value of backward collision risk
δ_{fc}	/	Quantitative value of forward collision risk
δ_r	/	Quantitative value of rollover risk

c) Reward Function

To evaluate the quality of output decision-making, an evaluation is materialized and quantified by establishing a reward function. Since driving decision-making is a multi-objective optimization problem involving safety, comfort and other objectives, the reward function is designed as

$$R_t = r_1 + r_2 + r_3 + r_4 \tag{8}$$

where R_t denotes the total reward function. r_1, r_2, r_3, and r_4 denote the safety distance reward function, the anti-rollover reward function, the comfort reward function, and the penalty function.

First, to prevent the commercial vehicle from colliding, a safety distance reward function is designed as

$$r_1 = \begin{cases} \omega_f\left(d_{sf} - d_0\right) + \omega_r\left(d_{sr} - d_0\right) & d_{sf} \geq d_0, d_{sr} \geq d_0 \\ \frac{1}{2}\left[\omega_f\left(d_{sf} - d_0\right) + \omega_r\left(d_{sr} - d_0\right)\right] & \frac{1}{2}d_0 \leq d_{sf} \leq d_0, \frac{1}{2}d_0 \leq d_{sr} \leq d_0 \\ 0 & else \end{cases} \tag{9}$$

where d_0 denotes the safety distance threshold, ω_f and ω_r denote the safe distance weighting factors, respectively.

Secondly, to prevent the vehicle from rollover, an anti-rollover reward function is designed as

$$r_2 = \sin\left(\frac{\pi}{2} \cdot \frac{|a_{lat}|}{a_{thr}}\right) + \sin\left(\frac{\pi}{2} \cdot \frac{|\theta_{roll}|}{\theta_{thr}}\right) \tag{10}$$

where a_{thr} and θ_{thr} denote the preset lateral acceleration threshold and roll angle threshold, respectively.

Thirdly, to ensure a better driving comfort, the excessive impact should be avoided. A comfort reward function is designed as

$$r_3 = \omega_j \left| a_{lon}(t+1) - a_{lon}(t) \right| \tag{11}$$

where ω_j denotes the comfort weight factor.

Finally, a penalty function is designed as

$$r_4 = \begin{cases} -100, \text{ collision occured} \\ -100, \text{ rollover occured} \\ 0, \text{ no collision or rollover occured} \end{cases} \tag{12}$$

2) Designing a Network Architecture for Decision-Making Model

An actor-critic framework is implemented to construct the decision-making model, including actor-network and critic-network. The actor-network takes the state-space information as input and outputs the action-space, which is the control value of the brake pedal opening, throttle opening and steering wheel angle. The critic-network takes state-space information and action decisions as inputs to output the current value of state-action (as illustrated in Fig. 2).

For actor-network, a hierarchical structure is built to extract features from various information in state-space. Firstly, the network encodes the motion state information and driver control information by using a few fully connected/ReLU layers, respectively. To map the above variables to fixed-length feature vectors, max- pooling layers are used behind fully connected layers. In the meanwhile, the network encodes the risk degree which is exported from the risk assessment module by using a fully connected layer. Then, feature vectors h_1, h_2, and h_3 are concatenated and mapped to a few fully connected layers to output actions.

For critic-network, the result from encoded state-space is mapped to a fully connected/ReLU layer and actions from actor-network are mapped to a fully connected layer. Then they are concatenated and mapped to a few fully connected layers to output the Q-value.

IV. Experimental Results

To verify the feasibility of the proposed framework, validation experiments were conducted based on the Prescan driving simulation platform. The simulation environment is imported from a real scene through Open Street Map, which is a part of Nanjing Airport Highway, China. The training scene and validation scene are shown in Fig. 3.

Experiments were conducted on a commercial vehicle. The commercial vehicle was equipped with a high precision differential GPS with centimeter-level position accuracy sampled at 50 Hz, an IMU sampled at 100 Hz, a lane marker sensor sampled at 20 Hz, and two millimeter-wave radars with 100 Hz update rate. The mounting positions of the sensors are shown in Fig. 4. All experiments were performed on a computer equipped with 64 GB RAM, Intel Core i7-7700k CPU with 32 cores, and a single GeForce GTX 2080 GPU.

