

Research on Traffic Signal Control Algorithm Based on Deep Reinforcement Learning

Hanfeng Xue

School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shanxi, China
1817532845@qq.com

Zhen Mu

School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shanxi, China
3116308331@qq.com

Pingping Liu

School of Computer Science and Engineering
Xi'an Technological University
Xi'an, 710021, Shanxi, China
1341369601@qq.com

Abstract—Aiming at the significant deficiencies of traditional traffic signal control algorithms in multi-intersection collaboration, unexpected event response and system generalization ability, this paper proposes an intelligent traffic signal control method that integrates a bidirectional gated recurrent unit BGRU with deep reinforcement learning DRL. The method adopts BGRU to model the historical traffic flow data in time sequence and accurately predict the traffic state; and based on this, it constructs deep Q-network intelligences to dynamically optimize the multi-intersection signal timing strategy. The experimental validation on SUMO simulation platform shows that the proposed method effectively improves the control performance. Compared with the traditional fixed-cycle control and adaptive control methods, the proposed method reduces the average vehicle waiting time by 58.8% and 42.2%, and improves the intersection access efficiency by 83.8% and 52.4%, respectively. The study provides new ideas for building an efficient and intelligent urban traffic management system.

Keywords—Deep Reinforcement Learning; Traffic Signal Control; BGRU; DQN

I. INTRODUCTION

With the continuous progress of urbanization and the growing number of cars, the pressure on the urban transportation system has become increasingly intense, and traffic congestion has become one of the key factors restricting the development of cities and affecting the quality of life of the residents. Traffic congestion during the peak period causes a lot of time and energy to be

wasted, and leads to air pollution and frequent traffic accidents, which seriously affects the efficiency of the city's operation and safety [1]. It also leads to air pollution and frequent traffic accidents, which seriously affects the efficiency and safety of urban operation. Enhancing the traffic efficiency of urban road networks has emerged as a central focus within intelligent transportation system research.

Traffic signal control, as one of the core links of urban traffic operation, has an extremely critical role in alleviating congestion at intersections and improving the efficiency of road use [2]. Currently, the traditional traffic signal control methods represented by timing control, inductive control and adaptive control are commonly used in practical applications [3]. Timing control employs historical traffic data to design fixed signal schemes [4]. While operationally simple, it lacks dynamic adaptability and struggles with real-time traffic fluctuations. Induction control offers limited real-time responsiveness but often relies on simplistic rule-based logic, leading to failures in complex scenarios. Adaptive systems like SCOOT/SCATS achieve partial dynamic optimization, yet remain incapable of comprehensive real-time adaptation. Adaptive control systems such as SCOOT and SCATS can realize dynamic optimization to a certain extent [5], however, most of these systems rely on a large number of sensors and rules, and the model itself

lacks the ability of self-learning and evolution, which makes it difficult to adapt to new types of road network structures or sudden traffic events [6]. The performance of traditional methods in complex traffic scenarios has obvious limitations, and there is an urgent need to introduce new control methods that are smarter, more flexible and generalizable.

Based on the above background, a traffic signal control algorithm integrating dual attention mechanism, BGRU timing prediction model and deep Q-network is proposed [7]. To expand, node attention and edge attention are used to jointly model the key areas and traffic relations in the urban road network to achieve spatial perception of the traffic state [8], the BGRU model is introduced to predict the traffic flow in time sequence to improve the system's ability to predict future state changes [9], and the DQN algorithm in deep reinforcement learning is combined to learn and optimize the signal timing strategy [10], so that the system can adjust the signal scheme adaptively in different traffic conditions to achieve the goal of mitigating the traffic signal control system's impact. The system can adaptively adjust the signaling scheme under different traffic conditions to achieve the goal of relieving congestion and improving traffic efficiency.

II. TRAFFIC SIGNAL CONTROL ALGORITHM

A. Attention Mechanism Optimization

1) Nodal Attention Mechanism

The node attention mechanism dynamically identifies the degree of influence that each neighboring node has on the current node's state [11], which is done as follows, at the time of the feature aggregation phase, each neighboring node is assigned a weight coefficient, which is computed with the help of an attention function that can be trained, as shown in Equation (1).

