Multi-Agent Reinforcement Learning-based Signal Planning for Resisting Congestion Attack in Green Transportation

Yike Li, Wenjia Niu, Yunzhe Tian, Tong Chen, Zhiqiang Xie, Yalun Wu, Yingxiao Xiang, Endong Tong, Thar Baker, *Senior Member, IEEE* and Jiqiang Liu, *Senior Member, IEEE*

***Abstract*—Inefficient signal control will not only exaggerate traffic congestion, but also increase the fuel consumption and exhaust emissions. Thus, signal planning is highly important in green transportation. As the Connected vehicle (CV) technology has transformed today's transportation systems by connecting vehicles and the transportation infrastructure through wireless communication, the CV-based signal control system has seen significant studies recently. Unfortunately, existing signal planning algorithms in use are developed for the signal-intersection, showing low traffic efficiency in the multi-intersection collaborative planning due to ignoring the traffic correlation among the neighboring intersections. In this work, we target the USDOT (U.S. Department of Transportation) sponsored CV-based traffic control system, and implement a multi-intersection traffic network. We model the multi-intersection collaborative signal planning problem as a multi-agent reinforcement learning problem, and present an actor-attention-critic algorithm to improve transportation efficiency and energy efficiency in green transportation, as well as resist congestion attack. Experiment results on the multi-intersection traffic network indicates that 1) compared to the baseline, our approach reduces the total delay by as high as 44.24%; 2) our method transports more vehicles passing the intersections meanwhile reduces the total CO_2 emissions by 2.40%; 3) under the congestion attack, our approach shows robustness and reduces the total delay by as high as 64.33%.**

***Index Terms*—Green transportation, Signal plan, CV-based system, Multi-agent reinforcement learning.**

I. Introduction

RECENTLY, as urbanization has dramatically increased the traffic demand in cities globally, it also causes a sharp increase of traffic congestion, energy consumption and environment pollution. Improving transportation efficiency and meanwhile reducing carbon emissions and fuel consumption has been significant studies for traffic control methods. Connected vehicle (CV) technology transforms today's transportation system by connecting vehicles and the transportation infrastructure through wireless communication. In these CV-based transportation systems, the strong connectivity of CV-based systems greatly improves the mobility, safety and public agency operations [1], [2]. Meanwhile, by increasing traffic efficiency, it also made contributions in sustainable development and environmental protection.

I-SIG (Intelligent Traffic Signal System) [3], a CV-based transportation sponsored by the USDOT (U.S. Department of Transportation) [4], performs one of the most basic urban traffic operations, traffic signal control. I-SIG operates based on wireless communications to connect vehicles (On-Board Units (OBUs)) and infrastructure (Roadside Units (RSUs)). At the beginning of each signal cycle, vehicles use OBUs to broadcast Basic Safety Messages (BSM) including its real-time trajectory data, e.g., location and speed, to the surrounding vehicles and infrastructure. Then the RSU of the intersection sends the vehicle trajectories into the signal planning model, and the signal planning model generates signal plans for the coming signal cycle. The I-SIG has been tested on real road intersection and has shown to achieve a 26.6% reduction in total vehicle delay [5].

However, the signal planning in the I-SIG system, for example dual-ring version of the COP (Controlled Optimization of Phases) algorithm [6], only taking the traffic of the intersection itself into computation, ignoring the traffic correlation among the neighboring intersections. Thus, when the traffic increases sharply or malicious attack (e.g., congestion attack on the CV-based traffic signal control [7]) occurred, causing massive traffic congestions, the single-intersection signal planning will take a long time to disperse the heavy traffic flow.

We implement simulation in a 4-intersection signal control area using the conventional COP algorithm, and analyze the variation of traffic conditions within a period. As shown in Fig. 1, we found that, even a sudden change of traffic flow at time t could cause the congestion of the intersection and thus spread to the neighbor intersections. In this case, the way for conventional signal plan algorithms to dissipate the congestion is inefficiency, which calls for 59 minutes on average in our simulation.

For green transportation, the sustainability is often defined as the total energy consumption or emissions of vehicles passing the intersection [8]. However, when traffic congestion

The work was supported by the National Natural Science Foundation of China under Grant Nos. 61972025, 61802389, 61672092, U1811264, and 61966009, the National Key R&D Program of China under Grant Nos. 2020YFB1005604 and 2020YFB2103802.(*Corresponding authors: Endong Tong; Wenjia Niu.*)

Yike Li, Wenjia Niu, Yunzhe Tian, Tong Chen, Zhiqiang Xie, Yalun Wu, Yingxiao Xiang, Endong Tong and Jiqiang Liu are with the Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University, Beijing 100044, China (e-mail: yikeli@bjtu.edu.cn; niuwj@bjtu.edu.cn; tianyunzhe@bjtu.edu.cn; tongchen@bjtu.edu.cn; zhiqiangxie@bjtu.edu.cn; wuyalun@bjtu.edu.cn; yxxiang@bjtu.edu.cn; edtong@bjtu.edu.cn; jqliu@bjtu.edu.cn).

