Congestion-Aware Path Planning with Vehicle-Road Cooperation in AIoV

Guoqiang Liu, Muhammad Bilal, *Senior Member, IEEE*, Qingyang Wang, Xiaolong Xu*, *Senior Member, IEEE*

***Abstract*—With the gradually integration of Internet of Vehicles (IoV) and Artificial Intelligence (AI), Artificial Intelligence of Vehicles (AIoV) is emerging as a novel paradigm with advanced capability for information gathering and decision-making. Leveraging massive traffic information facilitated by vehicle-road coordination in AIoV, path planning has the potential to effectively mitigate existing traffic problems, such as road congestion, improving traffic performance. However, the dynamic nature of traffic flow and the complexity of road networks increase the difficulty of path planning, posing a serious threat to road safety. In response to this challenge, a reinforcement learning based path planning scheme with traffic flow prediction, named RPFP, is proposed. RPFP consists of two fundamental components: precise traffic flow prediction and intelligent path planning. Specifically, the temporal convolutional network (TCN) is innovatively integrated into the spatiotemporal graph neural network (STGNN), providing accurate traffic flow prediction by comprehensively capturing spatial and temporal patterns. Informed by predicted traffic congestion, a path planning method utilizing dueling double deep q-network (D3QN) algorithm is employed to navigate within complex road networks. Eventually, RPFP was evaluated for its effectiveness through comprehensive experiments conducted on real traffic datasets. The superiority of RPFP was further substantiated via comparisons with multiple baseline schemes.**

***Index Terms*—Path planning, Traffic flow prediction, Vehicle-road cooperation, AIoV, Reinforcement learning.**

I. Introduction

THE learning, perception, and decision-making capabilities in artificial intelligence (AI) systems and involvement of human cognitive abilities, endowing AI with the ability to execute complex thought processes with high accuracy [1], [2]. As an AI-integrated system designed to provide intelligent traffic management, intelligent transportation dystems (ITS) benefit from AI in information gathering and smart traffic decision-making. The Internet of Vehicles (IoV), as a significant component of ITS, undergoes further development propelled by advanced technologies such as edge computing, wireless communication, and sensor technologies [3]–[5]. Due to AI's capability to significantly enhance the efficiency of massive data collection and decision-making intelligence, it is gradually converging with IoV to create a new paradigm known as Artificial Intelligence of Vehicles (AIoV) [6]. In AIoV, data originating from diverse sensors with varying sources and types can undergo efficient and unified processing, thereby fostering enhanced collaboration between vehicles and the road infrastructure. Relying on broader vehicular-road coordination, AIoV plays an increasingly significant role in ITS, providing intelligent services for vehicles, such as advanced detection of road congestion and intelligent path planning [7].

Traffic congestion is becoming an increasingly severe global issue, primarily due to the continuous growth of vehicles worldwide and significant increase in travel frequency [8]. Mega-cities like Beijing, New York, Tokyo, London, and others experience particularly severe congestion, significantly impacting residents' satisfaction. Statistical data indicates that the annual global economic losses attributable to traffic congestion amount to several hundred billion dollars. In pursuit of congestion alleviation, improved traffic performance, and enhanced traffic safety, various smart traffic control methods ndergo extensive research. These encompass intelligent traffic signal optimization (ITSO), intelligent parking systems (IPS), and intelligent lane management (ILM) [9], [10]. Path planning, incorporating real-time traffic conditions, empowers vehicles to proactively evade congested segments, thereby averting congestion escalation. This approach is recognized as a viable solution to effectively address traffic congestion issues and enhance driving efficiency [11].

Relying on vehicle-road cooperation, vehicles can obtain road conditions, such as traffic flow, from roadside units (RSUs) during path planning, employing this information as the theoretical foundation for planning decision. However, only depending exclusively on real-time traffic flow for planned paths, to be traversed in the future, may pose challenges in achieving comprehensive and efficacious avoidance of traffic congestion. Therefore, accurate prediction of future traffic flow becomes imperative [12]. n consideration of the attributes of road networks, the prediction of traffic flow can be conducted along two dimensions: temporal and spatial. Deep learning prediction models, such as Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU), demonstrate proficiency in time-series modeling of traffic flow [13]. Leveraging historical traffic flow data as training data, these models effectively capture time-series features inherent in traffic flow. Concur-

*Corresponding author

Guoqiang Liu is with the School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China. E-mail: bwq0709@gmail.com.

Muhammad Bilal is with the School of Computing and Communications, Lancaster University, Bailrigg, Lancaster LA1 4WA, United Kingdom. E-mail: m.bilal@ieee.org.

Qingyang Wang is with the School of Reading, Nanjing University of Information Science and Technology, Nanjing 210044, China. E-mail: wqy3098684229@gmail.com.

Xiaolong Xu is with the School of Software, Jiangsu Province Engineering Research Center of Advanced Computing and Intelligent Services, also with Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET), Nanjing University of Information Science and Technology, Nanjing 210044, China. E-mail: xlxu@ieee.org.

rently, Graph Neural Networks (GNN) model the topological structure of intricate road networks, accounting for the mutual influence between roads. Through iterative information aggregation between adjacent nodes, GNNs can adeptly capture spatial features integral to understanding traffic flow [14].

Following the prediction of future traffic flow, congestion-aware path planning within complex road networks can be executed based on the prediction results [11]. Dependent solely on manual path planning by drivers for departure and destination points or centralized path planning by dispatchers, however, entails substantial resource consumption and encounters challenges in efficiently avoiding congested segments. Intelligent path planning methods grounded in machine learning (ML) and deep learning (DL) find extensive application in addressing this issue. Algorithms, including A-Star Search Algorithm (A*), ant colony optimization (ACO), simulated annealing, genetic algorithms, particle swarm optimization, among others, have been employed in intelligent path planning endeavors and have demonstrated effectiveness [15]–[17].

While numerous traffic flow prediction schemes exist, it remains crucial to account for both spatial and temporal features to ensure accurate predictions in complex road networks. Furthermore, the dynamic and intricate nature of traffic conditions presents challenges for conventional path planning algorithms, including the imperative need for real-time responsiveness and environmental adaptability. So, a congestion-aware intelligent path planning scheme with vehicle-road Cooperation in AIoV is proposed in this paper. The main contributions of our work are as follows:

- The path planning model, facilitated by vehicle-road cooperation, is constructed under AIoV environment.
- By innovatively incorporating temporal convolutional networks (TCN) into spatiotemporal graph neural networks (STGNN), accurate prediction of complex traffic flow in road networks is achieved.
- By employing dueling double deep q-network (D3QN) algorithm, congestion-aware intelligent path planning which superiors in adapting to dynamic environments is accomplished.
- Conducting comprehensive experiments on real traffic datasets and comparative analysis to prove the efficiency and superiority of the proposed scheme.

The remainder of the paper is organized as follows. In section II, the related researches is elaborated and summarized. Then, system framework and travel consumption model are formulated, followed by the definition of the problem to be addressed. The path planning scheme in AIoV is elaborated, including traffic flow prediction and D3QN based path planning, in section IV. In section V, experimental results are presented and analyzed. Finally, the conclusions are presented in section VI.

II. Related Work

In this section, the existing researches related to our work are reviewed from the aspect of Artificial Intelligence for vehicle-road collaboration, traffic flow prediction and route planning.