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T [Wh_i | Wh_j]))}{\sum_{k \in N(i)} \exp(\text{LeakyReLU}(a^T [Wh_i | Wh_k]))} \quad (1)$$

Where i and j denote feature representations of node i and neighbor j , W and a are trainable

parameters, and n represents the neighbor set of i . In this way, the model learns that certain key nodes have a higher weight on the current decision in different traffic scenarios, thus modeling the propagation and impact of traffic states more accurately.

2) Design of the Side Attention Mechanism

In addition to nodal influences, differences in the attributes of the roads themselves play a decisive role in the propagation process of traffic flows [12]. For example, trunk roads and side roads play different roles in undertaking traffic flow and generating congestion. For this reason, this paper designs the edge attention mechanism, which introduces dynamic weight control to the edges in the graph. The edge attention mechanism generates edge weight coefficients based on the static and dynamic attributes of roads together, as shown in Equation (2).

$$\beta_{ij} = \sigma(w^T \cdot \Phi(e_{ij}, h_i, h_j)) \quad (2)$$

Where E is the feature vector of edge I , Q is the edge node feature fusion function, O is the activation function, and W is the learning parameter. Through the edge attention mechanism, the system can more reasonably assess the actual traffic conditions on the path “from an intersection to another intersection”, and provide more accurate weight reference for signal timing.

3) Dual Attention Mechanism

In the model implementation, node attention and edge attention are simultaneously integrated into the message passing process of the graph neural network [13]. In each layer of propagation, when the message is incoming from the neighbor nodes, the weights of the edges are firstly adjusted by the edge attention mechanism, and then the weighted aggregation of the neighbor information is completed by the node attention mechanism. The overall node update process is shown in Equation (3).

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in N(i)} \alpha_{ij} \cdot \beta_{ij} \cdot Wh_j \right) \quad (3)$$

This mechanism achieves precise control of spatial relationship modeling, enabling the system to identify and amplify the impacts of critical nodes and critical paths, and effectively improving the spatial perception of traffic signal control decisions.

B. Traffic Timing Prediction Model

1) Basic Principles of GRU

In traffic signal control systems, accurate prediction of future traffic conditions plays a key role in reaching the optimal control strategy. Most of the traditional traffic prediction methods use statistical modeling, classical machine learning like SVR, decision tree, or standard recurrent neural networks like RNN, GRU, LSTM to build models, which can depict time series patterns to a certain extent. However, there are still many problems in the complex urban transportation environment, such as relying on unidirectional historical information, slow response to sudden traffic changes, and insufficient prediction accuracy.

This work introduces a bidirectional GRU architecture to capture temporal dependencies in traffic data, optimizing the prediction module for improved accuracy and robustness

2) BGRU Model

To make up for the shortcomings of unidirectional GRU, this paper introduces the BGRU structure, i.e., on the basis of the original GRU, two GRU networks are constructed simultaneously in the forward and backward directions to process the temporal information from the past to the present and from the past to the present, respectively. The core idea of the BGRU lies in the fact that a more comprehensive representation of the temporal features is formed by splicing or weighted fusion of the hidden states in the two directions.

This bi-directional structure can effectively improve the modeling ability of global time dependence, so that the model can focus on the antecedents of the current traffic situation, and also “backward” the reasonableness of the current decision-making from the future trend, so as to improve the overall prediction performance.

3) Model Training and Fusion Strategies

During the training process, the BGRU receives the historical time series data output from the traffic state awareness module, whose input dimensions are generally related to traffic flow, average speed, road load, and queue length, and its output is the distribution of future states within the predicted time window, such as the traffic flow value per minute in the next 5 minutes, for example.

In order to prevent information redundancy and high model complexity, this paper applies the attention fusion mechanism in the output stage to weight the bidirectional states, rather than direct splicing, to reduce dimensionality and enhance generalization capability. The fusion formula is shown in equation (4).

$$h_t^{\text{final}} = \alpha_t \cdot \vec{h}_t + (1 - \alpha_t) \cdot \overleftarrow{h}_t \quad (4)$$

where A is output by a trainable attention network that dynamically adjusts the importance of positive and negative information.