Thar Baker is with the Department of Computer Science, College of Computing and Informatics, University of Sharjah, Sharjah, UAE (e-mail: tshamsa@sharjah.ac.ae).

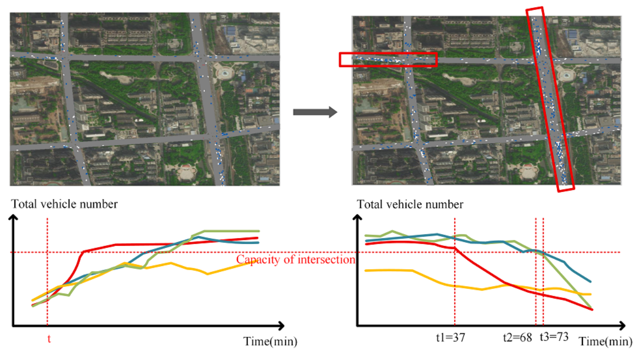


Fig. 1: The shortcoming of single-intersection signal planning in a multi-intersection scenario.

occurs in an intersection, the travel time for vehicles will be much longer, which cause a waste of both time and energy. What's more, vehicles across the traffic congestion are always in the idle state, which make the total energy consumption and exhaust emissions higher than the flowing state. As the traffic congestion increases the delay time of vehicles, adds social costs and causes environmental pollution, it is a problem that needs to be solved by the signal planning methods.

Researches have applied the reinforcement learning (RL) technique [9]–[12] to signal control problem. Recent studies [13] have applied deep reinforcement learning techniques, such as Deep Q-learning (DQN), for traffic light control problem. However, traditional reinforcement learning is difficult to apply due to the following challenges: (1) how to represent environment; (2) how to model the correlation between environment and decision; and (3) how to compatibility with multi-intersection environments. Wei et.al [14] proposed a deep reinforcement learning model for traffic light control, based on surveillance cameras. However, as traditional sensors such as cameras have poor performance in bad weather and poor lighting, the CV-based RL model leads to more superior and stable performance. For green transportation, there are Deep Learning (DL) methods used in intelligent transportation for smart cities [15]. However, few of the above RL-based methods focus on both of the traffic and efficiency and the environmental problems in green transportation systems.

To tackle the above environmental problem of green transportation, we define energy efficiency to evaluate the performance in environmental protection in green transportation respectively. Higher energy efficiency means to transport more vehicles passing the intersection in unit energy consumption, thus to reduce the environmental problems caused by traffic congestion. In the CV environment, it is easy to get global vehicle flow data, which makes multi-intersection signal planning a whole thing. An important insight brought by this work is to model the signal planning of multiple intersections in a region as a multi-agent reinforcement learning (MARL) problem [16], [17]. In the MARL, signal control in one intersection is controlled by an agent in RL model, each agent aims to optimize its local signal plan to minimize the delay time of vehicles, and thus to get the highest reward. Multiple agents are in cooperative relationships, we further introduce the attention mechanism to the RL model, which in a manner similar to a differentiable key-value memory model.

We evaluate our MARL-based method with random traffic flow scales with the simulator PTV VISSIM [18], a commercial-grade traffic simulation software, which is used to generate the traffic data as input of the RL model. Under random traffic flow scales, we evaluate the performance of the COP and our method in two aspects: the traffic efficiency and the energy consumption. For the traffic efficiency, we found that our method (1) reduced the total delay by 44.24% in common traffic conditions; (2) is robust to the congestion attacks, with the total delay reduced 64.33%. For the energy consumption, our method (1) transported more vehicles meanwhile reduced the total CO_2 emission by 2.40%; (2) kept the fuel consumption at the same level. The result shows that our method achieves higher effectiveness and robustness than the conventional COP algorithm.

Our major contributions are highlighted as follows:

- We make the first attempt to explore the multi-intersection signal planning in CV-based transportation system to improve traffic efficiency and energy efficiency as well as resist the congestion attack, by modeling the collaborative planning as a multi-agent reinforcement learning problem with shared rewards.
- We present an actor-critic based multi-agent reinforcement learning algorithm that trains decentralized policy for each agent, in which a centralized attention critic is used to dynamically select agents should be focused on for each agent at every time step.
- Comprehensive experiments on different traffic networks demonstrate the effectiveness of our MARL-based signal planning method in traffic efficiency and energy efficiency, reducing the total delay and the total CO_2 emission by 44.24% and 2.40%, respectively. Besides, under congestion attack, our method shows great robustness and reduces the total delay by as high as 64.33%.

The rest of the paper is structured as follows: Section II introduces necessary background. Section III proposes a multi-agent reinforcement learning-based signal planning framework. Section IV reports our experiments in traffic networks containing different number of intersections and evaluations on the traffic efficiency and the energy consumption metrics. In Section V, we discuss the related works. Finally, Section VI concludes the work of this paper and discusses works in the future.