Development of Artificial Intelligence technologies, including deep learning (DL) and reinforcement learning (RL), promotes popularization of intelligent services. Also, vehicle-road collaboration connects vehicles with intelligent roadside units, improving insufficiency of on-board computing power and providing extra environmental information for vehicles. To this end, combination of Artificial Intelligence and vehicle-road collaboration has attracted great attention to provide delay-sensitive Internet of Vehicles (IoV) services. To help emergency vehicles arrive faster, Ding et al. [18] proposed a learning-based cooperative vehicle-road scheduling framework named LEVID. In LEVID, a real-time route planning module was designed based on the artificial potential field method and a traffic signal control module was designed using graph attention RL. The two modules interacted with each other and made decisions iteratively. Chen et al. [19] proposed vehicle-infrastructure cooperation based approach to detect road hazards. They generalized vital features from few road hazard data through a meta-learning paradigm, designed a lightweight detection model to adapt to vehicles' low computing power, and used knowledge distillation to reduce complexity of model and data. The proposed approach remarkably improves the accuracy and traffic flow prediction. Cui et al. [20] presented a data-driven Cloud-Fog-Edge Collaborative Driver-Vehicle-Road framework to provide high-quality intelligent transportation system services while protecting customers' privacy. Cao et al. [21] reduced delay of IoV services by accurately determining vehicle network conditions based on the fuzzy theory and effectively offloading service requests based on RL.

As a key part of intelligent transport system, accurate traffic flow prediction makes great contributions on mitigating congestion and making safer and cost-efficient travel. Ma et al. [22] developed a novel traffic flow prediction model to further improve the accuracy of short-term traffic flow prediction. They obtained a stable time series as input of model by conducting time series analysis as well as smoothing and standardization processing on traffic flow data, and established an improved long short-term memory network (LSTM) model based on LSTM and bidirectional LSTM. Li et al. [23] developed a novel multisensor data correlation graph convolution network model, called MDCGCN, to improve traffic flow prediction accuracy. Compared with existing work, MDCGCN enables to conduct medium and long-term traffic prediction through analyzing the complex and changeable spatial–temporal correlation among roads. Different from former studies mainly focused on flow forecasting in a regular grid region, Ali et al. [24] designed a deep hybrid spatio-temporal dynamic networks, called DHSTNet, to forecast both the inflow and outflow of irregular grid regions at the same time. They also developed and integrated graph convolutional network with DHSTNet model to concurrently capture all dependencies in the irregular grid regions. Wang et al. [25] proposed an accurate locality-sensitive hashing-based traffic flow prediction approach to ensure high-accuracy and time-efficient traffic flow prediction while protecting customers' privacy.

As another significant part of intelligent transport system, reasonable route planning can help minimise logistics and tour costs. Teng et al. [26] provided a detailed overview and discussion of the latest motion planning methods for autonomous vehicles. These methods contribute to enhancing the convenience of autonomous vehicles and further promoting their global proliferation. In [27], the problems caused by traffic congestion are thoroughly analyzed, highlighting the effectiveness of traffic control, including path planning, in addressing these issues. Additionally, [27] introduced path planning in mixed traffic environments. Lin et al. [28] proposed a distributed-learning-based vehicle routing decision algorithm using multi-agent RL to alleviate the traffic congestion. The proposed algorithm not only greatly shorten decision delay but also effectively reduces waiting time of vehicles on the road. Liang et al. [29] designed a preserving-privacy route planning scheme in vehicular ad-hoc network based on oblivious transfer. In the scheme, vehicles deduce information of RSUs with the help of the certification authority who does not know the source of deduced information. After fast authentication is achieved between vehicles and RSUs, communication between adjacent vehicles is established. The scheme provides fast route planning for vehicles while well preserving customers' privacy. In order to optimize route planning in the air-ground integrated network, Cai et al. [30] proposed an optimization strategy of accompanying graph navigation for unmanned devices. The proposed strategy helps reduce the power consumption and CO2 gas emissions of unmanned devices. Aiming at minimizing the earliness and tardiness of a specified delivery time while minimizing the total completion time, Nishida et al. [31] developed a heuristic solution procedure to derive a conflict-free routing problem through which earliness and tardiness penalties are minimized while the total completion time of each task is also minimized under dynamic task arrivals.

Although extensive research has been conducted on path planning within IoV, only a limited number of studies have considered harnessing the powerful vehicle-road collaboration. Furthermore, most path planning schemes rely solely on real-time road conditions, making it challenging to adapt to future traffic congestion. Therefore, to further enhance path planning effectiveness and improve driver satisfaction, it is necessary to conduct future traffic flow predictions based path planning with vehicle-road collaboration.

III. SYSTEM MODEL AND PROBLEM DEFINITION

A. System Framework

The Vehicle-Road Cooperative System utilizes advanced wireless communication and IoV technologies to achieve dynamic real-time bidirectional interaction between vehicles and roads. Base stations (BSs) and RSUs utilize collected road information and employ Artificial Intelligence to provide intelligent services such as route planning to vehicles. Fig. 1 illustrates the framework of Vehicle-Road Cooperative System (VRCS). Table I enumerates the primary notations utilized in this paper.

In VRCS, intelligent vehicles equipped with wireless communication capabilities to communicate with infrastructure at the roadside. The RSUs are distributed along the sides of the road and communicate with vehicles within their coverage area through wireless links. BS is situated in an open area and is connected to multiple RSUs through wired links, facilitating real-time data exchange between them. Due to the specific geographical characteristics, the road network can be modeled as $G = \{V, E\}$. V is the set of road nodes, $V = \{v_1, \dots, v_i, \dots, v_N\}$, v_i represents the i_{th} node in G and N is the number of nodes. $E = \{e_1, \dots, e_j, \dots, e_M\}$ is the set of edges between nodes in G and M is the number of edges. Taking into account practical factors, i.e., the nodes within G are not necessarily interconnected, the introduction of the adjacency matrix A is necessary. $A \in \mathbb{R}^{N\times N}$ and the value of $A_{i,j}$ is based on the following equation that

$$A_{i,j} = \begin{cases} 1, & v_i \quad connects \quad to \quad v_j \\ 0, & otherwise \end{cases}. \tag{1}$$

The matrix A contains only 0s and 1s, making A an N-dimensional Boolean matrix. Due to the non-complete connectivity of nodes in G, it is evident that $M \leq N^2$.

TABLE I
SUMMARY OF NOTATIONS

Notation	Description
G	The road network
V	The set of road nodes
E	The set of edges in road network
N	The number of road nodes
S	The matrix recording speed information of roads
M	The number of edges
A	The adjacency matrix recording the road connectivity
$< o, p, d >$	The triplet represents the origin point, current location and destination
w_k	The k_{th} possible route between origin point and destination
s_i	The allowing speed of $v_i \in V$ at current
l_i	The length of $v_i \in V$
c_i	The time consumption for traveling through v_i at s_i
α_i^k	The variable that indicates whether v_i belongs w_k
ρ_k	The total travel time for w_k
σ_k	The total distance of the route w_k
η_i	The congestion level of v_i
γ_k	The overall congestion level w_k
$\mathcal{G}$	The spatial prediction function
St	The state space
Ac	The action space
Po	The state transition probability
Re	The reward function

B. Travel Consumption Model

Vehicles traveling on the road have information such as the origin, destination, and current location. Therefore, the travel information of a vehicle can be represented using a triplet $< o, p, d >$, where o is the origin point, p and d are the current location and destination, respectively. As shown in Fig. 1, during the process of traveling from the origin to the destination, there may be multiple routes. $W = \{w_1, \dots, w_k, \dots, w_K\}$ is adopted to represent the set of all possible routes, where w_k is a list containing the roads traversed on the k_{th} possible route. The objective of this scheme is to identify the most suitable route from numerous possibilities, taking into account factors such as total travel time and total travel distance.

