Energy Proceedings
Vol 47, 2024
ISSN 2004-2965

Hierarchical Eco-Driving Based on Safety Off-line Reinforcement Learning for P2-P3 Hybrid Electric Truck

Yahui Zhao[1], Zhong Wang[1], Yahui Zhang [2*], Yunfeng Song [2]
1 School of Vehicle and Energy, Yanshan University, Qinhuangdao, 066004, China
2 School of Mechanical Engineering, Yanshan University, Qinhuangdao,066004, China
(Corresponding Author Email: zhangyahui@ysu.edu.cn (Yahui Zhang))

ABSTRACT

Along with the rapid evolution of intelligent transportation systems (ITS) and network technology, vehicles have access to richer traffic data, paving the way for more efficient driving controls now. A novel hierarchical eco-driving strategy which is tailored specifically for hybrid electric truck navigating complex multi-intersection scenarios is proposed. Initially, a simulation scene is designed to simulate realistic truck-following scenarios. Subsequently, an upper-layer truck-following strategy is devised utilizing the safe off-line deep deterministic policy gradient(SDDPG) algorithm. This strategy is fully use of insights from leading vehicles and traffic signal data. Specifically, logical judgement module considering safety constraints are integrated into training processing to minimize collision risks. In addition, safe reward function is set to direct the agent to learn the safer action. Moving to the lower layer, an energy management strategy is proposed using deep reinforcement learning (DRL) techniques. A unique reward shaping function is introduced to guide the learning process effectively. Ultimately, the proposed methodology demonstrates a remarkable fuel-saving rate of 97.46% compared to dynamic programming (DP) approach by simulation.

Keywords: hybrid electric truck, truck-following, SDDPG, energy management strategy

1. INTRODUCTION

The pursuit of carbon peak and carbon neutrality is pivotal national strategies, which result in significant pressure on the transportation sector. It is crucial to accelerate transportation to achieve carbon peak because it is benefit for fostering high-quality development and facilitating green transformation. Hybrid power systems offer essential solutions to address energy management challenges, particularly in flexibly distributing energy within multiple power sources to meet the demand of power. Consequently, the research on energy management strategy (EMS) remains to be pivotal in automotive development. Furthermore, energy-saving control for hybrid electric vehicles (HEVs) now extends beyond EMS benefit from rapid advancements in intelligent connected technologies. Now Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) are widely utilized to enhance efficiency[1].

Eco-driving utilizes the information extracted from the surroundings to reduce vehicle fuel consumption and emissions through reasonable vehicle speed control and power distribution methods while ensuring safety, so as to achieve the purpose of energy conservation and emission reduction. Eco-driving has become one of the popular research directions now. Some studies have taken the driving terrain information into consideration to design energy management strategies. The results show that it can improve the energy saving effect by considering the surrounding environment information [2-3]. Some scholars integrate traffic information into driving strategies to design a collaborative optimization strategy. The results demonstrate that the approach achieves over 90% accuracy based on the dynamic programming (DP) method, while ensuring the vehicle tracking performance[4]. In addition, the ecological driving problem can also be constructed as a multi-agent hierarchical optimization problem, and a large number of scholars have conducted research on this[5-7].

Although previous studies have been carried out, many studies only consider the energy-saving effect ignoring safety separately or design complex learning processing[8-9]. In fact, many studies have proven that security constraints have a great impact on agent learning[10-11]. Because of the above phenomenon，a new SDDPG algorithm is proposed including a security mechanism. Compared with the previous TTC learning

This is a paper for the 10th Applied Energy Symposium: Low Carbon Cities & Urban Energy Systems (CUE2024), May. 11-12, 2024, Shenzhen, China.

algorithm, this method directly allows the agent to learn the appropriate acceleration, so that it can respond more directly to the velocity change under safe conditions. In addition, the design of the reward function should not be static because of the different demand at each time. Therefore, in order to explain the learning situation of the agent better, a new reward function form is proposed.

The rest of this paper is organized as follows. Detailed powertrain model of hybrid light truck and traffic flow model are illustrated in Section 2. Then, the hierarchical control structure is proposed in Section 3. Section 4 presents the experiment results. The relevant discussions is illustrate in Section5. Finally, conclusions and further research directions are outlined in Section 6.