C. Traffic Timing Prediction Model

1) Double DQN

Traditional DQN cannot distinguish between “whether the current state itself is important” and “the relative advantage of an action in the current state”. For this reason, this paper adopts the Dueling DQN structure, which splits the Q-value into two parts, namely, state value and action advantage, as shown in Equation (5).

$$Q(s, a) = V(s) + A(s, a) - \frac{1}{|A|} \sum_a A(s, a') Q(s, a) \quad (5)$$

This structure designs two separate sub-network branches in the network to learn V and A independently before combining them for the final Q value.

In this way, the model learns that certain key nodes (e.g., congested intersections, high-traffic intersections) have a higher weight on the current decision in different traffic scenarios, and thus

more accurately models the passing value of the traffic state. This optimization structure enables the model to better understand the importance of the traffic state itself and still make reasonable choices when the intersection state changes drastically, thus improving the strategy generalization capability.

2) Optimize Experience Playback

In traffic control scenarios, certain state transfers (e.g., sudden congestion, abnormal traffic) contribute much more to policy learning than regular states. In order to improve the efficiency of sample utilization, this paper adopts Prioritized Experience Replay (PER) to rank the importance of samples based on the TD error, as shown in Equation (6).

$$P(i) = \frac{|\delta_i|^\alpha}{\sum_k |\delta_k|^\alpha} \quad (6)$$

where R denotes the TD error of the i th sample and α controls the degree of prioritization. This mechanism makes high-informative samples sampled more frequently, which accelerates the policy convergence and improves the learning efficiency.

3) Reward Function Design

The design of reward function as the learning signal of reinforcement learning directly affects the strategy direction. In order to avoid the bias brought by a single objective, this paper constructs a multi-objective weighted reward function as shown in (7).

$$R = \alpha \cdot QueueLength + \beta \cdot WaitingTime + \gamma \cdot PhaseSwitchPenalty \quad (7)$$

Where $QueueLength$ denotes the average queue length, $WaitingTime$ denotes the average vehicle waiting time, and $PhaseSwitchPenalty$ denotes the penalty term for frequent phase switching. By adjusting the weight coefficients to achieve a balanced consideration of different optimization objectives, so that the strategy learned by the model is more in line with the actual needs.

III. SIMULATION EXPERIMENTS

A. Simulation Environment

In order to effectively verify the performance of traffic signal control strategies, this study relies on the microscopic traffic simulation platform SUMO to build a complete simulation environment. SUMO, as an open-source and highly scalable traffic simulation tool, can accurately simulate the moving behavior of each vehicle in urban traffic, and is widely used in the research and development of intelligent transportation systems.

SUMO has flexible traffic modeling capabilities, including road network modeling, traffic flow definition, traffic signal control and vehicle behavior modeling, etc., the software's simulation engine can support second-level control accuracy, which can be used to study the real-time control strategies, such as perceptual adaptive signal scheduling, traffic optimization based on deep reinforcement learning, etc., SUMO also supports interfaces with Python and other languages, such as TraCI, which facilitates integration with custom controllers, machine learning models, and other systems. The simulation environment selected for this simulation is a typical urban road network consisting of one intersection in the center of a city, with the simulation area as shown in Figure 1.

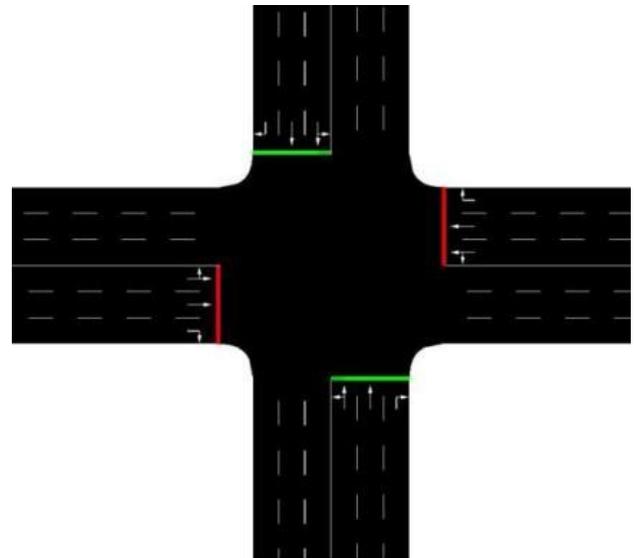


Figure 1 Single Fork Simulation Road

The spacing between intersections is about 250 meters, the road network contains straight, left turn, right turn and straight two-way three lanes, these lanes were extended in the four directions of southeast, northwest and north, ensuring compliance with statutory traffic requirements and practical conditions, the speed limit of the lanes is set to 70km/h, and the maximum speed of the vehicle is 60km/h.