Traffic speed is chosen as the traffic information for roads, as traffic congestion levels and other information related to

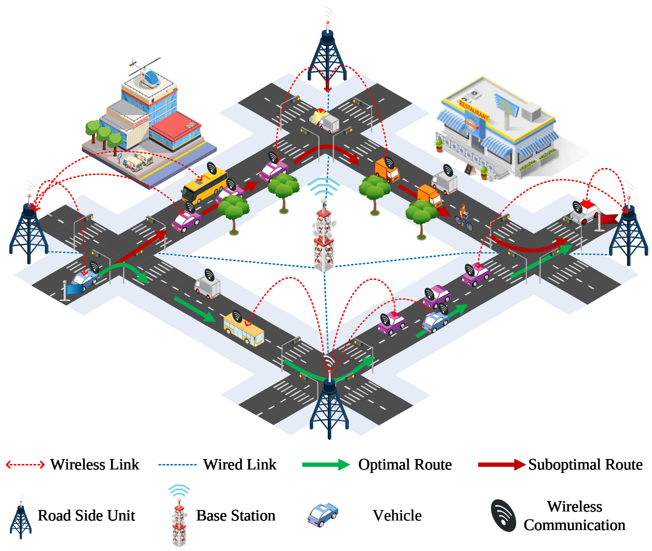


Fig. 1. The framework of vehicle road cooperation system in AIoV.

traffic flow can be inferred from traffic speed. The speed information of roads is recorded using the matrix $S \in \mathbb{R}^{N \times T}$, where N is the number of nodes and T is the time span of the recorded data. So, s_i is the speed of v_i during the current time period. [1] Since the planned routes will be executed in future time intervals, accurate prediction of future traffic speeds can be achieved by utilizing historical traffic speed data. This approach aims to enhance road planning effectiveness by relying on future traffic speed. The prediction process can be represented by the following equation that

$$s_i = f([s_i^{T-t}, \ldots, s_i^{T}]), \tag{2}$$

where f is the prediction function and will be detailed in section IV. The length of the roads is recorded in the set $L = \{l_1, \ldots, l_i, \ldots, l_N\}$, where l_i is the length of v_i. Therefore, the time consumption for traveling through v_i on route w_k can be calculated using equation (3) that

$$c_i = \frac{l_i}{s_i}. \tag{3}$$

Therefore, the total travel time for w_k can be calculated using the following equation that

$$\rho_k = \sum_{i=1}^{N} c_i \times \alpha_i^k, \tag{4}$$

where

$$\alpha_i^k = \begin{cases} 1, & v_i \in w_k \\ 0, & otherwise \end{cases}. \tag{5}$$

It is worth noting that unexpected events, such as traffic accidents, are not accounted for as additional waiting time for vehicles. Total distance of the route can be calculated using the following equation that

$$\sigma_k = \sum_{i=1}^{N} l_i \times \alpha_i^k. \tag{6}$$

[1]In this manuscript, according to the definition of the road network G, during time period t, "node speed" is equivalent to the traffic speed of the corresponding road segment, as well as "node length".

Another equally important factor is the congestion level of the road, which often significantly influences the driving experience of the driver. The congestion level of the v_i can be calculated based on the following equation that

$$\eta_i = 1 - \frac{s_i}{s_i^{'}}, \tag{7}$$

where $s_i^{'}$ is the theoretical speed of v_i. When there are no vehicles on the road, v_i is the maximum speed at which a vehicle can travel on that road. This equation signifies that as the actual speed s_i of v_i becomes smaller than the theoretical speed $s_i^{'}$, the congestion level η_i of v_i increases. The overall congestion level of w_k is determined by the following equation that

$$\gamma_k = \frac{\sum_{i=1}^{N} l_i \times \eta_i \times \alpha_i^k}{\sum_{i=1}^{N} l_i \times \alpha_i^k}. \tag{8}$$

This not only takes into consideration the congestion level of the road but also includes the length of the road as a weight in the calculation. This aligns more with reality, as roads with higher congestion levels and greater lengths tend to have a more significant impact on reducing the driving experience for the driver.

C. Problem Definition

After accurately modeling the three most critical factors in the path planning process, namely driving time, driving distance, and road congestion level, the objective of optimal path planning becomes evident. It is to select the path that minimizes the overall combination of these three factors from numerous possible routes. The comprehensive value of the three factors for route w_k can be calculated through a weighted sum, as outlined in the following equation that

$$\Gamma_k = w_1\rho_k + w_2\sigma_k + w_3\gamma_k, \tag{9}$$

where w_1, w_2, w_3 are adjustable parameters used to control the importance of different factors.

Hence, the formalized problem is given as

$$\min \Gamma_k, k \in [1, K], \tag{10}$$

$$s.t. \qquad |w_k| \leq N, \tag{11}$$

where $|w_k|$ representsthe number of roads included in route w_k and $|w_k| \leq N$ restricts that the number of roads in route w_k must be less than the total number of roads in G.

IV. Optimal Path Planning Scheme in AIoV

This section provides a detailed presentation of the proposed scheme RPFP to address the optimization problem modeled in section III. The section is divided into two parts. The first part focuses on precise prediction of traffic flow using spatio-temporal graph neural networks. And the second part involves optimal path planning with reinforcement learning.

A. Traffic Flow Prediction using STGNN

The efficacy of STGNN lies in its capacity to concurrently address interdependencies within spatial and temporal dimensions, rendering it well-suited for intricate traffic network configurations, thereby garnering considerable attention in the field of traffic flow prediction. The graph neural network integrated into STGNN adeptly captures the spatial intricacies of road networks by assimilating information from proximate nodes. Concurrently, the time series prediction module meticulously explores the temporal dynamics intrinsic to the model.

It can be inferred from section III that the traffic flow record embedding vectors of graph G is $S \in \mathbb{R}^{N\times T}$, so the embedding vectors of graph G at time t is $S_t = [s_1^t, \ldots, s_i^t, \ldots, s_N^t]$, as illustrated on the left side of Fig. 2. Utilizing GNNs enables the aggregation of flow information between different nodes through the spatial connectivity of the graph G. The specific process is illustrated as follows.

$$\begin{aligned} S_t^{'} &= \mathcal{G}(S_t) \\ &= \mathcal{G}([s_1^t, \ldots, s_i^t, \ldots, s_N^t]), \end{aligned} \tag{12}$$

where $\mathcal{G}$ is the objective function and is depicted as the spatial prediction module in Fig. 2. The more detailed process of information updating between nodes is defined by the following equations.

$$s_i^{t^{'}} = \sum_{j\in\varrho(i)} \mathcal{G}(s_i^t, e_{(i,j)}, s_j^t), \tag{13}$$

where $\varrho(i)$ is the set of all neighbor nodes of e_i and $e_{(i,j)}$ is the importance of the features of node e_j to e_i, indicating the connectivity between roads in the proposed traffic flow prediction scenario. It is noteworthy that $e_{(i,j)}$ is usually $\neq e_{(j,i)}$, which implies that the influence between node e_j and node e_i is asymmetrical, signifying a difference in their mutual impact on each other. This is due to certain roads being one-way streets, a real-world constraint that determines this characteristic.

