Intelligent Control Scheme for Traffic Signals at Multi-intersections based on Pedestrian Safety

Weiming Liang*

School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing, China

* Corresponding Author Email: b20032229@njupt.edu.cn

Abstract. Traffic congestion has become a more and more serious problem. Meanwhile, pedestrian safety problem is another issue under the circumstance of multiple modes of vehicles that needed to be dealt with. And these two problems can be effectively solved if the traffic signal is set reasonably. This paper proposes an intelligent control scheme for traffic signals at multi-intersections based on pedestrian safety, which is designed by using multi-agent deep reinforcement learning and self-attention mechanism. And Centralized training decentralized execution is also applied in this paper which can effectively downsize the action dimension to make the model learn much faster. Though pedestrian safety and vehicle efficiency are two contradictory factors, this paper dexterously designs the reward function to let the model learn the optimal timing scheme. And the model outperforms the fixed-timed scheme and deep Q-learning scheme in the simulation traffic environment, with fewer conflicts between pedestrians and vehicles, shorter average delay time and higher average speed of vehicles at the same time.

Keywords: Deep reinforcement learning, traffic signal, pedestrian safety.

1. Introduction

With the increasing pace of urbanization and the diversification of urban transportation modes, problems related to urban traffic congestion are becoming more and more serious. How to improve the operational efficiency of the road network by controlling intersection signals has been a research hotspot in recent years. Pedestrian is one of the most crucial parts among the transportation system, yet most researches focused on improving the efficiency of a particular type of vehicle, without considering potential issues of pedestrian safety. In crossings with signs or signals, 22% of pedestrian fatalities occurred, and more than 30% of those were attributed to illegal crossing behavior, according to statistics on traffic accidents [1]. Therefore, intelligent control of multi-intersection traffic signals that concerning about road network efficiency and pedestrian safety becomes rather important.

Traditional controlling strategies of traffic signals are mostly divided into three different categories. One is the fixed allocation time method, this type of methodology sets the duration of traffic signals artificially by summarizing the existing traffic patterns [2]. The second is adaptive methods, in which traffic signals are switched based on artificially set rules when specific conditions are met [3, 4]. Thirdly, optimization-based methods such as Max Pressure, which selects the signal phase with the highest traffic pressure every time [5]. The disadvantage of traditional methods is that their strategies cannot be dynamically adjusted when the traffic environment is complex.

In recent years, the application of artificial intelligence techniques in the field of traffic control effectively solves the problems faced by traditional methods, among which the multi-agent approach in deep reinforcement learning (DRL) is the most popular [6]. Multi-agent modelling enables cooperation between agents for strategy learning. The traditional modelling of multi-agent treats the whole system as a whole and its learning strategy is a mapping from global states to joint actions. However, when dealing with complex traffic issues, its state space and action space grow exponentially with the expansion of system scale, which makes effective learning impossible. Centralized Training, Distributed Execution (CTDE) learning model is capable solving the problem above, which uses global information for training and each agent only relies on local information when making decisions. It is effective in downscaling the size of action space, which means easier learning for strategies. Wu et al. designed a method based on MADDPG for traffic signals control [7,

8]. Zhang et al. proposed a Nash-Stackelberg hierarchical game model based on the importance of different intersections and game relationships in the road network, which solve the Nash equilibrium computational complexity problem caused by the multi-intersection game [9]. Wu et al. formulated a traffic organization scheme based on traffic lights, and implemented a trajectory reconstruction at the end of the execution of each act of the agents. The trajectory reconstruction is applied in changing the travelling path of vehicles without changing the Orientation-Destination (OD) pairs. And the final return of agents is calculated according to the scheme and reconstructed trajectory [10].