2. SYSTEM MODEL DESCRIPTION

2.1 Truck-following model

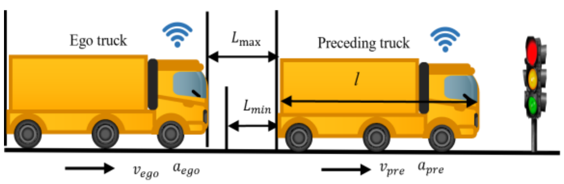

Fig. 1 Truck-following scene

As the above Fig. 1 shows, the two truck in the truck-following scene should meet the following equation:

$$\begin{cases} v_{ego} = a_{ego}t \\ x_{ego} = v_{ego}t + \frac{1}{2}a_{ego}t^2 \end{cases} \tag{1}$$

where a_{ego}, v_{ego} are the acceleration and the velocity of the ego car. And the x_{ego} is the travel distance. As shown in the figure above, L_{max} and L_{min} represent the maximum distance and the minimum distance between the two vehicles respectively.

2.2 Powertrain model

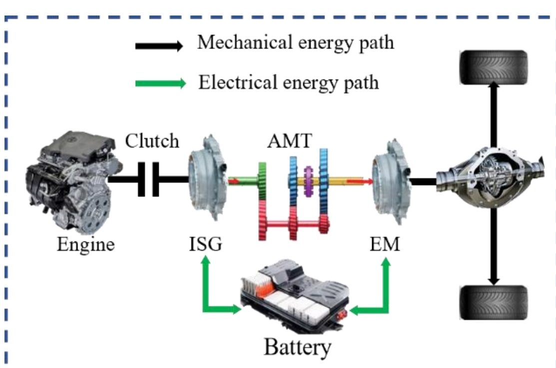

Fig. 2 Powertrain model

The research object of this paper is a coaxial hybrid electric truck equipped with an intelligent connected system. The hybrid system is shown in Fig. 2. The main components include engine, clutch, battery pack, two motor / generator, gearbox, wheel and so on. The main parameters are shown in Table 1.

Table 1

Vehicle parameters of truck

Parameters	Value
Weight	5000kg
Frontal area	$3.8m/s^2$
Air resistance coefficient	0.6
Tire radius	0.38m
Rolling resistance coefficient	0.012
Max power of EM motor	92kW
Max power of ISG motor	130kW
Max power of engine	65kW

The research simplifies the calculation process for the driving force demand of the truck by ignoring the performance of internal mechanical components and thermal energy. According to the longitudinal dynamics of the vehicle, the driving force be expressed as the following equation:

$$F_t = mgf\cos\theta + mg\sin\theta + \frac{1}{2}C_d\rho Av^2 + \delta m\frac{dv}{dt} \tag{2}$$

where F_t is the driving force demand, m is the mass, g is the gravity acceleration, f is the rolling resistance coefficient, ρ is the air density, $\theta$is the angle of slope, C_d is the drag coefficient, A is the front area, $v$is the velocity, $\delta$is the correction coefficient of rotating mass. The state of charge(SOC) is defined as the equation:

$$S\dot{O}C = \frac{U_{oc} - \sqrt{{U_{oc}}^2 - 4PR}}{2QR} \tag{3}$$

where U_{oc} is the voltage of battery, P is the power, R denotes the resistance, Qmeans the battery capacity.

3. HIERARCHICAL ECO-DRIVING FRAMEORK

The hierarchical framework of eco-driving consists of two parts, namely the upper-layer truck-following strategy passing through multi-intersection and the lower-layer energy management strategy. The overall framework is shown as Fig.3.

3.1 Truck-following strategy

In the upper layer, truck-following strategy is equipped with safe off-line deep deterministic policy gradient(SDDPG) agent including safe constrain in