To mitigate challenges such as gradient explosion or vanishing gradients frequently encountered in the training process of traditional sequential prediction networks like RNNs and LSTMs, this scheme creatively employs TCN as the means for temporal prediction. TCN utilizes convolutional operations instead of recursive operations to process input data and employs dilated causal convolutional layers to expand the receptive field. The specific prediction function can be referenced from the temporal prediction module in Fig. 2. Its mathematical form is as follows:

$$\widehat{S_t^{'}} = \mathcal{F}(w * S_{t-d}^{'} + b) \tag{14}$$

where $\mathcal{F}$ is the activation function, w is the convolutional kernel and b represents the bias term. To attain a comprehensive insight into the Dilated Causal Convolution operation, one is encouraged to consult the ensuing equation that

$$\widehat{S_t^{'}} = \sum_{i=0}^{k-1} w_i * S_{t-di}^{'}. \tag{15}$$

Here, k denotes the dimension of the convolutional kernel, while d signifies the dilation factor.

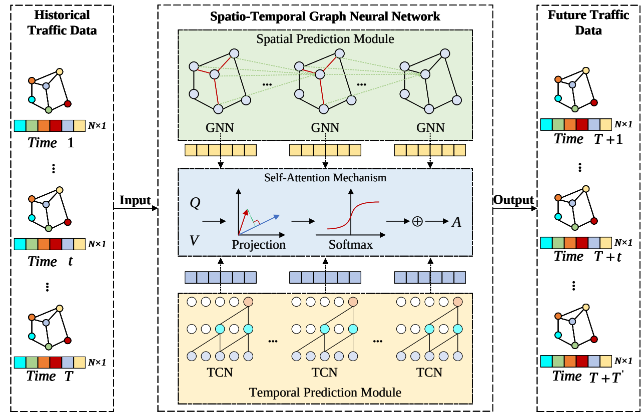


Fig. 2. Traffic flow prediction using STGNN integrated with TCN.

Utilizing the STGNN, a method for predicting traffic flow that concurrently addresses both spatial and temporal dimensions has been developed. The integration of TCN as the temporal prediction model effectively mitigates the challenge of vanishing gradients. For comprehensive details regarding the proposed scheme, kindly refer to Algorithm 1.

Algorithm 1: Traffic Flow Prediction using STGNN

Input: Historical traffic flow data S, road adjacency matrix A and the length of roads L.
Output: Traffic flow prediction result $\widehat{S_T^{'}} = [\widehat{s_1^{'}}, \ldots, \widehat{s_t^{'}}, \ldots, \widehat{s_T^{'}}]$

Construct an undirected graph $G = (V, E)$ based on A; **for** *each episode* **do**
    for $(i = 1; i <= N; i++)$ **do**
        for $e_j \in \varrho(i)$ **do**
            Computing the influence of node e_j on e_i with respect to l_j and l_i;
        end
        Aggregating spatial information from neighboring nodes by equation (13) that $s_i^{t^{'}} = \sum_{j\in\varrho(i)} \mathcal{G}(s_i^t, e_{(i,j)}, s_j^t)$;
        Leveraging the temporal dimension's information aggregation, as delineated by equation (15) that $\widehat{S_t^{'}} = \sum_{i=0}^{k-1} w_i * S_{t-di}^{'}$.
    end
    Obtain the prediction result $\widehat{S_t^{'}}$.
end
Return the traffic flow prediction results;

B. Congestion-Aware Path Planning in AIoV

This section furnishes an elaborate introduction to the path planning scheme, segmented into distinct sections: the definition of markov decision process (MDP), and the path planning scheme based on D3QN.

*1) **Definition of MDP**:* The primary aim of reinforcement learning algorithms is to iteratively train the agent through interactions with the environment, empowering it to autonomously make decisions that lead to the maximization of cumulative rewards. In the context of path planning, these algorithms engage in experiential learning by interacting with the environment, considering factors such as the current road distance to the destination and the fundamental conditions of roads connected to the current road. The ultimate objective is to cultivate the agent's capability to select paths characterized by shorter distances, reduced time requirements, and lower congestion levels as the planned route.

The Markov Decision Process (MDP) serves as a fundamental model widely utilized in reinforcement learning, with prevalent applications in disciplines like economics and operations research. Key components of an MDP consist of four essential elements: state space (St), action space (Ac), state transition probabilities (Po), and reward function (Re), succinctly represented as $M = (St, Ac, Po, Re)$. Subsequently, detailed explanations for the aforementioned four elements in the context of path planning will be presented.

- **State Space:** Throughout the vehicle's trajectory, the evolution of the state space is contingent upon alterations in the occupied road segment. Consequently, the definition of the state space is intricately linked to the current road segment. The precise delineation is articulated as follows:

$$St(i) = \{[e_i], [Cg(i)], [Ds(i)], [d]\}, \tag{16}$$

where e_i signifies the present road. $Cg(i)$ denotes the congestion level of the road linked to e_i, while roads not connected to e_i are regarded as having an infinite congestion level. $Ds(i)$ represents the length of the road associated with e_i, with roads not connected to e_i considered to have an infinite length. The variable d signifies the destination of the ongoing journey.

- **Action Space:** In the domain of path planning, the agent's actions align with the count of roads linked to the current road segment. Generating distinct numerical outputs signifies the road that the agent intends to navigate to in the subsequent step. Ac is used to represent the action space.
- **State Transition Probabilities:** The state transition probabilities pertains to the probability of transitioning from the current state $St(i)$ to the next state $St(i+1)$, contingent on the actions output by the agent. The intricate computational process will be elucidated through the following equations.

$$Pb = (St(i+1)|St(i), ac(i)), \tag{17}$$

where $ac(i)$ is the action.

- **Rewards Mechanism:** As the agent iteratively refines its path planning strategy, the output action at each step plays a pivotal role in reward calculation. Opting for a road segment that is shorter in distance, exhibits lower congestion, and brings the agent closer to the destination results in a positive increase in the reward. The specific calculation process is outlined as follows:

$$Re = \zeta * Ds(i) + \tau * Cg(i) + \chi. \tag{18}$$

In the above equation, parameters ζ and τ are employed to modulate the reward and can be adjusted according to distinct scenarios. Additionally, χ symbolizes the reward penalty, exhibiting a positive value when the subsequent road is closer to the destination and a negative value when it is farther away from the destination.

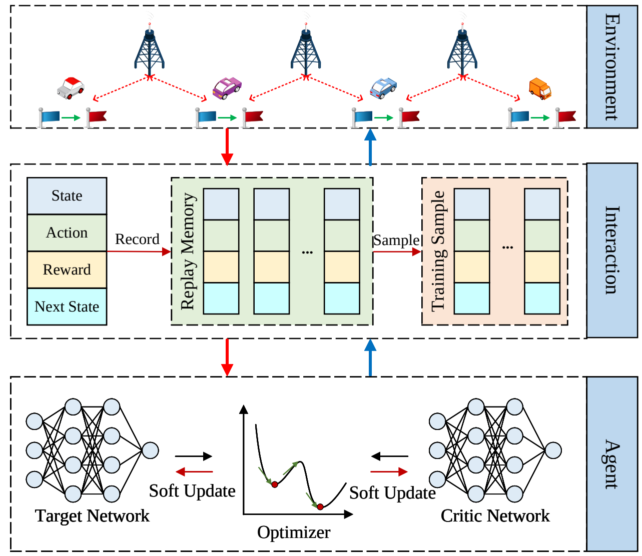


Fig. 3. Path planning scheme based on D3QN.