In the research field of safe pedestrian crossing, Zhuang et al. proposed a pedestrian manual activation control scheme, in which pedestrian manually press the activation button to change the traffic signal state when there is a need to cross the road. However, this pedestrian-vehicle separation control strategy actually badly affects the overall efficiency of the traffic signal control system [11]. Turky et al. proposed a genetic algorithm-based traffic signal control scheme under the condition of considering pedestrian factors. The effectiveness of candidate chromosomes is evaluated to regulate the signals by representing the waiting number of vehicles and pedestrians through a fitness function. However, in the case of unfair situations between the two, the priority of vehicles is set to be higher than that of pedestrians [12]. Liu et al. proposed a traffic signal control scheme based on distributed multi-agent reinforcement learning to minimize the waiting time of vehicles and pedestrians. However, it neglects the problem that pedestrians may not be able to fully pass through the intersection during the green light, which contributes to conflicts with vehicles [13]. Zhang et al. solved the above problem by introducing a dynamic all-red phase at the end of each green flash period and adaptively adjusting its duration according to the number of rule-violated pedestrians. However, this leads to a long duration of each dynamic all-red phase, which greatly affects the efficiency of vehicle traffic [14]. Zhang et al. compromised the amount of conflict with vehicles caused by pedestrian violations and the speed of vehicle to guide traffic signals to learn the behavior of pedestrians at intersections which contributed to the learning of optimal traffic light timing scheme. However, this study only considered the pedestrian and vehicle problems at a single intersection and lacked overall practicality [15].

In order to coordinate the efficiency of vehicles and pedestrian safety at multi-intersection circumstances, this paper proposes an intelligent control scheme for traffic signals at multi-intersections based on pedestrian safety. Through real-time observations of the traffic flow at multiple intersections, this paper comprehensively considers the contradiction between pedestrian violations and vehicle efficiency, and design the reward function in the DRL architecture, which can intelligently guide the learning to regulate the traffic signal cycle length, phase sequence, and phase duration, so as to achieve safe and efficient intelligent control of traffic signals at multiple intersections. At the same time, CTDE is used to reduce the dimensionality of the action space so that the strategies can be learned more easily, and attention mechanism is added to the deep network so that the neural network automatically pays attention to the important state components in order to enhance the network's perceptual ability.

2. Method

The well-known platform Simulation of Urban MObility (SUMO), which is an open source, highly portable, microscopic, and continuous traffic simulation software created to manage massive networks, is used in this article. It has a graphical user interface, allows numerous grid format inputs, and supports different road network designs. All experience data mentioned in this paper is generated by the help of this traffic simulation package. And this paper uses an Application Programming Interface (API) called TraCI provided by SUMO to interact with the algorithm in python.

2.1. Setting of Simulation Environment and Parameters

2.1.1 Intersection attribute setting

Each intersection is in a cross shape, divided into east-west and north-south. There are two-way 6 lanes in each direction, 3 in each way – left or straight, straight, right according to the driving direction from left to right. Besides, there is a sidewalk next to the right-turn lane. The speed limit is 30km/h on the driveway, 1m/s on the sidewalk and 1.2m/s on the pedestrian crossing. Total length of each direction is 400m, that is, 200m for each side, see Fig. 1. This paper doesn't take the circumstance that vehicles can turn right during the red light into consideration.