II. Backgrounds

A. Signal Control Concepts

The signal control scenario in a single intersection is shown in Fig. 2. There are 8 traffic signals in each intersection, called phases. Each phase is configured with the green light lasting time t_g, the yellow light lasting time t_y, and the clearance red light lasting time t_r in the signal control. The Signal control is performed by setting t_g of the phase sequence, which is calculated by the signal planning algorithm. The output of the signal planning algorithm is called a signal plan, including (1) the sequence of phases, and (2) t_g for each phase, as shown in Fig. 3. Number 1 to 8 are phases, and the green, yellow, and clearance red light periods for each phase are filled with the corresponding colors.

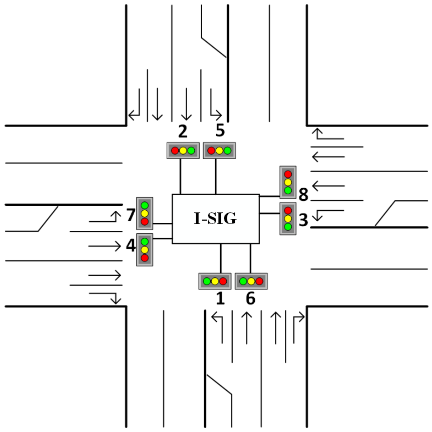


Fig. 2: Signal control scenario in single intersection.

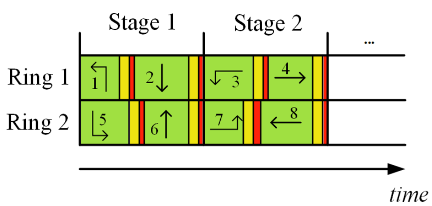


Fig. 3: Illustration of a signal plan for 8 phases.

The COP Algorithm is the core signal planning algorithm used in the I-SIG system. The goal of the COP algorithm is to reduce the total delay for all vehicles in the intersection. The input of the COP algorithm is each approaching vehicle's estimated arrival time at the intersection, which is defined as the estimated remaining time for a vehicle to reach the stop bar of its current lane. Based on the arrival time, COP uses dynamic programming to calculate an optimal signal plan for 8 phases with the least estimated total delay. After each stage of signal control, the following signal status is returned as feedback for continuous COP planning.

The use of COP is chosen by the I-SIG designer, the team of USDOT-selected signal control experts, based on a 2015 paper published in Transportation Research Part C [19]. The COP algorithm is chosen because it is very suitable for the CV environment: its input is the arrival time for individual vehicles instead of aggregated traffic information, and thus can best leverage the per-vehicle trajectory data in the CV environment to effectively handle traffic dynamics. However, Chen et. al [7] has shown that the limitation of planning stage in COP unexpectedly leaves the I-SIG system vulnerable to congestion attacks.

B. Green Transportation

Green transportation is defined as "the transportation service with a fewer negative impact on human health and the environment compared to existing transportation services" [20]. In this context, green transportation has three major dimensions: environment, economy, and society. From an environmental point of view, green transportation has not adverse impact on the environment. From an economic perspective, as transportation is the main factor for growth, development, and employment, green transportation must be energy efficiently and at a price that is friendly to the community. Finally, green transportation must be safe for the community and not harm public health [21].

However, the current public transportation service is considered unsustainable due to the extensive use of the fossil fuels driven transportation system, including private vehicles. All its emissions are responsible for environmental pollution, in which CO_2 emission constitutes about 65% of the total emission [22]. Thus, the green transportation has been extensively studied by researches in recent years. For instance, the electrification of the vehicle is now picking up speed, aiming to limit CO_2 emissions and improve energy efficiency [23]. Alternative fuels for the transportation system would also be an alternative option [24]. Solutions for low-latency failover traffic engineering are proposed to steer the traffic in failure scenarios [25], which can be used to provide reliable communication in CV environment. For the future intelligent transportation systems, solutions on autonomous vehicles smart-platooning [26] and electric vehicle energy prediction [27] are also proposed to improve the green transportation. In most cases, existing infrastructures and transportation systems is also need to be innovated.

C. Multiagent Markov Decision Processes

Markov Games [9] is a multi-agent extension of Markov Decision Processes. They are defined by a set of states, S, action sets for each of N agents, $A_1, \dots, A_N$, a state transition function, $T : S \times A_1 \times \dots \times A_N \to P(S)$, which defines the probability distribution over possible next states, given the current state and actions for each agent, and a reward function for each agent that also depends on the global state and actions of all agents, $R_i : S \times A_1 \times \dots \times A_N \to \mathbb{R}$. We will specifically be considering a partially observable variant in which an agent, i receives an observation, $o_i \in O_i$, which contains partial information from the global state, $s \in S$. Each agent learns a policy, $\pi_i : O_i \to P(A_i)$ which maps each agent's observation to a distribution over it's set of actions. The agents aim to learn a policy that maximizes their expected discounted returns:

$$J_i(\pi_1) = \mathbb{E}_{a_1 \sim \pi_1, \dots, a_N \sim \pi_N, s \sim T}\left[\sum_{t=0}^{\infty} \gamma^t r_{it}(s_t, a_{1t}, \dots, a_{Nt})\right] \tag{1}$$

where $\gamma \in [0, 1]$ is the discount factor that determines how much the policy favors immediate reward over long-term gain.