2) The Path Planning Scheme based on D3QN: DQN is a reinforcement learning algorithm adept at effectively tackling problems characterized by discrete action spaces, utilizing deep neural networks to navigate high-dimensional spaces. Nevertheless, the application of DQN often encounters the challenge of "maximization bias." In response to this, Double DQN was introduced, successfully alleviating the issue by modifying the calculation method for target values. Analogous to Double DQN, Dueling DQN represents an enhanced iteration of DQN. Through the introduction of the "dueling" architecture, Dueling DQN decomposes the Q-value function into two functions: the state value of the given state and the advantage values for each action. By enhancing the modeling of state values and action advantages, Dueling DQN notably enhances the performance and learning efficiency of DQN.

The scheme implemented herein leverages the Dueling Double Deep Q Network (D3QN), amalgamating key principles from both the Double DQN and Dueling DQN algorithms. This amalgamation is designed to augment learning stability and mitigate the challenge of overestimation, rendering it particularly efficacious in reinforcement learning tasks marked by heightened uncertainty and complexity. This rationale underscores the selection of D3QN for path planning in this scheme.

- **Environment of Path Planning:** As illustrated in Fig. 3, RSUs collect the environment data essential for reinforcement learning from the trajectories of vehicles within their designated coverage areas. These RSUs furnish the requisite tuples to the agent during the training phase. Initially, the agent initializes network parameters and an experience replay memory. Subsequently, guided by the environment, the agent performs actions, calculates rewards, and observes state transitions. The acquired training experiences are stored in the memory. Upon reaching its capacity limit, training samples are randomly extracted from the memory to train the two networks.

- **Training of Networks in D3QN:** Within the framework of D3QN, the training objective for the Critic network is to acquire an approximation of the state-action value function, endeavoring to align its output Q-values closely with the actual optimal Q-values. The loss function of the primary critic network is calculated by

$$L(\theta^C) = \frac{1}{N}\sum_{k=1}^{N}\left[Q(St_k(i), Ac_k(i); \theta^c) - y_k\right]^2,$$
$$Q(s,a;\theta) = V(s;\theta) + A(s,a;\theta) - \frac{1}{|\mathcal{A}|}\sum_{a'} A(s,a';\theta), \tag{19}$$

where θ^c is the parameters of critic network, N is the size of the sample batch.$Q(St_k(i), Ac_k(i); \theta^c)$ denotes the output of the Critic network, signifying the Q-value obtained by selecting action $Ac_k(i)$ in state $St_k(i)$. The target value is calculated by

$$y_k = r_k(i) + \gamma \max_{a'} Q(St_{k+1}(i), a'; \theta^T), \tag{20}$$

where a' represents any action that can be taken in the next state $St_{k+1}(i)$. And $r_k(i)$ is the reward in the current state, γ represents the discount factor.

- **Network Parameters Updating:** In D3QN, the Critic network is updated using the gradient descent method. Iteratively, the network parameters are adjusted based on the computed loss value, employing a predetermined learning rate until the model converges.

$$\theta^{C^*} \leftarrow \theta^C - \alpha\nabla_{\theta^C} L(\theta^C). \tag{21}$$

Equation (21) elucidates the process of updating the parameters of the critic network, with the parameters of the target network being copied from the critic network after several training steps.

Algorithm 2 shows the details of the path planning scheme in IoV.

The complexity of Algorithm 2 is $O(E \times T \times f(N))$, where E represents the number of episodes, T represents the number of time steps, and $f(N)$ denotes the forward propagation complexity of the neural network with N being the number of parameters. The above offers an intricate elucidation of the path planning scheme.

Algorithm 2: Path Planning based on D3QN

```
Input: Space of state St, Space of action Ac and the
       network of D3QN.
Output: Strategies for path planning
Initialize network parameters and the buffer memory;
Size of buffer memory π = 0;
for each episode do
    Reset the reinforcement learning environment;
    for t in episode do
        The critic network outputs the action;
        Compute the reward using equation (18);
        if π reaches the preset value then
            Sample from buffer memory for training;
            Compute y_k using
              y_k = r_k(i) + γ max_a' Q(St_{k+1}(i), a'; θ^T);
            Getting the loss as L(θ^C) =
              (1/N) Σ_{k=1}^{N} [Q(St_k(i), Ac_k(i); θ^c) − y_k]^2;
            Updating the network parameters with
              θ^{C*} ← θ^C − α∇_{θ^C} L(θ^C);
            Updating θ^T after several steps with
              θ^T ← θ^C.
        end
    end
end
Return θ^T, θ^C and the network of D3QN;
Output the strategies for path planning.
```

V. EXPERIMENT AND ANALYSIS

Experimental setup and comparative experiments are presented. The efficacy and superiority of the proposed approach are inferred through the analysis of experimental results.

A. Experimental Setup

Two datasets are employed for the evaluation of traffic flow prediction. Without loss of generality, the dataset utilized is the traffic flow dataset.

1) SZ-taxi. This dataset was collected from taxi trajectories in Shenzhen over a period of 31 days, from Jan.1 to Jan.31, 2015. The data comprises traffic flow information collected at fixed time intervals for roads within the region. The included adjacency matrix delineates the connectivity between roads.
2) Los-loop. This dataset is obtained from detection sensors on Los Angeles highways, covering the period from Mar.1 to Mar.7, 2012. Similar to the SZ-taxi dataset, it also incorporates an adjacency matrix illustrating the connectivity between roads.

In the experiments, 70% of each dataset is utilized as the training dataset, with the remaining 30% serving as the validation dataset. Experiments are implemented in python 3.10, Pytorch 2.0 and conducted on the server with Intel i7-8700 CPU (3.2 GHz, 6 cores), 32-GB DRAM, and NVIDIA RTX 3090 GPU.

B. Analysis of the Prediction Performance

To precisely capture the spatial and temporal characteristics of road traffic flow, this approach employs STGNN-TCN. Furthermore, to assess the effectiveness and superiority of the proposed method in traffic flow prediction, four distinct schemes have been implemented. The particulars of each comparative approach are outlined as follows.

- Traffic Flow Prediction with TCN (TCN)
  The scheme utilizes TCN as the traffic flow prediction model, employing one-dimensional convolutional layers with dilated kernels to extract temporal features from time-series data [32]. This configuration enables accurate prediction of traffic flow.

- Traffic Flow Prediction with GRU (GRU)
  GRU, as an enhanced iteration of RNN, adeptly tackles challenges such as gradient explosion or vanishing gradients [33]. Its proficiency in long-term time series prediction renders it suitable for applications like traffic flow prediction.
- Traffic Flow Prediction with GNN (GNN)
  GNN exhibits proficiency in handling data with a graph structure, enabling exploration of mutual influences between roads within complex road networks. This capability is particularly well-suited for scenarios involving traffic flow prediction.
- Traffic Flow Prediction with RNN integrating TCN (RNN-TCN)
  RNN, a fundamental neural network architecture, is commonly employed for inferring and predicting time series data. Through integration with TCN, RNN can achieve enhanced training efficiency and improved performance in traffic flow prediction.
- Traffic Flow Prediction with STGNN incorporating GRU (STGNN-GRU)
  STGNN excels at concurrently capturing features from both spatial and temporal dimensions, rendering it particularly suitable for intricate road scenarios with spatial distribution characteristics. In the STGNN employed in this scheme, GRU serves as the temporal prediction model.