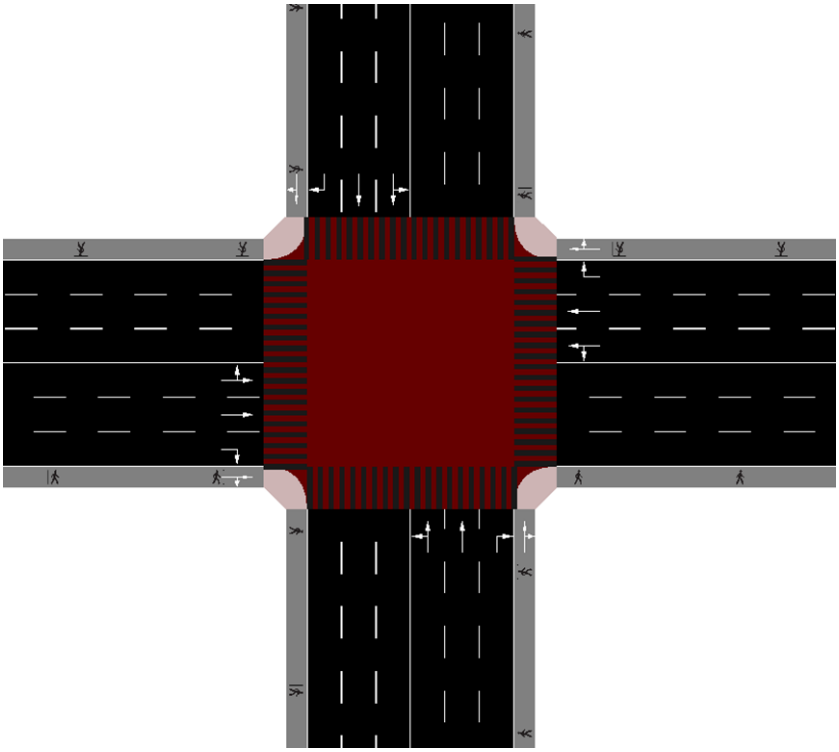


Fig. 1 Overview of a single intersection

Multi-intersection scenarios is composed by 4 intersections in a 2*2 shape, where the intersection is identical, that is, with the same initial setting, see Fig. 2.

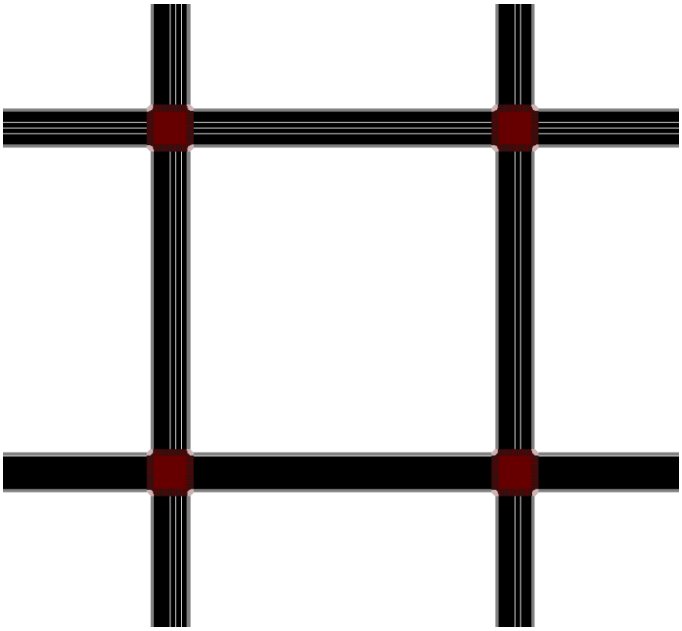


Fig. 2 Multi-intersection scenario

2.1.2 Initial attribute setting of traffic light

The attributes of traffic light in each intersection are set according to SUMO's default 3-direction setting, from innermost to outermost are left or straight, straight and right.

2.1.3 Setting of traffic flow

The arrival of vehicles and pedestrians follows the Poisson distribution, and the scenarios such as "morning rush hour", "evening rush hour" and "off-peak time" are effectively simulated according to the arrival probability of vehicles and pedestrians in different time periods, so as to reflect the real traffic flow in a day.

In reference of the setting of Zhang etc., this paper compresses the time of daily traffic flow to 3 hours(10800s) for simulation experience [15]. Traffic flow statistics are collected in 8 periods, which are 0~1350s, 1350~3150s, 3150~4050s, 4050~5400s, 5400~6300s, 6300~7650s, 7650~8550s and 8550~10800s. During the 3 hours, vehicles randomly enter the entrance of the intersection and select the lane in advance, and randomly select the target lane when exiting the intersection. Pedestrians randomly enter the entrance of the intersection and go straight to one of the four direction through the intersection.

2.2. Evaluation Metric

This study compares the effectiveness of several traffic signal control techniques using three distinct metrics, illustrated as followed: First is the average waiting time of vehicles. The waiting time is defined as the duration between a vehicle enter the entrance of an intersection and it successfully go through the waiting line. The second one is the amount of conflicts. It's the amount of pedestrians who are still walking on the crossing when the pedestrian light is turning red. It's a crucial factor of accidents. The last one is the average speed of all vehicles during the 3-hour period.