D. Actor-Critic and Soft Actor-Critic

For many practical problems, the computation of the exact value function is intractable, analytically and numerically, due to the enormous size of the state space. Actor-critic methods [28] aim to ameliorate this issue by using a function approximation of the expected returns, by approximating the value function and restricting the search for a good policy to a smaller family of policies. One specific instance of actor-critic methods learns a function to estimate expected discounted returns, given a state and action, $Q_\psi(s_t, a_t) = \mathbb{E}[\sum_{t'=t}^{\infty} \gamma^{t'-t} r_{t'}(s_{t'}, a_{t'})]$, learned through off-policy temporal-difference learning by minimizing the regression loss:

$$\begin{aligned} &\mathcal{L}_Q(\psi)(s_t, a_t) = \mathbb{E}_{(s,a,r,s')\sim D}[(Q_\psi(s, a - y))^2] \\ &where\ y = r(s,a) + \gamma \mathbb{E}_{a' \sim \pi(s')[Q_{\bar{\psi}}(s',a')]} \end{aligned} \quad (2)$$

where $Q_{\bar{\psi}}$ is the target Q-value function, which is simply an exponential moving average of the past Q-functions and D is a replay buffer that stores past experiences.

To encourage exploration and avoid converging to nonoptimal deterministic policies, recent approaches of maximum entropy reinforcement learning learn a soft value function by modifying the policy gradient to incorporate an entropy term [13]. The loss function for temporal-difference learning of the value function is also revised accordingly with a new target:

$$y = r(s,a) + \gamma \mathbb{E}_{a' \sim \pi(s')}[Q_{\bar{\psi}}(s',a') - \alpha \log(\pi_{\bar{\theta}}(a'|s'))] \quad (3)$$

III. Methodology

A. Problem Definition

In this work, we try to model the signal planning of multiple intersections as a multi-agent reinforcement learning problem. For green transportation, we focus on (1) the traffic efficiency and (2) the energy efficiency of the green transportation system. Traffic efficiency is a common indicator for evaluating a transportation system. Higher energy efficiency means to transport more vehicles passing the intersection in unit energy consumption, thus to reduce the environmental problems caused by traffic congestion. In evaluation, except for the evaluation in traffic control performance, we also make estimation of the environmental impact of the traffic congestion by defining the energy consumption metrics for green transportation. In detail, (1) for the traffic efficiency, an important insight brought by this work is to model the signal planning of multiple intersections in a region as a multi-agent reinforcement learning (MARL) problem; (2) for the energy efficiency, we implement analysis on the fuel consumption and CO_2 emission metrics.

Considering the example in Introduction, the conventional signal algorithm is based on the queuing vehicle data of 8 phases in independent signal intersection without taking correlated intersections into consideration. As Fig. 1 shows that it is ineffective and vulnerable to congestion attacks in the practical scenario, which will further increase the fuel consumption and the CO_2 emissions for green transportation, the main goal for agents is reducing total time consumption of vehicles to prevent traffic congestions. In our proposed method, multiple agents are in cooperative relationships. Based on the current state, each agent generates optimal signal plan that can improve the local traffic condition while the agent group aims to improve the global traffic condition in long term. In the end, agents achieve a policy optimizing the overall signal plans. This is based on the following assumptions:

- Each agent can obtain traffic flow data from their neighbors in CV-based intelligent signal planning system.
- The optimal policy of single agent is not conflict with the optimal target of agent group.

In the real-world scenario, it is estimated that the market penetration rate of CV technology needs 25-30 years to reach at least 95%, which means that the portion of the equipped vehicles is less than 95%. Thus, it is hard for an agent to get global traffic data from all of the other agents as the system cannot record trajectory data of the unequipped vehicles. To deal with the problem in transition period, we introduce the Observation space in POMDP [29], in which the state of agent is unknown, but the observed value of the states is known. For each agent, it observes the traffic data from all of the other agents with different 'attention weights' [30], [31]. Thus, we introduce the Multi-Actor-Attention-Critic (MAAC) [32] in which the attention mechanism functions in a manner similar to a differentiable key-value memory model and each attention head focuses on a different weighted mixture of other agents. Intuitively, each agent queries the other agents for information about their observations and actions and incorporates that information into the estimate of its value function.