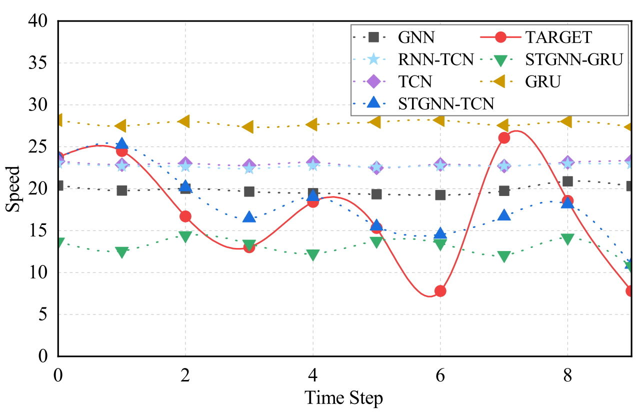


Fig. 4. Traffic flow prediction with different schemes.

Fig. 4 portrays the predictive performance of various schemes over ten time intervals, with the red solid line denoting the original traffic flow data. The graph clearly indicates that STGNN-TCN demonstrates superior predictive performance, closely aligning with the trends and values of the original data. STGNN-GRU exhibits some adherence to the trends, but it deviates notably from the actual data, with only a few data points matching. Conversely, methods focusing solely on a single dimension face challenges in accurately capturing the evolving patterns in traffic flow.

To further illustrate the performance of different schemes in traffic flow prediction tasks, we leverage more intuitive data and employ various evaluation metrics. The specific details are outlined below.

- Mean Absolute Error (MAE)
  MAE is the average of the absolute errors, providing a better reflection of the actual prediction value errors. MAE is calculated as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|. \tag{22}$$

- Mean Absolute Percentage Error (MAPE)
  MAPE represents an enhancement over MAE, as it computes the percentage error between actual and predicted values. This mitigates the influence of data range, providing a more robust evaluation metric. The calculation equation is shown as

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|. \tag{23}$$

- Root Mean Square Error (RMSE)
  RMSE quantifies the average deviation between predicted and actual values, sharing the same unit as the target values. The calculation equation of RMSE is that

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}. \tag{24}$$

The performance of the mentioned schemes under various evaluation metrics is presented in the table II. Observing Table II, it is evident that across various metrics, STGNN-TCN demonstrates the best performance, followed by STGNN-GRU. A preliminary analysis attributes this trend to the consideration of features in both temporal and spatial dimensions exclusively in these two schemes among the five experimental setups. The other three schemes, focusing solely on features in a single dimension, exhibit comparatively poorer performance, ranked from best to worst as TCN, RNN-TCN, GCN, and GRU. Furthermore, through joint validation on two datasets, the proposed traffic flow prediction scheme consistently attains optimal results.

C. Analysis of the Path Planning Performance

To attain superior path planning results, we propose a path planning strategy grounded in the reinforcement learning algorithm D3QN. In an effort to identify the optimal reinforcement learning algorithm for discrete action problems and to compare the effectiveness of the reinforcement learning-based path planning strategy with other approaches, we further implement the following strategies.

- Path Planning Scheme Based on DQN (DQN)
  This scheme is rooted in the classical reinforcement learning algorithm DQN, which amalgamates Q-learning with deep neural networks, rendering it effective for addressing discrete action problems in deep reinforcement learning.
- Path Planning Scheme Based on DDQN (DDQN)
  Double DQN (DDQN) represents an improvement over DQN, employing two Q networks: one for action selection and the other for estimating the value of that action [34]. This architecture is designed to address the overestimation problem.

TABLE II
THE DIFFERENT METRIC VALUES OF DIFFERENT SCHEMES

	SZ-taxi			**Los-loop**		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE
GRU	5.02	0.24	6.51	5.89	0.11	8.66
GCN	5.39	0.27	6.82	5.47	0.10	7.84
RNN-TCN	3.92	0.18	5.42	5.87	0.11	8.67
TCN	3.87	0.18	5.32	5.81	0.10	7.55
STGNN-GRU	3.57	0.16	4.98	5.25	0.10	7.76
STGNN-TCN	**3.04**	**0.13**	**4.52**	**4.85**	**0.09**	**6.59**

- Path Planning Scheme Based on Dijkstra (Dijkstra)
  Dijkstra is a greedy search algorithm that iteratively expands the most promising path at each step, ensuring optimality in finding the shortest path within a weighted graph.
- Path Planning Scheme Based on ACO (ACO)
  Swarm intelligence, inspired by the foraging behavior of ants, utilizes a probabilistic approach to simultaneously explore multiple paths. However, it does not guarantee the discovery of the absolute optimal path; instead, it can only find statistically good solutions.

In reinforcement learning, the intricacies of the environment present challenges for a single-layer network in the path planning scheme. Therefore, a critical initial step involves determining the appropriate number of network layers. Fig. 5 below illustrates the variations in rewards and variance for D3QN with different network layer configurations in a complex environment.

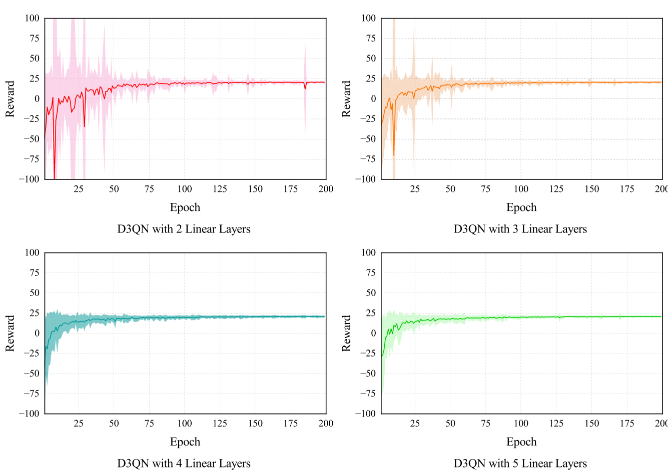


Fig. 5. The rewards and variances of D3QN with different network layers.

From the Fig. 5, it is evident that when the network has only two layers, D3QN obtains rewards with significant oscillations and high variance. As the number of layers increases to three, the oscillation amplitude of the rewards decreases, and the variance reduces, but the ideal performance is not yet achieved. D3QN with four and five layers shows better performance in terms of reward trends and variance, with little difference between them. Considering that a higher number of network layers would incur more time and resource consumption during training, after analysis, it is determined that D3QN requires a four-layer network for executing the path planning task.

After determining the appropriate number of network layers, the subsequent step involves comparing reinforcement learning algorithms. Three distinct reinforcement learning algorithms are utilized with the same number of network layers and training rounds. From Fig. 6, it is evident that under three different routes (30, 60, 90), the average rewards achieved by D3QN are higher than those of DQN and DDQN. This suggests that D3QN is the most suitable for executing path planning tasks. As shown in Fig. 7, with an increase in the number of routes, the variance of all three strategies gradually decreases. Moreover, D3QN exhibits the smallest reward variance, indicating that its performance is more stable. D3QN demonstrates a reduction in variance by 36% compared to DQN and 8.4% compared to DDQN. Considering both Fig. 6 and 7, D3QN outperforms the other two reinforcement learning algorithms across the three different routes.