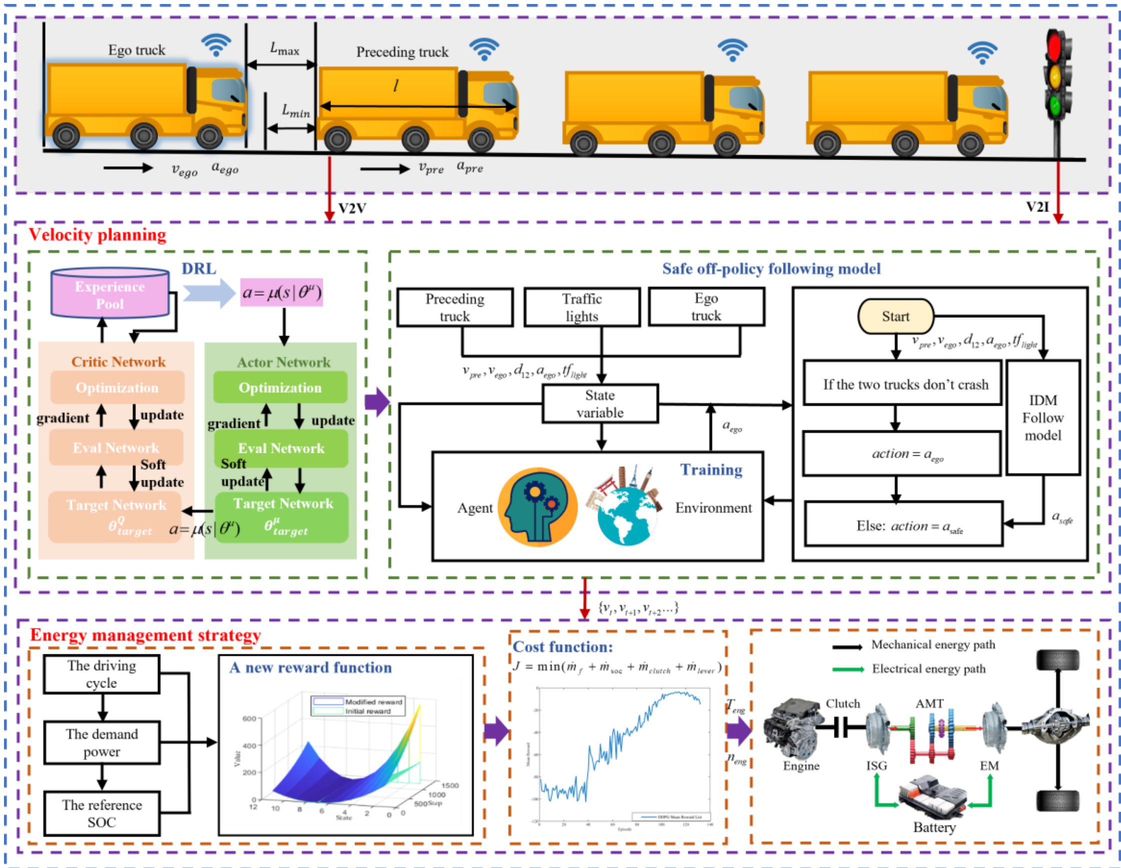

Fig. 3 The overall hierarchical eco-driving framework

learning processing. The state variables are shown in Table 2.

Table 2

The list of variables state

Vehicle state	Sensor state	Traffic light state
v_{ego}	v_{pre}	b_{TL}
a_{ego}	a_{pre}	d_{TL}
	d_{pre}	$t_{TL,gbegin}$
		$t_{TL,gend}$
		$v_{TL,max}$
		$v_{TL,min}$

where v_{ego} and a_{ego} are velocity and acceleration; v_{pre} , a_{pre} and d_{pre} are the state of preceding vehicle; b_{TL} reflects the phase of traffic lights; d_{TL} is the distance between the ego car and the target traffic light; $t_{TL,gbegin}$ and $t_{TL,gend}$ reflect the beginning time and ending time of target traffic light; $v_{TL,max}$ and $v_{TL,min}$ are the maximum and minimum velocity which can drive through traffic intersections safely.

As we know, many trained agents often learn actions that beyond the safe boundary, so in this section a logical judgement module with intelligent driver model(IDM) is designed as shown in Fig. 4.

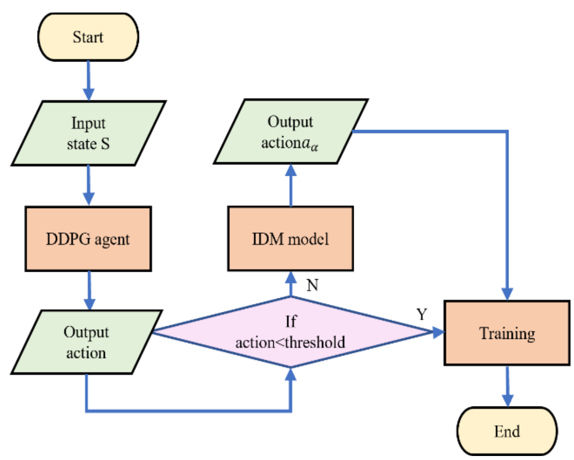

Fig. 4 The logical judgement module with IDM

Except for that, reward function related to safety rules is designed to direct the agent to learn safe acceleration. The reward function is set as :