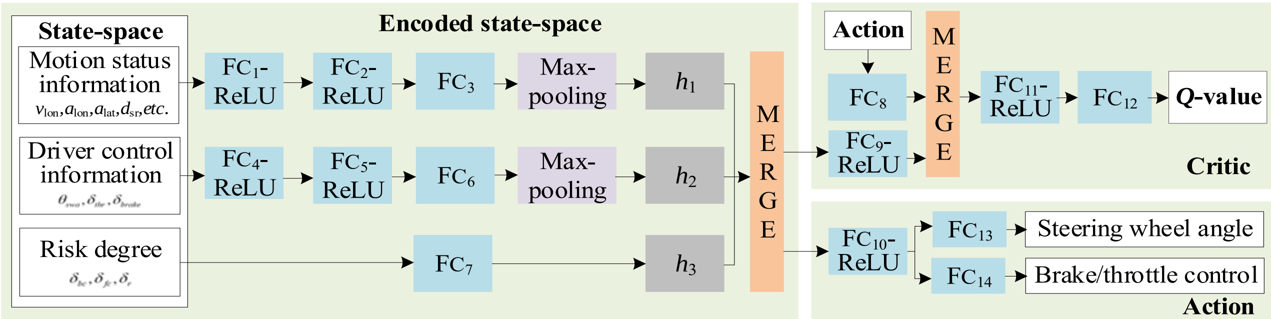

Fig. 2 Framework of the proposed actor-critic network

Fig. 3 Simulation environment of the highway scenario

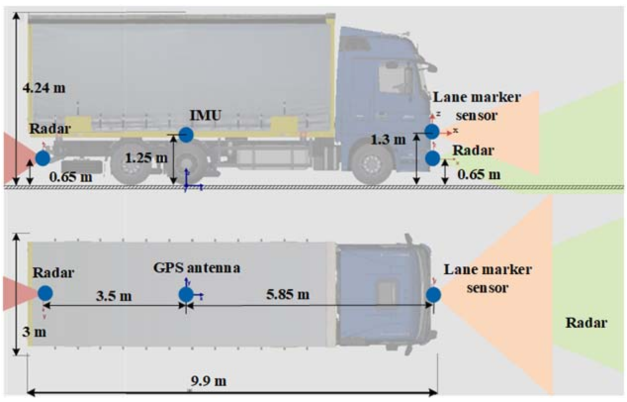

Fig. 4 Installation position of sensors on the commercial vehicle

A. Training Setup

For the part of feature extraction, the numbers of neurons in fully connected layer FC_1, FC_2, FC_3, FC_4, FC_5, FC_6, and FC_7 are 48, 48, 24, 48, 48, 24, 48 respectively. For the actor-network, the numbers of neurons in fully connected layer FC_8, FC_9, FC_{11}, FC_{12} are 48, 48, 24, and 24. The activation function of each layer is a linear rectification unit (ReLU). Adam optimizer [10] is utilized and the learning rate is set to 0.0001. The discount factor is set to 0.95, and the smooth factor is set to 0.0015.

For the critic network, the number of neurons in the layers FC_{10}, FC_{14} is 48. The activation function of each layer is ReLU. Critic network's learning parameter is the same as actor-network but for learning rate 0.001. The training batch size is set to 32. The maximum number of sequence states stored in the experience pool is 100000.

To improve the convergence speed of action-network and critic-network, the termination condition is set. In the case of retrograde, collision, rollover, and leaving the road, the current episode will be terminated and a new episode will be started for training. Besides, if the reward does not increase within 100 steps, the new round will be restarted.