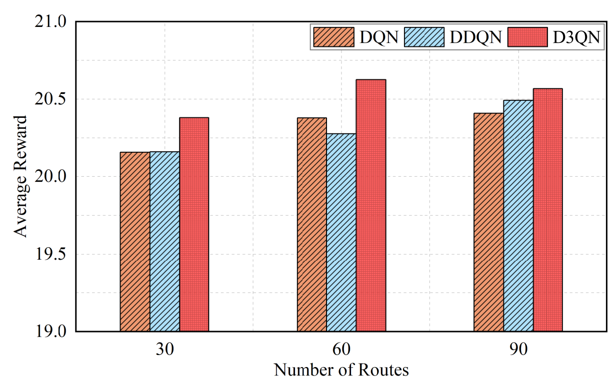


Fig. 6. The reward of DQN, DDQN, D3QN under different routes.

We conducted a comparison between Dijkstra and ACO with the path planning scheme proposed in this paper, utilizing criteria such as the distance of the path and the time taken to traverse it. Five pairs of start and destination points were randomly selected as test paths. As depicted in Fig. 8, D3QN outperforms in terms of planning path distance, selecting the shortest path. ACO performs well, while Dijkstra shows the poorest performance. Concerning distance, D3QN demonstrates reductions of 12.6% and 4.63% in comparison

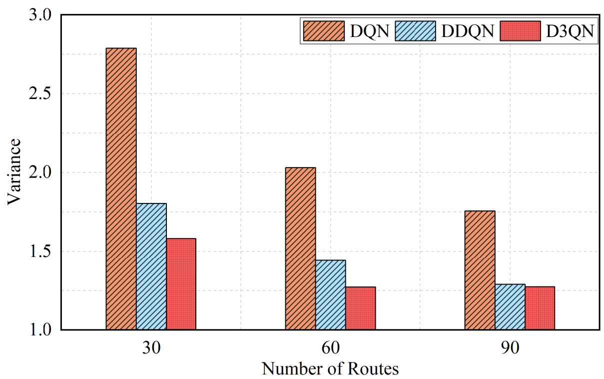


Fig. 7. The variance of DQN, DDQN, D3QN under different routes.

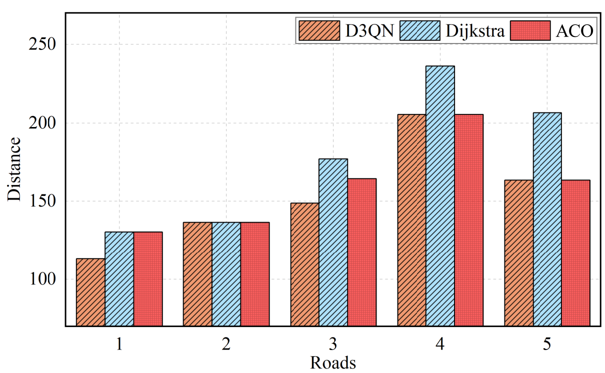


Fig. 8. The distance of planned path under D3QN, Dijkstra, ACO.

to Dijkstra and ACO, respectively. The analysis concludes that the poor performance of Dijkstra is attributed to its tendency to reach local optimal solutions in large-scale maps or long path planning scenarios.

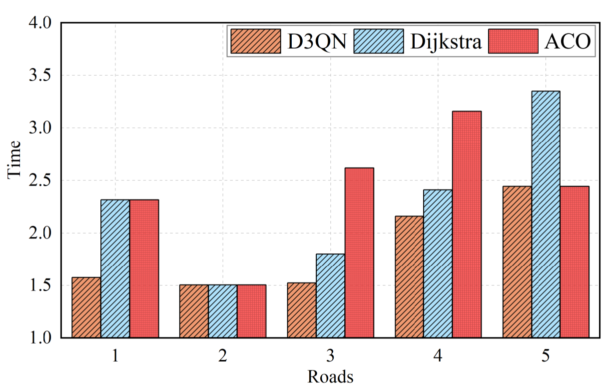


Fig. 9. The time consumption of planned path under D3QN, Dijkstra, ACO.

Fig. 9 illustrates the time consumption for five test paths under three different path planning schemes. It can be observed that in most cases, D3QN performs well, ACO occasionally matches the time spent by D3QN, and Dijkstra remains the least efficient. Concerning travel time, D3QN demonstrates reductions of 16.9% and 21% in comparison to Dijkstra and ACO, respectively.

VI. Conclusion

In addressing the challenges of achieving high-quality path planning in complex road networks and enhancing the driving experience, a vehicle-road cooperation framework is proposed in AIoV. Additionally, a congestion-aware path planning scheme, named RPFP, has been formulated. RPFP utilizes STGNN-TCN for efficient and accurate traffic flow prediction, serving as the decision foundation for path planning. The path planning algorithm based on D3QN can achieve targeted path planning in complex decision environments. In future work, we will focus on the collaborative control strategy of signal lights in traffic flow optimization in coordination with the present study.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant (No. 62372242 and 92267104), and in part by Natural Science Foundation of Jiangsu Province of China under Grant (No. BK20211284). This work is also supported by Postgraduate Research & Practice Innovation Program of Jiangsu Province (No. KYCX24_1532).

References

[1] R. Mishra, H. P. Gupta, R. Kumar, and T. Dutta, "Leveraging augmented intelligence of things to enhance lifetime of uav-enabled aerial networks," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 586–593, 2022.

[2] D. Javeed, T. Gao, M. S. Saeed, and M. T. Khan, "Fog-empowered augmented intelligence-based proactive defensive mechanism for iot-enabled smart industries," *IEEE Internet of Things Journal*, 2023.

[3] X. Zhou, M. Bilal, R. Dou, J. J. Rodrigues, Q. Zhao, J. Dai, and X. Xu, "Edge computation offloading with content caching in 6g-enabled iov," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[4] P. Arthurs, L. Gillam, P. Krause, N. Wang, K. Halder, and A. Mouzakitis, "A taxonomy and survey of edge cloud computing for intelligent transportation systems and connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[5] L. Xu, X. Zhou, M. A. Khan, X. Li, V. G. Menon, and X. Yu, "Communication quality prediction for internet of vehicle (iov) networks: an elman approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 19 644–19 654, 2021.

[6] J. Zhang, J. Pu, J. Xue, M. Yang, X. Xu, X. Wang, and F.-Y. Wang, "Hivegpt: human-machine-augmented intelligent vehicles with generative pre-trained transformer," *IEEE Transactions on Intelligent Vehicles*, 2023.

[7] S. Aggarwal and N. Kumar, "Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges," *Computer Communications*, vol. 149, pp. 270–299, 2020.

[8] C. Lin, G. Han, J. Du, T. Xu, L. Shu, and Z. Lv, "Spatiotemporal congestion-aware path planning toward intelligent transportation systems in software-defined smart city iot," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8012–8024, 2020.

[9] P. W. Shaikh, M. El-Abd, M. Khanafer, and K. Gao, "A review on swarm intelligence and evolutionary algorithms for solving the traffic signal control problem," *IEEE transactions on intelligent transportation systems*, vol. 23, no. 1, pp. 48–63, 2020.

[10] W. Wang, T. Qie, C. Yang, W. Liu, C. Xiang, and K. Huang, "An intelligent lane-changing behavior prediction and decision-making strategy for an autonomous vehicle," *IEEE transactions on industrial electronics*, vol. 69, no. 3, pp. 2927–2937, 2021.

[11] V.-L. Nguyen, R.-H. Hwang, and P.-C. Lin, "Controllable path planning and traffic scheduling for emergency services in the internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12 399–12 413, 2021.