Three traffic load scenarios are simulated to comprehensively assess the signal control system's performance and adaptability "low load", "medium load" and "high load". In the process of simulation experiment, quantifying the traffic control system's performance and adaptive capacity, three kinds of traffic load scenarios with different intensities are introduced, specifically "low load", "medium load" and "high load". These three types of traffic scenarios, low load scenarios simulate a more sparse vehicle, smooth passage of the morning and evening non-peak hours, medium load scenarios represent the daily traffic state, when the traffic flow is at a moderate level, the intersection pressure is relatively balanced, and high load scenarios simulate the peak or emergency under the dense traffic conditions, traffic flow and queuing phenomenon is very obvious. By virtue of the system simulation in a variety of different traffic densities, the control effect of the model can be examined in normal working conditions, and its robustness and generalization ability can be evaluated in complex and extreme traffic conditions, which gives a key basis for the system to adapt to the real-life traffic fluctuations.

B. Design of Assessment Indicators

In order to comprehensively and fully evaluate the operation effect and intelligent decision-making ability of traffic signal control system in different scenarios, this paper selects five types of key performance indicators, as shown in Table 1, and starts from the two levels of vehicle micro-behavior and system macroscopic scheduling, to carry out a systematic, multi-dimensional analysis and comparison.

TABLE I. INDICATOR MEANING AND ASSESSMENT DIMENSIONS

Indicator name	Explanation of meaning	Assessment dimensions
Average vehicle waiting time	Average length of time a bicycle stays at a red light, in seconds	Accessibility
Average vehicle travel time	Total time required from entry to exit, in seconds	System-level scheduling efficiency
Signal switching frequency	Number of signal changes per unit of time, unit: times/minute	Decision stability and control smoothing

C. Ablation Experiment Program

To validate the proposed method's performance advantages across varied traffic conditions, this paper elaborately designs three groups of representative ablation experimental scenarios, as shown in Table 2, which correspond to the typical rule-based control, perceptual control and intelligent learning control methods in the current traffic control field. By systematically evaluating the performance of each scheme under the same simulation scenario, it can more intuitively present the improvement effect of this paper's method in terms of scheduling efficiency, system stability and generalization ability.

Scheme 1 is the fixed-time control strategy (FIXED), which pre-sets the duration of each signal phase and carries out the operation according to a fixed cycle, and its control logic does not rely on any real-time traffic state information. As a typical representative of the traditional methods, this strategy has the advantages of easier implementation and lower cost, but lacks of dynamic response capability and intelligent adjustment mechanism, which makes it difficult to cope with the traffic situation in the future. However, it lacks the ability to respond dynamically and the mechanism of intelligent adjustment, which makes it difficult to cope with complex road conditions.

Scheme 2 is vehicle-aware adaptive control (ADAPTIVE), with the help of real-time detection of the queue length of each lane and traffic density and other indicators, the signal phase to carry out relatively flexible adjustment, the method does not have the long-term learning ability, and can not be predicted traffic trends, in the high load, sudden

congestion and other conditions there is still a performance bottleneck.

Scenario 3 is the Deep Reinforcement Learning control method (DB-DRL) proposed in this paper, which utilizes BGRU timing feature prediction and improved DQN reinforcement learning strategy, and this method achieves efficient scheduling and global optimal signal allocation for different traffic scenarios by using the intelligent body to learn the optimal control strategy by itself.

TABLE II. ABLATION EXPERIMENT FORM

Program name	fixed phase	perceptual ability	learning to predict ability
FIXED	✓	✗	✗
ADAPTI VE	✓	✓	✗
DB-DRL	✓	✓	✓

IV. RESULTS OF THE EXPERIMENT

A. Access Efficiency Analysis

The results of the experiments are shown in Fig. 2, Fig. 3 and Fig. 4, respectively. DB-DRL shows strong learning ability and adaptability during the simulation process, and in the overall level, its control effect has a large optimization potential, and in some rounds, DB-DRL succeeds in achieving a shorter average queue length, and has the ability to improve the control strategy according to the feedback from the environment continuously. It has the ability to continuously improve the control strategy based on the feedback from the environment.