B. Performance Evaluation

In Fig. 5, the distribution of reward after training the network for 950 episodes (208713 of total steps) is illustrated, in which the red line represents the variation of average rewards and the blue line represents the variation of episode reward. During the training process, the *Q*-value (as illustrated in the yellow line in Fig. 5) increases gradually and tends to be gentle. In the beginning, the network explores randomly with a minor reward of about 1000. Collision or rollover of the vehicle results in a reward of -100 in each step. With the training going on, the reward gradually increases. From the 120th episode, reward concentrates around 2500, and the distribution shifts to high reward, but the agent still gets out of track sometimes. From the 320th episode, the change of reward gradually flattens and the vehicle rarely collides or rolls over, which means the network is converged.

In Figs. 6 (a) and (b), control quantities of action decisions in the last training episode are shown. From the figure, we can see that the change of the curve is relatively gentle. The variation of control quantities (steering wheel angle, brake pedal, and throttle) at each moment is not changing dramatically. Moreover, there is no sudden steering, emergency braking, and other actions that affect the safety of the vehicle.

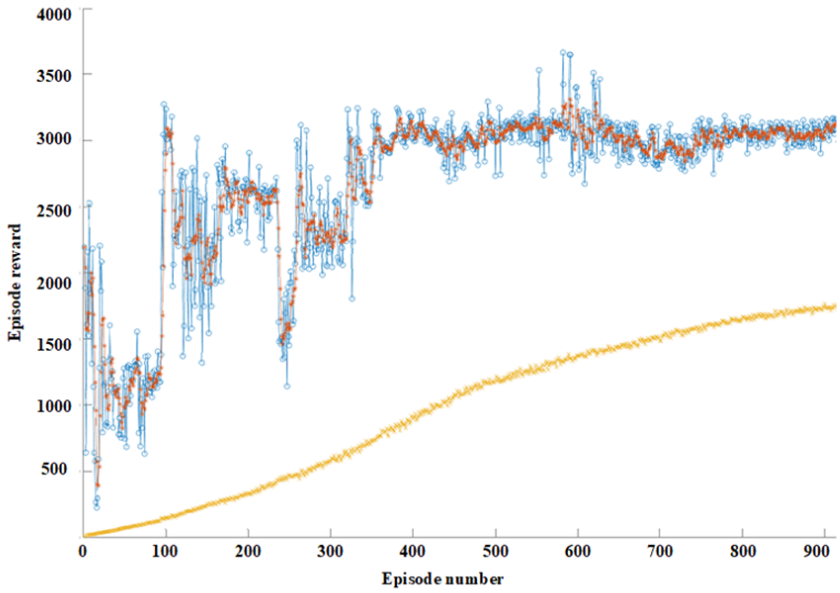

Fig. 5 Episode reward for driving decision-making

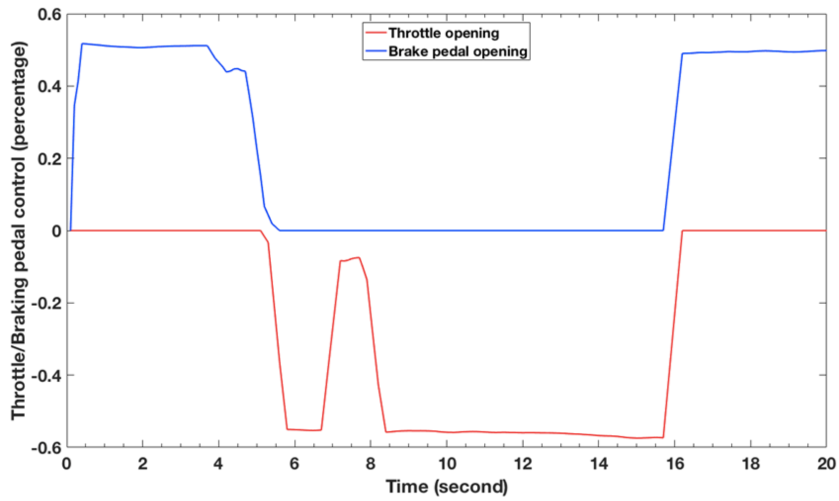

Fig. 6 (a) Control quantities of throttle/braking pedal opening

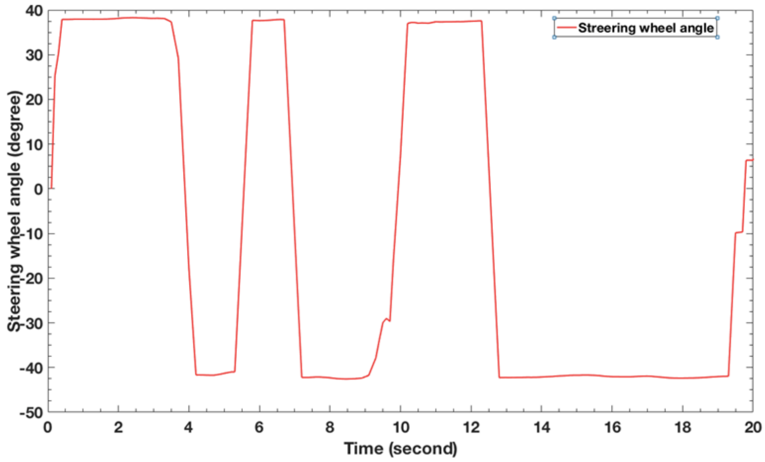

Fig. 6 (b) Control quantities of steering wheel angle

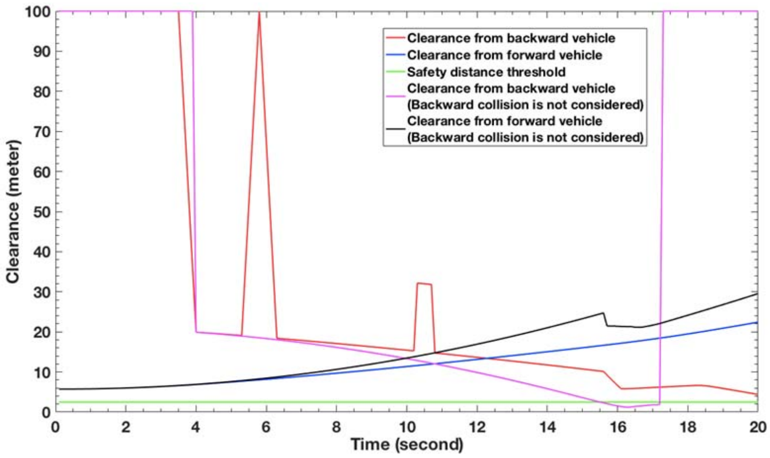

Fig. 7 (a) Clearance results in scenario 1

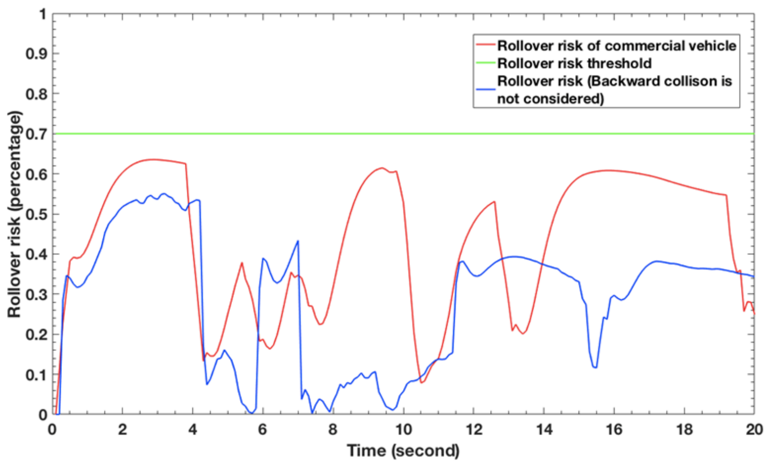

Fig. 7 (b) Rollover risk results in scenario 1

For intuitive evaluation of the network learning, two scenarios are selected in which the speed of surrounding vehicles is changing and constant randomly. Firstly, the simulation results for the surrounding vehicle driving at a dynamic speed are indicated. The commercial vehicle can follow the forward vehicle and keep a safe distance. In the running process, the clearances from surrounding vehicles are shown in Fig. 7 (a). The curve of rollover risk is illustrated in

Fig. 7 (b). It is worth noting that clearance equals 100 means no vehicle in the direction.