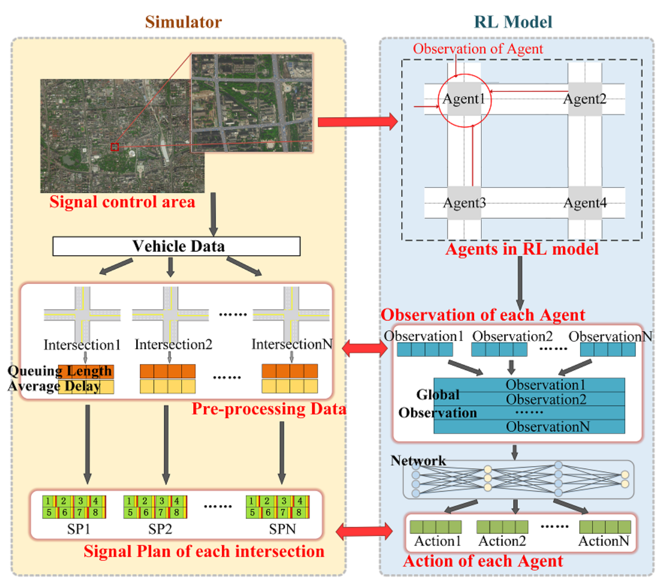


Fig. 4: Multi-agent reinforcement learning-based signal planning framework.

Fig. 4 illustrates the framework of our proposed method. N agents correspond to the N intersection signal control area. To make the RL training process more compatible with the real-time scenario, we propose a novel training framework. The training process are implemented under the real-time interaction between the RL model and the simulator. The preprocessed vehicle data of each section is regarded as the observation of each agent, then the generated global observation is used as the input of the network for policy update. Each agent obtains an action according to their observations and policy, which is the signal plan of the intersection, meanwhile the current policy is updated. Then the current signal plan is sent back to the simulator to execute the next signal cycle. Finally the RL model outputs an global policy aiming at optimizing the signal plan of the whole area.

We consider the standard 8 phase signal planning scenarios, which means that the traffic flow data and the signal plan are both phase-based, and phase settings only consider the traffic input roads. In transition period, the equipped vehicle data for each lane are assigned into three regions: (1) queuing region, including vehicles waiting in the queue with zero speed, (2) slowdown region, including vehicles slowing down because of the front vehicles, and (3) free-flow region, including vehicles far away from the queue so that they behave independently. Among the three regions, the queuing region is more important for signal planning, since the queuing vehicles are stopped and 'waiting' while the goal of signal plan is to minimize the time of 'waiting' in the intersection. To simplify the problem, in each intersection, the input of the learning model is the number of queuing vehicles of each phase and the delay time of all vehicles under the current signal planning stage. Each agent aims to output the signal plan for the intersection to minimize the average delay of passing vehicles. We assume the traffic light on each phase has only two states: red light and green light.

B. Proposed Method

We model the signal planning of the N-intersection area as a multi-agent reinforcement learning problem. As shown in Fig. 4, our method is composed of a simulator and a multi-agent RL model. In the simulator, we gain the road conditions at each time step, which is regarded as the environment of RL model at each training episode. In the RL model, we set N agents according to the N-intersection area. The set of agents is $Agents = \{Agent_1, \ldots, Agent_N\}$, the signal plan in the ith intersection is controlled by $Agent_i$. Each agent observes the environment and executes its action a_i, which is corresponding to the signal plan of the ith intersection in the simulator. The signal plan for all intersections will be written back to the simulator as the beginning of the next signal planning cycle.

The observation, action and reward in our method are defined as follows:

- **Observation.**
  The state of an intersection is presented by the traffic condition of the intersection in real world. As the portion of the equipped vehicles is less than 95%, each agent cannot get the trajectory data of all vehicles. Thus, the observation describes the traffic data which can be observed by agents, which is part of the state set.
  The observation space is $O = \{o_1 \times \ldots \times o_N\}$, in which o_i is the observation of $Agent_i$. $o_i = \{\boldsymbol{V}_i, \ldots, \boldsymbol{V}_N\}^T$ is composed of a 9 tuple, $\boldsymbol{V}_i = \{nq_{i1}, nq_{i2}, \ldots, nq_{i8}, delay_i\}$ is the traffic data of $Agent_i$, where nq_{ik} is the number of queuing vehicles in the phase k, and $delay_i$ is the average of delay for vehicles to pass in the ith intersection.
- **Action.**
  $A = \{a_1 \times \ldots \times a_N\}$ is the action space, in which $a_i = [SP]_{2\times 4}$ is the action of $Agent_i$ recording the signal plan in the ith intersection. SP_{ij} is the green light duration of 8 phases under the order shown in Fig. 3.
- **Reward.**
  The reward is calculated by Eq. 6, which is defined as a weighted sum of the following factors:
  1) Delay time (D). D_i is the sum of delay time of vehicles in the ith intersection.
  2) Queuing time (Q) L_i is the total number of queuing vehicles in the ith intersection, Q_{ij} is the queuing time of vehicle j in the ith intersection. A vehicle with a speed of less than 0.1 m/s is considered as queuing, $\ell_{ij} = 0, 1$ denotes whether vehicle j is queuing. Thus, the queuing time in the ith intersection is defined as Eq. 4.

$$
\begin{aligned}
Q_i &= \sum_{j=1}^{L_i} Q_{ij} * \ell_{ij}, where \\
\ell_{ij} &= \begin{cases} 1, vehiclespeed \leq 0.1 \\ 0, vehiclespeed > 0.1 \end{cases}
\end{aligned} \tag{4}
$$