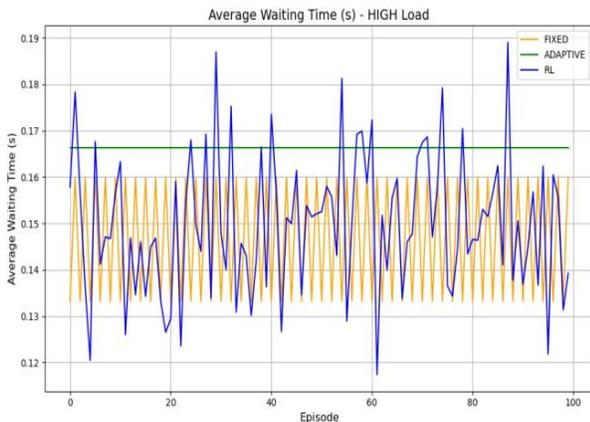


Figure 2 Average waiting time for high load vehicles

Under high load scenarios, the FIXED strategy has a stable waiting time but insufficient responsiveness, the ADAPTIVE strategy has some adjustment capability but lacks a learning mechanism, and the DB-DRL strategy demonstrates strong adaptability and optimization potential, with a shorter average waiting time, which effectively mitigates the peak congestion problem.

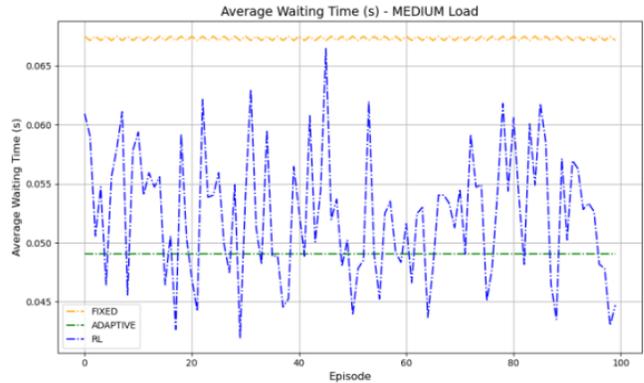


Figure 3 Average waiting time for medium-load vehicles

Under medium load, the fixed-time control strategy cannot adjust the signal phase, resulting in low efficiency; the adaptive method is slightly improved, but still passive in the face of changes; DB-DRL can dynamically learn and quickly respond to changes in the traffic state, and exhibits superior waiting time control in multiple weekly periods.

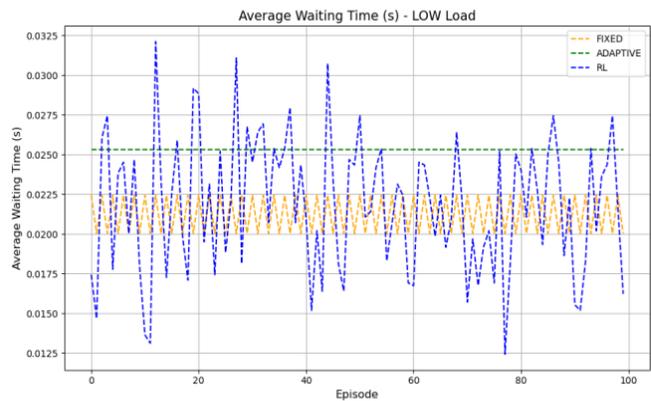


Figure 4 Average waiting time for low load vehicles

When the traffic flow is low, there is little difference in the overall performance of each strategy, and the waiting time of the DB-DRL strategy remains at a low level despite slight

fluctuations, showing good adaptability and generalization ability to sparse traffic environments, and strong global optimization capability.

B. System-level Scheduling Efficiency Analysis

Figures 5, 6, and 7 show the average vehicle travel times of the three control methods under high, medium, and low load traffic scenarios, respectively, reflecting the impact of different strategies on system-level scheduling efficiency.