[12] X. Wang, Y. Ma, Y. Wang, W. Jin, X. Wang, J. Tang, C. Jia, and J. Yu, "Traffic flow prediction via spatial temporal graph neural network," in *Proceedings of the web conference 2020*, 2020, pp. 1082–1092.

[13] D. A. Tedjopurnomo, Z. Bao, B. Zheng, F. M. Choudhury, and A. K. Qin, "A survey on modern deep neural network for traffic prediction: Trends, methods and challenges," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 4, pp. 1544–1561, 2020.

[14] F. Li, J. Feng, H. Yan, G. Jin, F. Yang, F. Sun, D. Jin, and Y. Li, "Dynamic graph convolutional recurrent network for traffic prediction: Benchmark and solution," *ACM Transactions on Knowledge Discovery from Data*, vol. 17, no. 1, pp. 1–21, 2023.

[15] C. Liu, L. Wu, W. Xiao, G. Li, D. Xu, J. Guo, and W. Li, "An improved heuristic mechanism ant colony optimization algorithm for solving path planning," *Knowledge-Based Systems*, vol. 271, p. 110540, 2023.

[16] Y. V. Pehlivanoglu and P. Pehlivanoglu, "An enhanced genetic algorithm for path planning of autonomous uav in target coverage problems," *Applied Soft Computing*, vol. 112, p. 107796, 2021.

[17] Z. Yu, Z. Si, X. Li, D. Wang, and H. Song, "A novel hybrid particle swarm optimization algorithm for path planning of uavs," *IEEE Internet of Things Journal*, vol. 9, no. 22, pp. 22 547–22 558, 2022.

[18] L. Ding, D. Zhao, Z. Wang, G. Wang, C. Tan, L. Fan, and H. Ma, "Learning to help emergency vehicles arrive faster: A cooperative vehicle-road scheduling approach," *IEEE Transactions on Mobile Computing*, 2022.

[19] C. Chen, G. Yao, L. Liu, Q. Pei, H. Song, and S. Dustdar, "A cooperative vehicle-infrastructure system for road hazards detection with edge intelligence," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[20] Z. Cui, Z. Lu, L. T. Yang, J. Yu, L. Chi, Y. Xiao, and S. Zhang, "Privacy and accuracy for cloud-fog-edge collaborative driver-vehicle-road relation graphs," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[21] S. Cao, D. Liu, C. Dai, C. Wang, Y. Yang, W. Zhang, and D. Zheng, "Reinforcement learning based tasks offloading in vehicular edge computing networks," *Computer Networks*, p. 109894, 2023.

[22] C. Ma, G. Dai, and J. Zhou, "Short-term traffic flow prediction for urban road sections based on time series analysis and lstm_bilstm method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5615–5624, 2021.

[23] W. Li, X. Wang, Y. Zhang, and Q. Wu, "Traffic flow prediction over muti-sensor data correlation with graph convolution network," *Neurocomputing*, vol. 427, pp. 50–63, 2021.

[24] A. Ali, Y. Zhu, and M. Zakarya, "Exploiting dynamic spatio-temporal graph convolutional neural networks for citywide traffic flows prediction," *Neural networks*, vol. 145, pp. 233–247, 2022.

[25] F. Wang, G. Li, Y. Wang, W. Rafique, M. R. Khosravi, G. Liu, Y. Liu, and L. Qi, "Privacy-aware traffic flow prediction based on multi-party sensor data with zero trust in smart city," *ACM Transactions on Internet Technology*, vol. 23, no. 3, pp. 1–19, 2023.

[26] S. Teng, X. Hu, P. Deng, B. Li, Y. Li, Y. Ai, D. Yang, L. Li, Z. Xuanyuan, F. Zhu, and L. Chen, "Motion planning for autonomous driving: The state of the art and future perspectives," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 6, pp. 3692–3711, 2023.

[27] J. Li, C. Yu, Z. Shen, Z. Su, and W. Ma, "A survey on urban traffic control under mixed traffic environment with connected automated vehicles," *Transportation research part C: emerging technologies*, vol. 154, p. 104258, 2023.

[28] K. Lin, C. Li, Y. Li, C. Savaglio, and G. Fortino, "Distributed learning for vehicle routing decision in software defined internet of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3730–3741, 2020.

[29] Y. Liang, Y. Liu, and B. B. Gupta, "Pprp: preserving-privacy route planning scheme in vanets," *ACM Transactions on Internet Technology*, vol. 22, no. 4, pp. 1–18, 2022.

[30] K. Cai, T. Dai, Q. Lin, X. Song, W. Zhou, Q. Zhou, J. Wei, F. Wang, H. Chen, and B. Li, "Route planning based on parallel optimization in the air-ground integrated network," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[31] K. Nishida and T. Nishi, "Dynamic optimization of conflict-free routing of automated guided vehicles for just-in-time delivery," *IEEE Transactions on Automation Science and Engineering*, 2022.

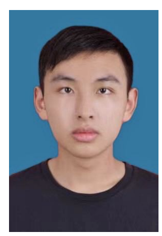
[32] Y. Ren, D. Zhao, D. Luo, H. Ma, and P. Duan, "Global-local temporal convolutional network for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 2, pp. 1578–1584, 2020.

[33] W. Shu, K. Cai, and N. N. Xiong, "A short-term traffic flow prediction model based on an improved gate recurrent unit neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 16 654–16 665, 2021.

[34] Z. Chu, F. Wang, T. Lei, and C. Luo, "Path planning based on deep reinforcement learning for autonomous underwater vehicles under ocean current disturbance," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 108–120, 2022.

Guoqiang Liu received the B.S. degree in software engineering from the Nanjing University of Information Science and Technology in 2022. He is currently pursuing the master's degree in software engineering. His research interests include mobile edge computing, Internet of Things, and deep learning.

Muhammad Bilal received the Ph.D. degree in information and communication network engineering from the School of Electronics and Telecommunications Research Institute (ETRI), Korea University of Science and Technology, Daejeon, South Korea, in 2017. From 2018 to 2023, he was an Associate Professor with the Division of Computer and Electronic Systems Engineering, Hankuk University of Foreign Studies, Yongin, South Korea. In 2023, he joined Lancaster University, Lancaster LA1 4YW, United Kingdom, where he is currently working as a Senior Lecturer with the School of Computing and Communications. His Network Optimization, Cyber Security, the Internet of Things, Vehicular Networks, Information-Centric Networking, Digital Twin, Artificial Intelligence, and Cloud/Fog Computing.



Qingyang Wang is currently pursuing the bachelor's degree in data science and big data technology at School of Reading of Nanjing University of Information Science and Technology in China. His research interests include deep learning and edge computing.

Xiaolong Xu received the Ph.D. degree in computer science and technology from Nanjing University, China, in 2016. He is currently a Full Professor with the School of Software, Nanjing University of Information Science and Technology. He has published more than 100 peer-review articles in international journals and conferences, including the IEEE TKDE, IEEE TPDS, JSAC, IEEE TSC, IEEE TFS, IEEE T-ITS, IJCAI, ICDM, ICWS, ICSOC, etc. He was selected as the Highly Cited Researcher of Clarivate (2021-2023). He received best paper awards from Tsinghua Science and Technology at 2023, Journal of Network and Computer Applications at 2022, and several conferences, including IEEE HPCC 2023, IEEE ISPA 2022, IEEE CyberSciTech 2021, IEEE CPSCom2020, etc. His research interests include edge computing, the Internet of Things (IoT), cloud computing, and big data.