We define P^i as the total time consumption in the ith intersection and the $R^i = -P^i$ as the reward of the ith intersection:

$$R^i = -P^i = -(\omega_1 * D_i + \omega_2 * Q_i) \tag{5}$$

Hence, by optimising the signal plan according to the observation o_i, each agent aims to minimize the time consumption P^i and thus get the highest local reward R^i. The global reward R is calculated by $R = \frac{\sum_{i=1}^{N} R^i}{N}$. In this situation, the global Q-value function is defined to encourage all agents to get higher long-run reward with sharing parameter:

$$Q = \sum_{i=1}^{N} (R^i + \sum_{i \neq j} \alpha_j R^j) \tag{6}$$

Where Q^i is the Q-value function of $Agent_i$, $\sum_{i \neq j} \alpha_j R^j$ is the contribution from other agents except $Agent_i$ itself. α_j is the attention weight. The attention weight α_j passes the similarity value between embedding e_i and e_j into a softmax:

$$\alpha_j \propto \exp(e_j^T W_k^T W_q e_i) \tag{7}$$

We set N attention heads focusing on different weighted mixture of other agents. Each attention head, using a separate set of parameters (W_k, W_q, V), gives rise to an aggregated contribution from all other agents to the $Agent_i$. Thus, each head can focus on a different weighted mixture of agents.

Algorithm 1 illustrates the policy updating process in each epoch.

Algorithm 1 Policy updating progress

Input: Intersection number N, optimization steps T, policy $\{\pi_i^t\}_{i=1,...,N}$
Output: Policy $\{\pi_i^{t+1}\}_{i=1,...,N}$

Initialize signal plan SP^0, signal plan period ST^0
for $t = 0$ to T **do**
  $VehicleData \leftarrow Simulator(SP^t, ST^t)$
  for $i = 1$ to N **do**
    $\boldsymbol{o}_i^t = \{ql_{i1}, ..., ql_{i8}, delay_i\} \leftarrow Preprocessing(VehicleData)$
  end for
  $\boldsymbol{O}^t = \{\boldsymbol{o}_1^t, ..., \boldsymbol{o}_N^t\}$
  for $i = 1$ to N **do**
    $\{(s_i^{t+1}, a_i^{t+1}, r_i^{t+1})\} \leftarrow roll(\boldsymbol{O}^t, \pi_i^t)$
    $\pi_i^{t+1} \leftarrow optimize\_policy(\{(s_i^{t+1}, a_i^{t+1}, r_i^{t+1})\}, \pi_i^t)$
  end for
  $\boldsymbol{A}^{t+1} = \{a_1^{t+1}, ..., a_N^{t+1}\}$
  $SP^{t+1}, ST^{t+1} \leftarrow SignalController(\boldsymbol{A}^{t+1})$
end for
return $\{\pi_i^{t+1}\}_{i=1,...,N}$

IV. Experimental Results and Analysis

In this section, first, we define metrics for the traffic control and the environmental protection performance. Then, we evaluate different methods under the normal traffic conditions and the congestion attacks and give analysis on their performance in environmental protection.

A. Evaluation Metrics

In our evaluation, we evaluate the green transportation system in two aspects: (1) the traffic efficiency, which indicates the ability of reducing delay time. (2) the energy consumption, which evaluates the performance in sustainable development and environmental protection.

1) Traffic Efficiency: For the traffic efficiency, we focus on the single intersection and the multi-intersection region whole region in the meanwhile. Traditional evaluation metrics in traffic area are no longer suitable for our work. Thus, for a region which contains K intersections, we propose the following novel evaluation metrics:

- **Volume-to-Capacity Ratio (V/C)** [33]. Vehicle capacity of an intersection is the maximum number of vehicles that could theoretically pass through. For the kth 8-phase intersection, the vehicle capacity is $C_k^{max} = \sum_{p=1}^{8} C_p^{max}$. N_k is the total number of vehicles passing through the kth intersection. Thus, the Volume-to-Capacity Ratio (V/C) of the kth intersection is $(V/C)_k = N_k / C_k^{max}$.
- **Delay Time (DT)**. veh_k^n is the nth vehicle passing the kth intersection in all N_k vehicles, then AT_k^n is the actual time cost for veh_k^n, FT_k^n is the free-flow travel time for veh_k^n. Thus, the average delay time for each vehicle passing the kth intersection is $DT_k = \sum_{n=1}^{N_k} (AT_k^n - FT_k^n)/N_k$
- **Intersection Congestion Degree (ICD)**. ICD reflects the congestion degree for a single intersection. Q_p^k is the vehicle number of queuing in the pth phase in the kth intersection and Q_{normal} is a constant that we set $Q_{normal} = 10$. Thus, the congestion degree for the kth intersection is calculated as $ICD_k = \sum_{p=1}^{8} Q_p^k / Q_{normal}^k$.
- **Region Congestion Degree (RCD)**. RCD reflects the congestion degree for the whole region. $I\bar{C}D$ is the average number of queuing vehicles in each intersection, $\bar{N}$ is the average vehicle number of each intersection, then the global congestion degree for the whole region is calculated by $RCD = \sum_{k=1}^{K} (ICD_k - I\bar{C}D)(N_k - \bar{N})/K$.