FIXED shows a highly stable trend in the figure, and there is an upper limit to the scheduling efficiency, which is difficult to effectively alleviate congestion during peak periods, and the system scheduling responsiveness is weak. ADAPTIVE is able to adjust the signal cycle according to some traffic conditions, but the adaptability of its scheduling rules is still limited, and it cannot adequately respond to the complex road conditions, and the system scheduling efficiency is relatively low. DB-BRL exhibits a much stronger scheduling potential. Although there is some fluctuation in the average passing time, the overall mean value is better than the other two strategies, and the passing time is significantly reduced in some cycles. This volatility comes from the strategy's continuous exploration and adjustment of complex road conditions during the training process, reflecting its strong learning ability and dynamic optimization capability. From the perspective of system efficiency, the DB-DRL strategy has higher scheduling flexibility and response speed, and is more capable of realizing global optimal control under high load.

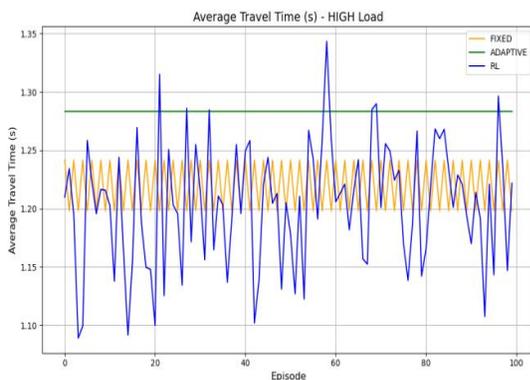


Figure 5 Average passage time for high-load vehicles

In the high-load environment, the FIXED and ADAPTIVE strategies have higher passing times and lack the ability to cope with complex traffic changes. the DB-DRL strategy has a more fluctuating but overall better average passing time, and significantly reduces the passing time in some rounds, reflecting its sensitivity and adaptability to high-load scheduling.

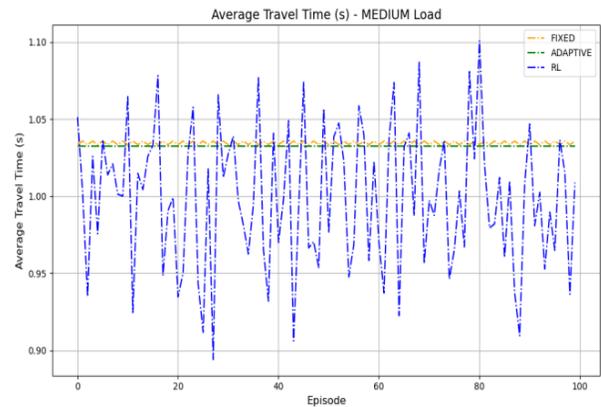


Figure 6 Average travel time for medium-load vehicles

In the medium-load scenario, DB-DRL still outperforms the two traditional control methods, and despite some fluctuations, its passing time is generally shorter, showing the ability to respond quickly to the changes in road conditions, and realizing the optimal scheduling strategy under different flow conditions.

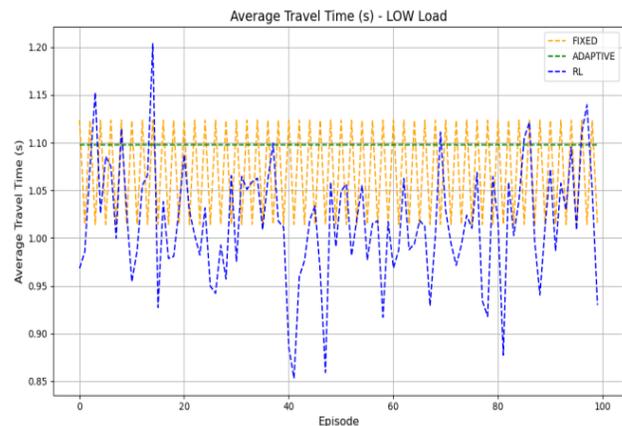


Figure 7 Average low load vehicle travel time

Under the low load scenario, the passage time of DB-DRL strategy fluctuates, but it is overall better than FIXED and ADAPTIVE, and is close to the minimum value in some stages. It indicates that its trained strategy has the ability to efficiently

regulate the sparse traffic state, and the overall scheduling performance is excellent.

C. Decision Stability and Control Smoothness Analysis

Figures 8, 9, and 10 show the average cumulative rewards of vehicles for the three control methods under high, medium, and low load traffic scenarios, respectively, reflecting the impact of different strategies on system-level scheduling efficiency.