As can be seen from Figs. 7 (a) and (b), for the case without considering the backward collision, the clearance from backward vehicle, the clearance cannot always be greater than the safety distance threshold. The safety of the commercial vehicle cannot be guaranteed. For the framework proposed in this paper, the variables which include clearance from the backward vehicle and forward vehicle are always higher than the threshold. Meanwhile, the rollover risk is always less than the threshold. The output driving decision-making is confirmed to be able to avoid collision and rollover accidents.

Secondly, the simulation results for the surrounding vehicle driving at a constant speed are validated. The results show that the commercial vehicle follows the forward vehicle and keeps a safe distance. In the running process of the commercial vehicle, the clearances from the forward vehicle and backward vehicle are shown in Fig. 8 (a). The rollover risk of the commercial vehicle is shown in Fig. 8 (b). As can be seen from the figure, when the backward anti-collision is not considered in reward function, the performance of driving strategy is poor. For the proposed framework, the clearance at each moment is higher than the threshold, and rollover risk is less than the threshold. The output driving strategy can ensure the safety of the commercial vehicle.

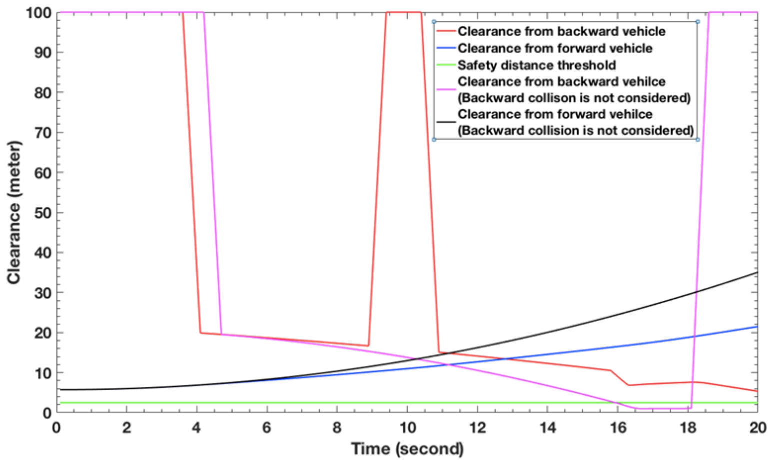

Fig. 8 (a) Clearance results in scenario 2

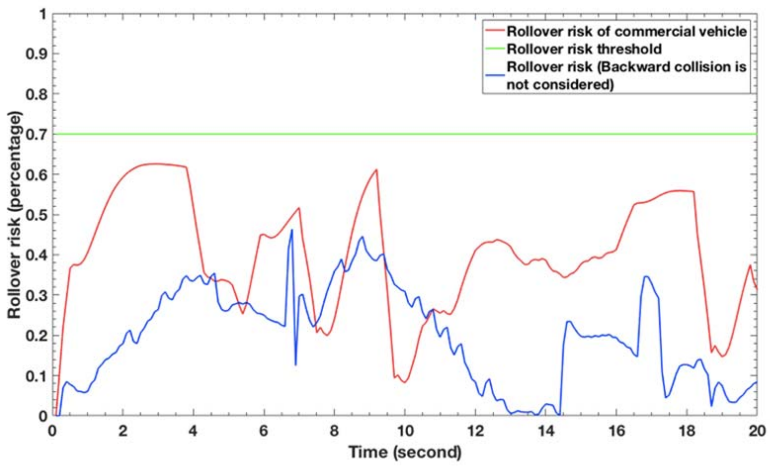

Fig. 8 (b) Rollover risk results in scenario 2

V. Conclusion

To enhance the safety of the commercial vehicle under the highway scenario, an effective decision-making method based on multi-objective optimization is proposed in this work. With the constructed risk assessment module, the risk of collision and rollover is significantly recognized. Reward function and network learning parameters are optimized to improve the performance of driving decision-making. The proposed framework comprehensively considers the impact of forward and backward obstacles on vehicle collision and rollover. The framework is theoretically and experimentally verified that it can output driving strategies under different driving conditions, and provide drivers with effective and reliable driving advice. Future work will focus on driving strategies under different road attachment conditions.