2.3. Intelligent Traffic Light Control Algorithm

This paper proposed an algorithm using multi-agent deep reinforcement learning with an attention layer between observation input and network's first hidden layer. With the help of attention mechanism, the model is able to distribute attention on different input components. These components can selectively emphasize informational features and suppress less useful features, resulting in a more accurate state input for the first hidden layer. At the same time, this paper proposed to use CTDE strategy for training, which is to say that each agent use local environment state and the neighboring states for training, and for execution they only use the information of the local environment.

2.3.1 State definition

State is that each agents receive a quantitative description of the local intersection environment at each time step. This paper defines a state input as:

$$s_{t,i} = \{delay_t[l], waitline_t[l], conflict_t[l]\}_{ji \in E, l \in L_{ji}}, \quad (1)$$

where $s_{t,i}$ is the state of intersection i at time step t. $delay_t[l]$ measures the average waiting time of vehicles on lane l at time step t, while $waitline_t[l]$ measures the total number of approaching vehicles along incoming lane l at time step t. $conflict_t[l]$ means the amount of conflict with vehicles caused by pedestrian violations, which is that the number of pedestrians who have not fully go through the crosswalk when the pedestrian traffic light is turning red, at incoming lane l at time step t.

2.3.2 Action definition

With reference of Prashanth and Bhantnagar's proposed method of traffic signals action setting, this paper defines each local action as a possible phase, and predefines a_t as a set of all feasible phases for each intersection, and each agent selects one of them to be implemented for a time step t [16, 17].

2.3.3 Reward definition

In order to effectively guarantee the safety of pedestrians, this paper sets that pedestrians passing through the intersection within the range of the green light is legitimate, while the presence of pedestrians at other stages on the crosswalk is regarded as a violation and as a conflict with the vehicles which are normally running. At this time, the number of pedestrians is called the amount of conflict between pedestrians and vehicles. Therefore, when evaluating whether the pedestrian is safe when crossing the intersection, this article sets the number of conflicts as a metric, each agent is meaning to reduce the number of conflicts by learning. When it turns to traffic efficiency evaluation, the metrics are the average speed of vehicles on each lane and the average waiting time of vehicles in each incoming lane.

To summarize, this paper pay full attention to the contradictions between pedestrian safety and vehicle traffic efficiency, alongside with experiences brought by different behaviors of pedestrian and vehicles, giving a reward definition as follows:

$$r_{t,i} = -\sum_{ji \in E, l \in L_{ji}} (conflict_{t+\Delta t}[l] + delay_{t+\Delta t}[l] + e^{-v}) \quad (2)$$

where v is the average speed after process of normalization.

2.3.4 Structure of deep neural network

As shown in Fig. 3, there are 4 state inputs—delay, waitline, conflict and neighbor policies, which will be put into a self-attention layer (this paper uses multi-head attention algorithm) for features extraction. Then there is a LSTM layer following in order to further mitigate the nonstationary caused by complex spatial-temporal data and the state dimension explosion, and it will also maintain or improve the model's attention. This paper proposes to separately train the actor network and critic network. For actor network, the output layer is softmax, whereas for critic network, it is linear.