The curve of FIXED is highly stable with minimal fluctuation, which is unable to adapt to the dynamic traffic, has limited control effect, and lacks optimization space. The reward of ADAPTIVE has almost no fluctuation, has insufficient optimization ability under high loads, and the system responds conservatively. The curve of DB-DRL has obvious fluctuation, and although the optimal value can be reached locally, there is an unstable trend in the whole. This indicates that the DB-DRL strategy is still in the stage of exploration and strategy updating during the training process, and this fluctuation also reflects the ability of the system to constantly try to find the optimal value, which is a manifestation of the gradual learning of complex traffic features by the intelligent body.

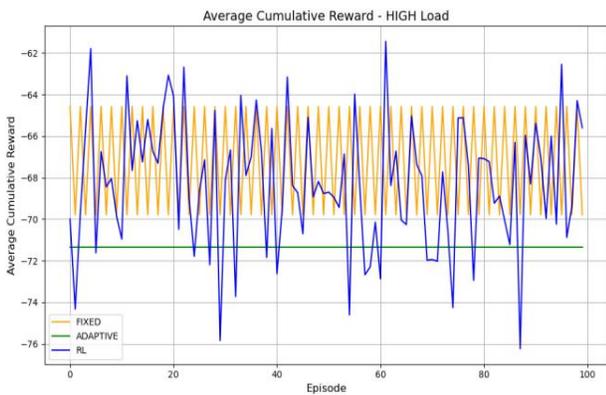


Figure 8 High load average cumulative rewards

The FIXED strategy has stable but low returns, ADAPTIVE is slightly better but lacks volatility, and the DB-DRL strategy has large fluctuations in cumulative rewards and reaches the optimal value in multiple rounds despite instability, suggesting that it has the ability to continuously learn and

optimize, and is able to tap into a better strategy in complex environments.

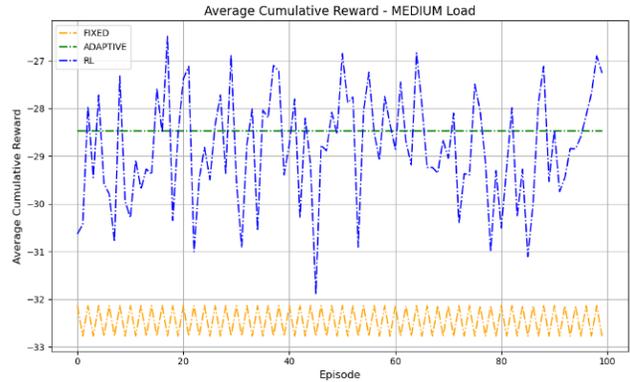


Figure 9 Average cumulative reward for medium Load

Under medium-load conditions, the DB-DRL strategy gradually shows strong learning ability, and the accumulated rewards show an upward trend, which is overall better than FIXED and ADAPTIVE despite fluctuations, reflecting the optimization potential of the system in strategy iteration.

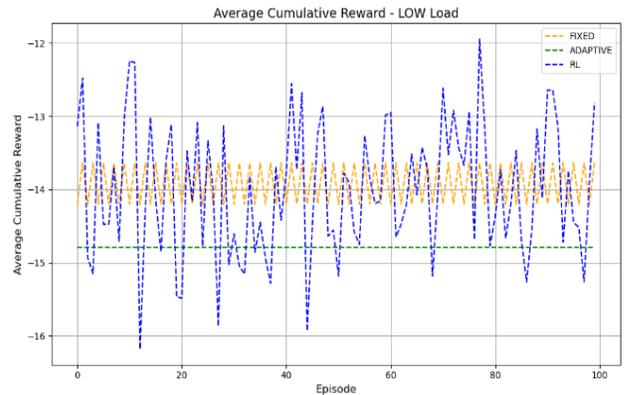


Figure 10 Low load average cumulative reward

Under low load conditions, the differences among the three strategies are narrowed, and FIXED and ADAPTIVE are stable but with low rewards. DB-DRL obtains higher cumulative rewards in most of the cycles, which indicates that it has a good generalization ability and can stably output a better strategy under various traffic load conditions.

Taken together, DB-DRL has the ability to continuously optimize, dynamically respond, and adapt to complex traffic situations, which is much better than traditional control strategies, despite

the instability of DB-DRL at the early stage of training. In the future practical deployment, with the improvement of model training and the enhancement of data support, DB-DRL emerges as a key technology for enhancing urban road network efficiency.