$$r = w_{fuel}r_{fuel} + w_{tr}r_{tr} + w_{jer}r_{jer} + w_{safe}r_{safe} \quad (4)$$

where w_{fuel} , w_{tr} , w_{jer} , w_{safe} are the weights of different parameters. r_{fuel} , r_{tr} , r_{jer} are reward functions that consider fuel consumption, traffic information and comfort. , r_{safe} represents safe direction on agent respectively.

3.2 Energy management strategy

The state space is defined as five dimensions space including the demand power P_{dem}, the state of battery SOC_{now}, the velocity of the host truck v_{ego}, the current position s_{ego} and the state of the clutch.

In order to learn a less fuel consumption strategy, this paper introduces a leverage function into the framework of the reward function. At the same time, a modified SOC reward is designed considering the demand relationship between SOC and fuel is flexible. For example, the truck should priorly use fuel in the beginning. With the constant driving, we want to restore the SOC to the initial value, so that the battery energy can be fully utilized. The modified reward function is set as:

$$r = w_f r_f + w_{clu} r_{clu} + w_{soc} r_{soc}(\beta, i) + w_{lev} r_{lev} \quad (5)$$

where w_f, w_{clu}, w_{soc}, w_{lev} are the weights of different parameters. r_f, r_{clu}, r_{soc}, r_{lev} are reward functions that consider fuel consumption, stop-start of clutch, SOC and action choice. β is the factor between SOC maintain and change, and *i* represents the step of current episode.

4. SIMULATION RESULTS

A multi-intersection road with traffic lights is established as Fig.5 . In order to validate the effect of the reinforcement learning, the Krauss algorithm and the base energy management algorithm which is lack of r_{lev} and r_{soc} is fixed have been compared.

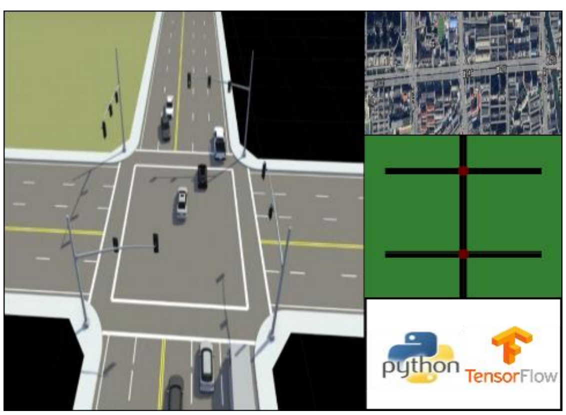

Fig. 5 The simulation scene

The training process of upper is shown in Fig. 6. The truck-following distance of the two cars is shown in Fig. 7. The acceleration and velocity compared with Krauss are shown in Fig.8. The truck-following effects obtained by the algorithms are shown in the Fig. 9.