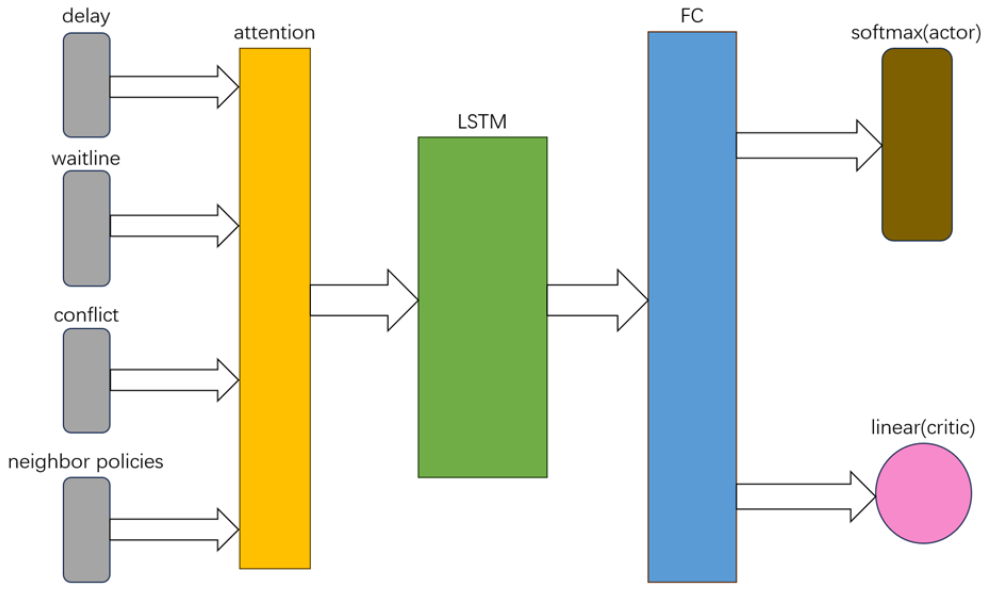


Fig. 3 Structure of deep neural network

3. Results and Discussion

In order to verify the effectiveness and feasibility of this study, this paper mainly focuses on the following two questions: The First one is whether this study can improve pedestrian safety, that is, whether the average amount of conflicts is significantly decreased compared with the other traffic signal timing schemes under the same simulation environment. Another one is whether the study simultaneously takes pedestrian safety and vehicle efficiency into account, that is, how is the overall performance of each intersection which is considered by two metrics: average speed and average delay of vehicles.

3.1. Traffic Signal Timing Scheme for Comparison

In order to evaluate the effectiveness of this paper, two typical traffic signal timing schemes are used for comparison. Each scheme is equipped the same setting of intersection, pedestrian flow and vehicle flow.

3.1.1 Fixed-timing scheme

Phase of traffic signals and the duration of every phase in every intersection is set fixed. And the setting is the same as in section 2.1.2. In this scheme, SUMO's interface TraCI is used to directly obtain data such as the amount of conflicts between pedestrians and vehicles at every moment and the average speed of vehicles to reflect the efficiency of vehicles and pedestrian safety at each intersection.

3.1.2 Deep Q-learning network(DQN) timing scheme

The state of this reinforcement learning algorithm is set by vehicle's location-velocity matrix and pedestrian's location-velocity matrix. The initial action is set to the first 10s of every sub-stage. After learning, every agent dynamically changes the phase duration, which can be 0 second, according to the intersection traffic flow. And the reward function is set as same as the one of this study.

3.2. Comparison of Case One

In order to verify whether pedestrians are in a safer traffic environment, where the amount of conflicts is decreasing, this paper make a comparative experiment between the multi-agent deep reinforcement learning with attention (MADRLA) proposed in this study and other two traffic signal timing schemes illustrated above.

As shown in Fig. 4, in comparison with the fixed-timing scheme, the learning-based timing schemes can adjust the optimal timing scheme in real circumstances according to the pedestrian flow and vehicle flow, and effectively reduce the conflict between pedestrians and vehicles. Though the performance of DQN scheme and MADRLA scheme are similar, the average amount of conflicts of MADRLA is 0.66, which is 16.40% less than that of DQN. In conclusion, MADRLA algorithm can effectively improve the safe-walking environment for pedestrians when crossing the crosswalks.