2) Energy Consumption: For energy consumption in the green transportation, we focus on the fuel consumption and the exhaust emissions. As the proportion of CO_2 in the car exhaust gas is about 95% or even higher, we only analyze the CO_2 emissions of vehicles in all of our experiments. Thus, we compute the following metrics:

- **Fuel Consumption (FC)**. The fuel consumption cost of vehicle is related to vehicle type k, driving speed v, mileage l, and fuel prices P. Thus, the fuel consumption of the traffic region is calculated as: $FC = \sum_k N_k * l * o_k(\bar{v})$, where N_k is the number of the k type vehicles, $o_k(\bar{v})$ is the fuel consumption of the k type vehicles at the average speed $\bar{v}$.
- **CO_2 Emission (CE)**. The CO_2 Emission is calculated as: $CE = N_k * l * e_k(\bar{v})$, where $e_k(\bar{v})$ is the CO_2 emission factor of the k type vehicles at the average speed $\bar{v}$, N_k is the number of the k type vehicles, l is the mileage.

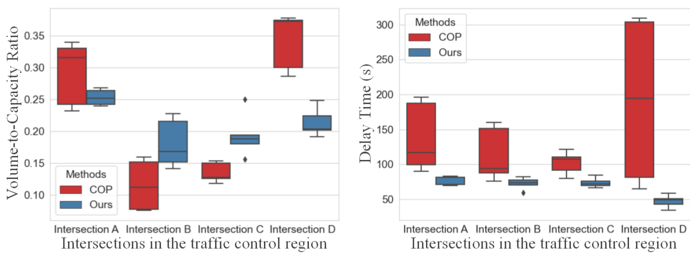
B. Experimental Setup

To make the RL training more compatible with real-time scenario, we consider a novel training framework. The training process are implemented under the real-time interaction between training and simulation, based on multi-agent particle environment framework and the simulator VISSIM. The multi-agent particle environment is introduced by [34], which is useful for creating environments involving complex interaction between agents, while keeping the control and perception problems simple, as we are primarily interested in addressing agent interaction.

Considering a traffic region containing N 8-phase intersections, the topological structure of the traffic region contains direct and indirectly traffic-flow relationships. Referring to the traffic capacity of medium-scale intersections in Beijing City, in our experiments, the traffic volume of in-directions of each traffic region is set to be 2500 ± 300. Accordingly, there are N agents for this task and each of them is set to have $N/2$ attention heads. In each training episode, we input the traffic flow data $\boldsymbol{V}^N$ from simulator into the model. For a region with N intersections, the output of the model is a $N*2*8$ signal control scheme for current episode and will be sent to the simulator again for continuous simulation. When the simulation finished, the reward value $\boldsymbol{R}^N$ and the vehicle data $\boldsymbol{V}^N$ as the input of the next stage are fed back to RL training process. All the experiments are trained with 6 random seeds.

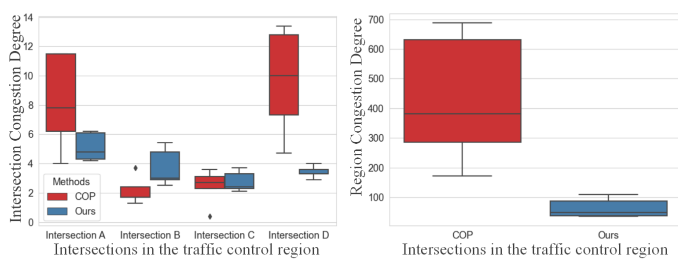
C. Results and Analysis

1) Traffic Efficiency Analysis: In this section, we compare the traffic efficiency of our methods and the conventional COP algorithm based on the proposed metrics in Section IV-A1. Under the same initial traffic inputs, we analyze the performance of different methods under normal traffic condition after 60-minutes simulation. It's clearly that our methods gets lower and more stable value of V/C and DT (see Fig. 5), which denotes that our method reduced the queuing number, and thus reduced the time costs for vehicles to pass the intersection. We further evaluate the ICD and the RDV to intuitively compare the variance in difference between intersection (see Fig. 6). Both of the ICD and the RDV are lower on average, which shows the stability of our method for the whole region.



(a) (b)

Fig. 5: Volume-to-Capacity Ratio (V/C) and Delay Time (DT) under normal traffic condition after 60-minutes simulation. The x-axis of (a) and (b) is the intersection denoted as A, B, C, D. The y-axis of (a) is the Volume-to-Capacity Ratio (V/C), y-axis of (b) is the Delay Time (DT)(seconds).