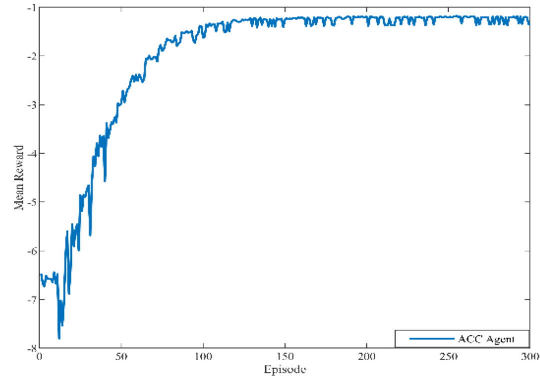

Fig. 6The training process of truck-following

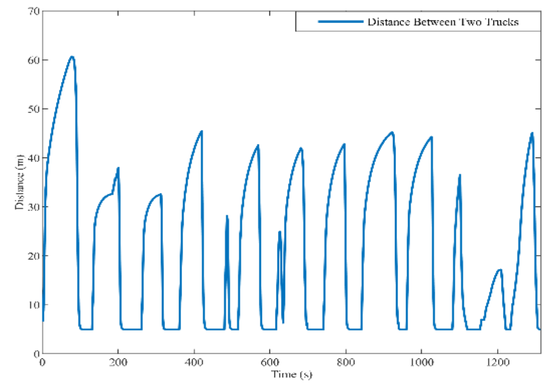

Fig.7 The Distance between two trucks

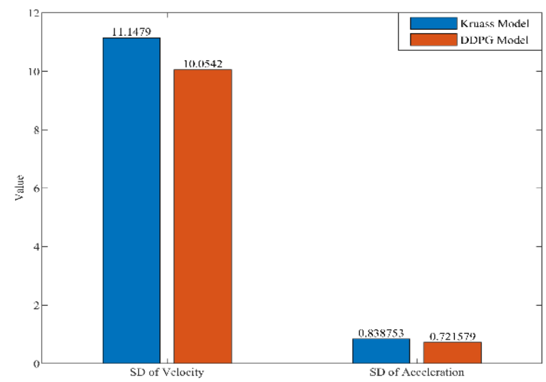

Fig. 8 Results comparison with Krauss

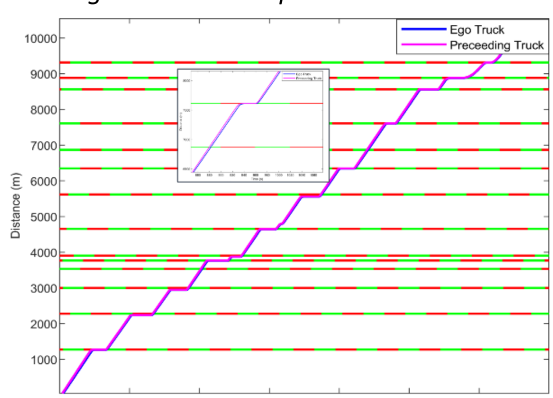

Fig. 9 The trip trajectory curves of two trucks

In addition, the training process of lower layer is expressed in the Fig. 10. The SOC curves of different strategy are shown in Fig.11.

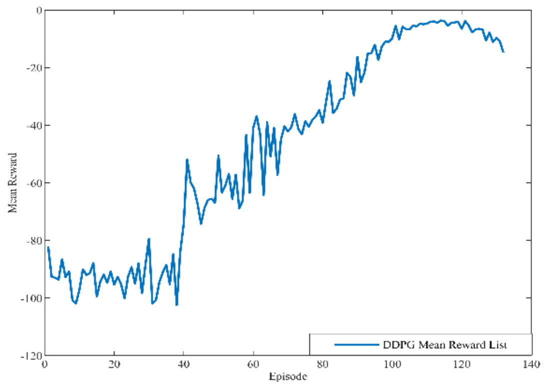

Fig. 10 The training process of lower layer

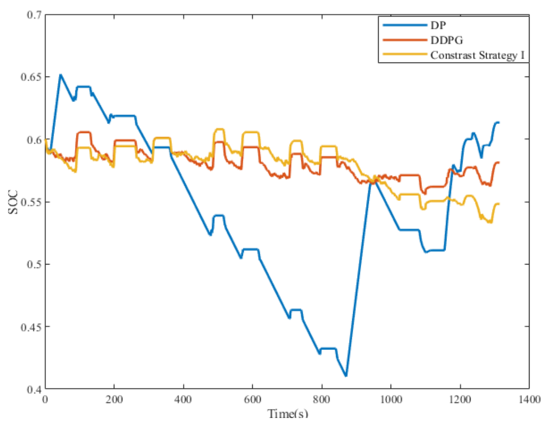

Fig. 11 The SOC curves of different strategies

The final fuel consumption of different methods is list in Table 3.

Table 3

The variables of state

Strategy	SOC final	Equivalent fuel consumption (g)	Percentage
DP	0.612	572.62	--
Contrast strategy I	0.548	601.08	95.03%
DDPG modified	0.581	587.76	97.46%

5. DISCUSSION

It can be seen from the Fig. 6 that the upper RL agent has learned a satisfactory strategy. Fig. 7 shows that the following distance between the two trucks is always greater than 0. It means that there will be no collision, which means that the strategy learned by the agent is very safe. It can be seen from the Fig. 8 that the RL strategy has more lower changes on acceleration compared with the Krauss, which means better comfort. Fig. 9 means that there is good truck-following strategy when driving through the intersection traffic lights. It can be seen from the Fig. 10 that the lower strategy agent has learned a satisfactory strategy. The battery SOC change curve is shown in the Fig. 11. It can be seen that the SOC terminal value of the proposed strategy is better than the baseline strategy, and the fuel efficiency can reach 97.46% of DP, which shows that the proposed strategy has good fuel economy.