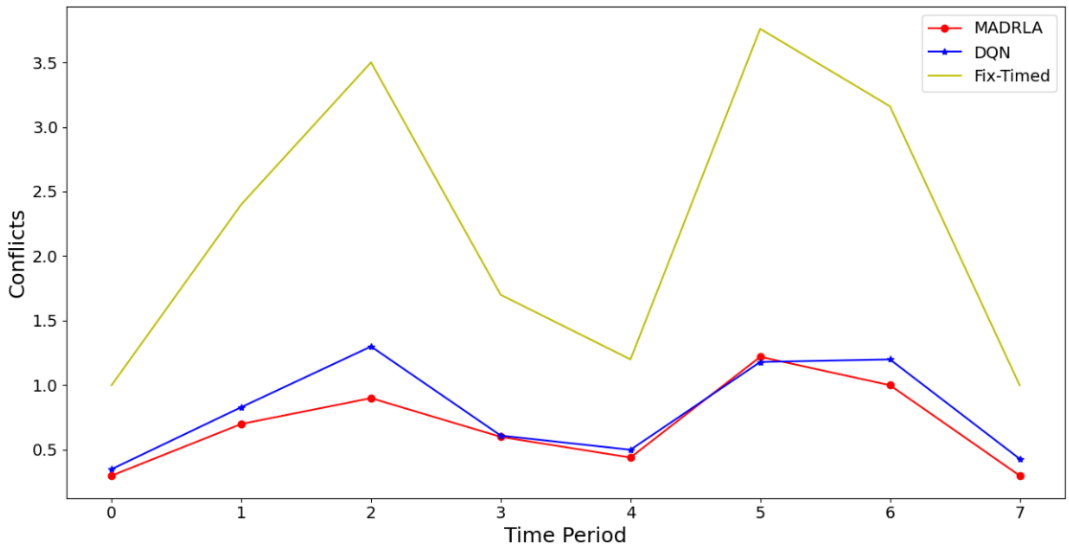


Fig. 4 Comparison of average conflicts between pedestrians and vehicles

3.3. Comparison of Case Two

In order to evaluate the overall performance of MADRLA, this paper design two different experiments. One is to compare the average speed of vehicles under the circumstances with pedestrians walking on the sidewalks and crossings. Another one is to compare the average delay without pedestrian factor.

3.3.1 With pedestrian factor

As shown in Fig. 5, MADRLA algorithm outperform the other two traffic signal timing schemes on the average speed of vehicles on the road net, especially during the peak time (such as time period 3 and 7). Through averaging the data of every period with MADRLA scheme, I get the average speed of vehicles on each road is 22.60km/h, which is 40.12% higher than that of DQN and 29.76% higher than that of fix-timed scheme, which illustrates that the efficiency of vehicle is improving. At the same time, the average speed of vehicles of DQN is 5.60% lower than that of time-fixed scheme, while the average amount of conflicts is 66.04% lower than that of time-fixed scheme as shown in Fig 4. It can be clearly seen that the DQN timing scheme sacrifices a certain amount of vehicle traffic efficiency to ensure the safety of pedestrians at the intersections.