ACKNOWLEDGMENT

This work was supported in part by the Primary Research & Development Plan of Jiangsu Province BE2019106, in part by the Program for Special Talents in Six Major Fields of Jiangsu Province under Grant 2017 JXQC-003, in part by the National Natural Science Foundation of China under Grant 61973079, in part by the National Key Research and Development Program of China under Grant 2018YFB1600803, in part by the National Natural Science Foundation of China under Grant 41904024, and in part by the Fundamental Research Funds for the Central Universities under Grant 2242019K40037.

REFERENCES

[1] T. Gindele, D. Jagszent, B. Pitzer, R. Dillmann and Ieee, “Design of the planner of Team AnnieWAY's autonomous vehicle used in the DARPA Urban Challenge 2007,” in: 2008 IEEE Intelligent Vehicles Symposium, Vols 1-3, 2008, pp. 862-871.

[2] C.R. Baker and J.M. Dolan, “Traffic interaction in the urban challenge: Putting boss on its best behavior,” in: 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008.

[3] S. Ulbrich, M. Maurer and Ieee, “Probabilistic Online POMDP Decision Making for Lane Changes in Fully Automated Driving,” in: 2013 16th International IEEE Conference on Intelligent Transportation Systems -, 2013, pp. 2063-2070.

[4] S. Noh, “Decision-Making Framework for Autonomous Driving at Road Intersections: Safeguarding Against Collision, Overly Conservative Behavior, and Violation Vehicles,” IEEE Trans. Ind. Electron., vol. 66, pp. 3275-3286, Apr. 2019.

[5] W. Liu, S.W. Kim, S. Pendleton and M.H.A. Jr, “Situation-aware decision making for autonomous driving on urban road using online POMDP,” in: Intelligent Vehicles Symposium, 2015.

[6] Z. Huang, J. Zhang, R. Tian and Y. Zhang, “End-to-End Autonomous Driving Decision Based on Deep Reinforcement Learning,” in: 2019 5th International Conference on Control, Automation and Robotics (ICCAR), 2019.

[7] I.H. Lee and B.C. Luan, “Design of Autonomous Emergency Braking System Based on Impedance Control for 3-Car Driving Scenario,” in: Sae World Congress & Exhibition, 2016.

[8] D. Li, D. Zhao, Q. Zhang and Y. Chen, “Reinforcement Learning and Deep Learning Based Lateral Control for Autonomous Driving,” IEEE Computational Intelligence Magazine, vol. 14, pp. 83-98, 2019.

[9] Lillicrap, T., J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D, “Silver and Daan Wierstra. Continuous control with deep reinforcement learning,” International Conference on Learning Representations (ICLR), 2016.

[10] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980, 2014.

[11] C.Z. Wu, L.Q. Peng, Z. Huang, M. Zhong and D.F. Chu, “A Method of Vehicle Motion Prediction and Collision Risk Assessment with a Simulated Vehicular Cyber Physical System,” Transportation Research Part C (Emerging Technologies), vol. 47(2), pp. 179-191, 2014.

W. Hu received the M.S. degree in mechanical engineering and automation from Northeast University, Shenyang, China, in 2017. He is currently pursuing the Ph.D. degree in instrument science and technology with Southeast University, Nanjing, China. His current research interests include prediction of driving maneuver, intelligent driving assessment, and intelligent transportation systems.

Xu Li (M’10) received the Ph.D. degree in instrument science and technology from Southeast University, Nanjing, China, in 2006. He is currently a Professor with the School of Instrument Science and Engineering, Southeast University. Since 2002, he has been involved in the research and development of multi-sensor integration navigation, environment perception, and control for automated vehicles. His current research interests include the collaborative perception and control of intelligent vehicle and infrastructure systems, information fusion, automated vehicles, and active safety.