V. CONCLUSIONS

In this paper, an intelligent system integrating graph structure modeling, bidirectional gated recurrent unit (BGRU) traffic flow prediction, and improved deep Q-network (DQN) reinforcement learning control is proposed and implemented for the difficulties of multi-intersection coordination and dynamic response in urban traffic signal control. The bidirectional time-series modeling of historical traffic states by BGRU achieves accurate prediction of short-term traffic flow and provides prospective support for reinforcement learning decision-making; on this basis, the state representation, reward function and action selection mechanism of DQN are optimized so that it can learn the optimal signal timing strategy efficiently and stably. The experimental validation on the real road network of SUMO simulation platform shows that compared with the fixed cycle control, this method significantly reduces the average waiting time of vehicles by 58.8% and improves the passing efficiency by 83.8%; compared with the adaptive control, it reduces the waiting time by 42.2% and improves the efficiency by 52.4%, and at the same time, effectively shortens the queue length, and shows good robustness and extensibility in the handling of emergencies and the coordination of multiple intersections. Robustness and scalability. This study provides an innovative theoretical approach and a scalable implementation framework for efficient intelligent traffic signal control, which has important engineering application prospects.

REFERENCES

- [1] Xu, S., Sun, C., & Liu, N. (2024). Road congestion and air pollution-analysis of spatial and temporal congestion effects. *Science of The Total Environment*, 945, 173896.
- [2] Majstorović, Ž., Tišljarić, L., Ivanjko, E., & Carić, T. (2023). Urban traffic signal control under mixed traffic flows: Literature review. *Applied Sciences*, 13(7), 4484.
- [3] Qadri, S. S. S. M., Gökçe, M. A., & Öner, E. (2020). State-of-art review of traffic signal control methods: challenges and opportunities. *European transport research review*, 12, 1-23.
- [4] Qadri, S. S. S. M., Gökçe, M. A., & Öner, E. (2020). State-of-art review of traffic signal control methods: challenges and opportunities. *European transport research review*, 12, 1-23.
- [5] Moganarangan, N., Balaji, N., Suresh Kumar, R. G., Balaji, S., & Palanivel, N. (2018). Study on static and dynamic traffic control systems. *International Journal of Pure and Applied Mathematics*, 119(12), 565-579.
- [6] *ng, J., Wei, D., Zhou, S., Wang, T., Huang, Y., & Chen, H. (2024). A comprehensive study on self-learning methods and implications to autonomous driving. *IEEE Transactions on Neural Networks and Learning Systems*.
- [7] Zhao, Z., Wang, K., Wang, Y., & Liang, X. (2024). Enhancing traffic signal control with composite deep intelligence. *Expert Systems with Applications*, 244, 123020.
- [8] Li, G., Wang, J., Zhao, Z., Chen, Y., Tang, L., & Li, Q. (2024). Advancing complex urban traffic forecasting: A fully attentional spatial-temporal network enhanced by graph representation. *International Journal of Applied Earth Observation and Geoinformation*, 134, 104237.
- [9] He, S., Sang, X., Yin, J., Zheng, Y., & Chen, H. (2023). Short-term runoff prediction optimization method based on BGRU-BP and BLSTM-BP neural networks. *Water Resources Management*, 37(2), 747-768.
- [10] Wu, Z., Wang, S., Ni, C., & Wu, J. (2024). Adaptive traffic signal timing optimization using deep reinforcement learning in urban networks. *Artificial Intelligence and Machine Learning Review*, 5(4), 55-68.
- [11] Xu, D., Cheng, W., Luo, D., Gu, Y., Liu, X., Ni, J., ... & Zhang, X. (2019, November). Adaptive neural network for node classification in dynamic networks. In 2019 IEEE International Conference on Data Mining (ICDM) (pp. 1402-1407). IEEE.
- [12] Zhang, M., Huang, T., Guo, Z., & He, Z. (2022). Complex-network-based traffic network analysis and dynamics: A comprehensive review. *Physica A: Statistical Mechanics and its Applications*, 607, 128063.
- [13] Sun, C., Li, C., Lin, X., Zheng, T., Meng, F., Rui, X., & Wang, Z. (2023). Attention-based graph neural networks: a survey. *Artificial intelligence review*, 56(Suppl 2), 2263-2310.