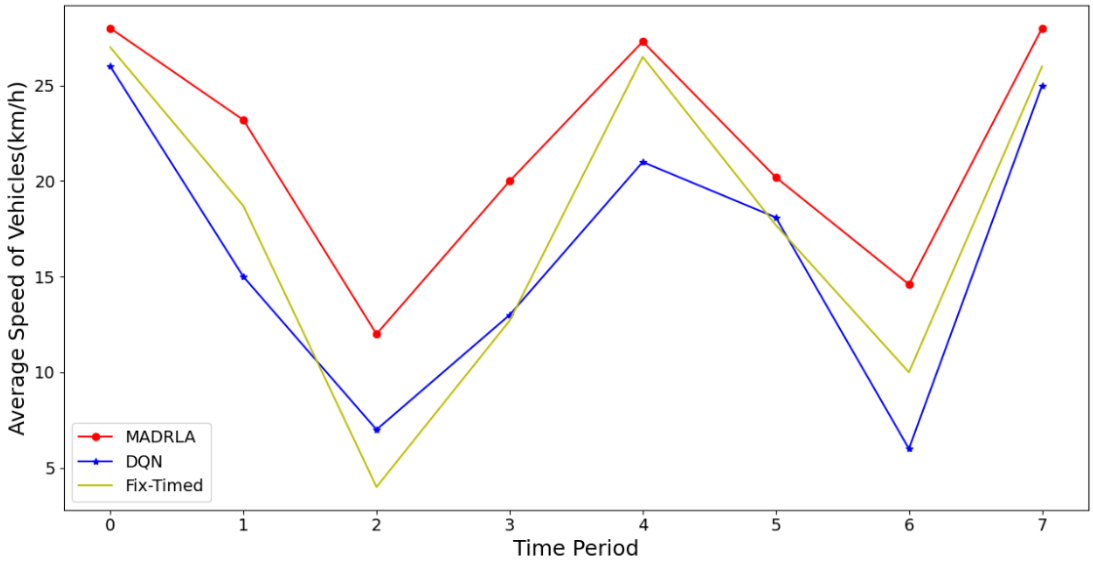


Fig. 5 Comparison of average speed of vehicles

3.3.2 Without pedestrian factor

In order to make comparison of the influence of pedestrian factor on vehicle efficiency, this paper designs an experiment that omits the pedestrian factor in MADRLA scheme and DQN scheme. Fig 6 shows the comparison of average delay of vehicles. To be clear, delay is the difference between vehicle's actual driving time and the driving time under maximum speed of the road.

As Fig. 6 shows, the delay of MADRLA scheme with pedestrian factor is 4.68s, which is increasing from 2.89s when MADRLA scheme takes the pedestrian factor away. And the delay of DQN scheme is 1.27s when it takes the pedestrian factor away, from 7.22s downing to 5.95s. One can say that there will be a certain level of negative influence on vehicle efficiency when scheme takes pedestrian safety into account.

Although there will be negative influence on vehicle efficiency when schemes contain pedestrian safety concern, MADRLA still outperforms when it comes to simultaneously consider pedestrian safety and vehicle efficiency than the other two timing schemes. As we can see in the comparison of average delay of vehicles, MADRLA scheme is 40.02% better than DQN scheme (without pedestrian factor), 53.75% better than DQN scheme (with pedestrian factor), and 48.32% better than the fixed-timed scheme.