6. CONCLUSIONS

This paper presents a hierarchical ecological driving strategy for hybrid electric trucks using SDDPG. Initially, a reinforcement learning agent is developed by the SDDPG algorithm to learn truck-following strategy at multi-intersection scene. At the same time, safety constraints and safe logic framework are introduced to enhance driving safety. Besides, it is helpful to ensure comfortable speed trajectories and appropriate following distances. Subsequently, agents equipped with energy management strategies are trained to optimize power allocation. A novel reward function design method is proposed to balance battery loss and fuel consumption effectively. Simulation results demonstrate that the hierarchical driving strategy demonstrates improved comfort compared with Krauss under the premise of safety. At the same time, , there is a significant improvement in SOC final value, which is better for the recycling of battery. Compared to the DP method, the proposed approach achieves 97.46% of the fuel consumption. Future work will focus on developing more complex scenarios and improving the universality of algorithm.

ACKNOWLEDGEMENT

This research is supported by National Natural Science Foundation of China (52302482), S&T Program of Hebei (236Z2202G), Hebei Natural Science Foundation (E2021203079), Science and Technology Foundation of Hebei Education Department (QN2022176), and Scientific Research Foundation of Hebei Province for the Returned Overseas Chinese Scholars (C20210323).

REFERENCE

[1] Shi, J., Xu, B., Zhou, X., & Hou, J. A cloud-based energy management strategy for hybrid electric city bus considering real-time passenger load prediction. J. Energy Storage 2022; 45, 103749.

[2] Li, K., Zhou, J., Jia, C., Yi, F., & Zhang, C. Energy sources durability energy management for fuel cell hybrid electric bus based on deep reinforcement learning considering future terrain information. Int J Hydrogen Energ 2024; 52, 821-833.

[3] Lin, X., Zhang, J., & Su, L. A trip distance adaptive real-time optimal energy management strategy for a plug-in hybrid vehicle integrated driving condition prediction. J. Energy Storage 2022; 52, 105055.
[4] Peng, J., Fan, Y., Yin, G., & Jiang, R. Collaborative optimization of energy management strategy and adaptive cruise control based on deep reinforcement learning. IEEE Trans. 2022; 9(1), 34-44.
[5] Liu, Y., Huang, B., Yang, Y., Lei, Z., Zhang, Y., & Chen, Z. Hierarchical speed planning and energy management for autonomous plug-in hybrid electric vehicle in vehicle-following environment. Energy 2022; 260, 125212.
[6] Wang, J., Du, C., Yan, F., Zhou, Q., & Xu, H. Hierarchical rewarding deep deterministic policy gradient strategy for energy management of hybrid electric vehicles. IEEE Trans. 2023.
[7] Xue, J., Jiao, X., Yu, D., & Zhang, Y. Predictive hierarchical eco-driving control involving speed planning and energy management for connected plug-in hybrid electric vehicles. Energy 2023; 283, 129058.
[8] Ganesh, A. H., & Xu, B. A review of reinforcement learning based energy management systems for electrified powertrains: Progress, challenge, and potential solution. RENEW SUST ENERG REV. 2022; 154, 111833.
[9] Hoel, C. J., Wolff, K., & Laine, L. (2018, November). Automated speed and lane change decision making using deep reinforcement learning. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 2148-2155).
[10] Ye, F., Cheng, X., Wang, P., Chan, C. Y., & Zhang, J. (2020, October). Automated lane change strategy using proximal policy optimization-based deep reinforcement learning. In 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 1746-1752).
[11] Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., & Amodei, D. Deep reinforcement learning from human preferences. Adv. Neural Inf. Process. Syst. 2017; 30.
[12] Xue, R., Zhao, Y., Wang, Z., Wang, Z., & Zhang, Y. (2023, October). Deep reinforcement learning based hierarchical eco-driving strategy for a series-parallel hybrid electric truck. In 2023 7th CAA International Conference on Vehicular Control and Intelligence (CVCI) (pp. 1-6).
[13] Treiber, M., Hennecke, A., & Helbing, D. Congested traffic states in empirical observations and microscopic simulations. Phys. Rev. E 2000; 62(2), 1805.