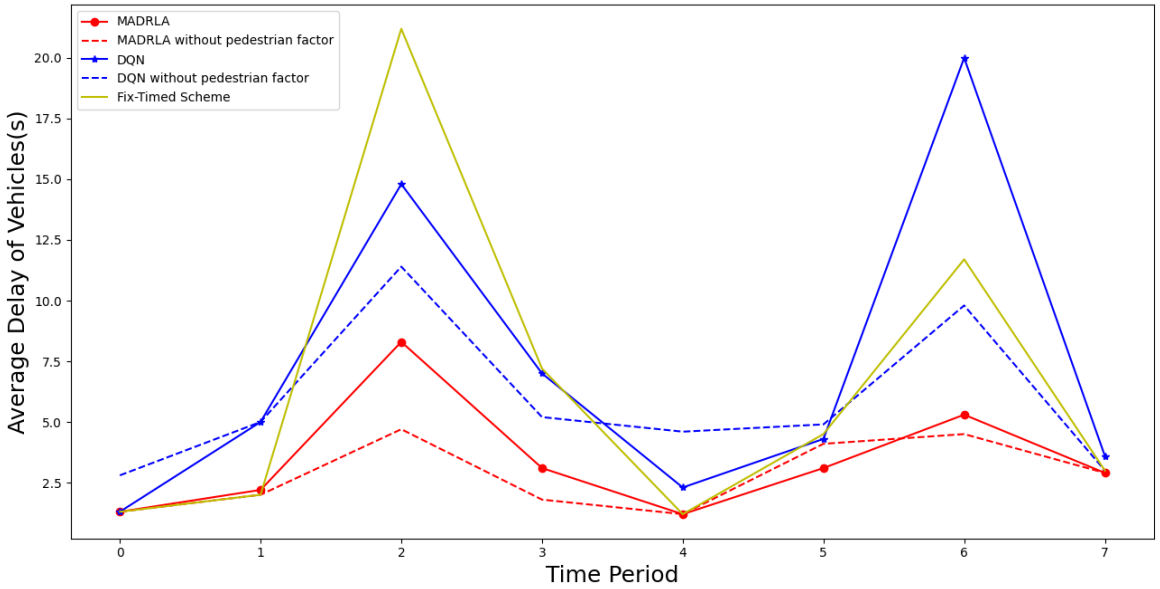


Fig. 6 Comparison of average delay of vehicles

4. Conclusion

This paper has presented an intelligent control scheme for traffic signals at multi-intersections based on pedestrian safety. The scheme is designed under the powerful benefit of multi-agent deep reinforcement learning, and it additionally adds a self-attention layer between the state input and deep neural network which enables the model to stay focus on important state components under the risk of state and action dimension explosion because of the complexity of the spatial-temporal traffic data. Centralized training and decentralized execution also play an important role in downsizing the action space and makes the model easier to learn. This paper verifies the performance of MADRLA scheme with a simulation traffic environment designed in SUMO, which is found that it outperforms the other two traffic signal timing schemes in the amount of conflicts between pedestrians and vehicles, the average speed of vehicles and the average delay of vehicles. Though there is a tradeoff between pedestrian safety and vehicle efficiency, the MADRLA scheme still performs well when it simultaneously concerns the safety of pedestrians and the efficiency of vehicles. There are some improvements can be made in the future, for example, implement this algorithm in real world circumstances, add more traffic modes into the traffic net but not only one type of vehicles like in this paper, make it much complex and closer to the real environment, etc.

References

[1] Road Traffic Accident Database 2009-2011: Traffic Police Department. SPF, Singapore, 2013.

[2] Alan J Miller. Settings for fixed-cycle traffic signals. Journal of the Operational Research Society, 1963, 14: 373–386.

[3] Isaac Porche, Stéphane Lafortune. Adaptive look-ahead optimization of traffic signals. Journal of Intelligent Transportation System, 1999, 3: 209–254.

[4] Dahai Wang, Qian Wang. Cycle lost time L is negative for the fine optimal adaptive TC. The 9th China Intelligent Transportation Annual Conference, 2014: 10.

[5] Varaiya P. The max-pressure controller for arbitrary networks of signalized intersections. In Advances in Dynamic Network Modeling in Complex Transportation Systems, 2013, 27-66.

[6] Guoyou Zhang, Shifeng Song. Traffic Light Control Optimization based one D3QN. Computer and Modernization, 2023(07): 30-35+47.

[7] Tong Wu, et al. Multi-Agent Deep Reinforcement Learning for Urban Traffic Light Control in Vehicular Networks. IEEE Transactions on Vehicular Technology, 2020, 8: 8243–8256.

[8] Lowe R, Wu Y, Tamar A, et al. Multi-agent actor-critic for mixed cooperative-competitive environments. In Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, 6382-6393.

[9] Zundong Zhang, et al. Road network traffic control reinforcement learning algorithms based on Nash-Stackelberg hierarchical game model. Journal of Southeast University (Natural Science Edition), 2023, 334-341.

[10] Haosheng Wu, Jiaoling Zheng, Maofan Wang. TR-light: traffic organization plan optimization algorithm based on multiple traffic signal lights reinforcement learning. Application Research of Computers, 2023, 39(02): 504-509+514.

[11] Xianglin Zhuang, Changxu Wu, Shua Ma. Cross or wait? Pedestrian decision making during clearance phase at signalized intersections. Accident Analysis and Prevention, 2018, 115-124.

[12] Turky A M, et al. The use of genetic algorithm for traffic light and pedestrian crossing control. International Journal of Computer Science and Network Security, 2009, 9(2):88-96.

[13] Liu Y, Liu L, Chen W P. Intelligent traffic light control using distributed multi-agent Q learning. IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017, 1-8.

[14] Zhang Y, et al. Pedestrian-safety-aware traffic light control strategy for urban traffic congestion alleviation. IEEE Transactions on Intelligent Transportation System, 2019, 22(1): 178-193.

[15] Qianlong Zhang, Zhiqun Hu, Hailin Xiao. Research on Intelligent Control Algorithm of Traffic Light Based on Pedestrian Safety. Computer Measurement & Control, 2022, 30(4): 114-120.

[16] Prashanth L, Bhatnagar S. Reinforcement learning with function approximation for traffic signal control. Intelligent Transportation Systems, IEEE Transactions on, 2011, 412–421.

[17] Anni Ren, Dake Zhou, Jinhao Feng, et al. Attention mechanism based deep reinforcement learning for traffic signal control. Application Research of Computers, 2023, 40(02): 430-434.