(a) (b)

Fig. 6: Intersection congestion degree (ICD) and Region Congestion Degree (RCD) under normal traffic condition after 60-minutes simulation. The x-axis of (a) and (b) is the intersection denoted as A, B, C, D. The y-axis of (a) is the Intersection congestion degree (ICD), y-axis of (b) is the Region Congestion Degree (RCD).

We found that, our method uniform the traffic flow in the whole region by relieving the congestion in local intersections to the global intersections. Thereby, our method get lower DT value in average. For the traditional method COP, since the signal planning between the intersection is independent, the congestion in a certain intersection is easy to spread to their neighbor intersections, resulting in a larger area of congestion. However, under the same traffic flow input, our approach can also gain a stability under large traffic flow.

2) Energy Consumption Analysis: In this section, we implement the energy consumption analysis of our method and the conventional COP algorithm based on the proposed metrics in Section IV-A2. As the fuel consumption and the CO_2 emission are different in different vehicle types, we consider the composition of traffic flow under normal traffic conditions, detailed settings are shown in TABLE I.

TABLE I: The proportion of different vehicle types and corresponding CO_2 emissions per kilometre in our experiments.

Vehicle type	**Proportion**	**CO_2 emissions (kg/kilometre)**
Car	0.800	0.207
Bus	0.120	0.069
HGV	0.075	0.143
Tram	0.015	0.042

We compare our method and the COP algorithm under the same traffic flow scale and traffic composition. The simulation duration is set to be 60 minutes. As TABLE II shows, under the same number traffic flow input, our method transported 15172 passing vehicles totally during 60-minutes simulation, higher than the COP by 2.4%, however, the CO_2 emission of our method is lower than the COP. As the goal of green transportation is to increase traffic efficiency with lower energy consumption and exhaust emissions, our method increased the traffic efficiency by transporting more vehicles, and meanwhile, reduced the CO_2 emissions with the fuel consumption are at the same level.

TABLE II: The proportion of different vehicle types and corresponding CO_2 emissions per kilometre in our experiments.

	Vehicle Number	**Fuel consumption (l)**	**CO_2 emissions (kg)**
COP	14816	1958.18	2924.63
Ours	**15172**	1986.26	**2946.80**

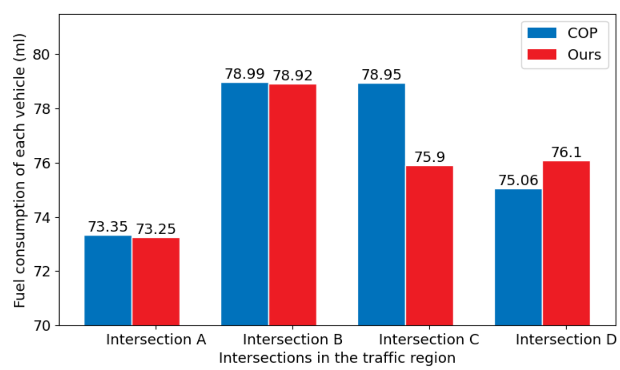


Fig. 7: The fuel consumption comparison of our method and the COP algorithm. The x-axis is the intersection denoted as A, B, C, D. The y-axis is the average of fuel consumption in 60-minutes simulation for each vehicle.

We further compared the average fuel consumption of each passing vehicle (see Fig. 7) and the CO_2 emission of each passing vehicle (see Fig. 8), from which we found that our method reduced both of the fuel consumption and the CO_2 emission. With the same number traffic flow input, the fuel consumption and the CO_2 emission of our method show more stability in each intersection and get lower value in total.

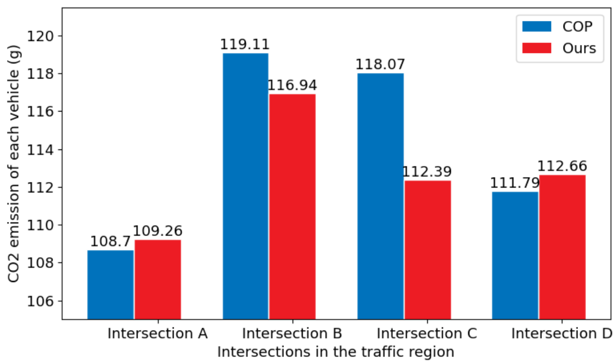
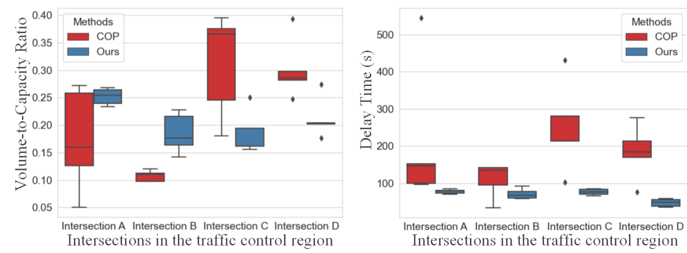


Fig. 8: The CO_2 emission comparison of our method and the COP algorithm. The x-axis is the intersection denoted as A, B, C, D. The y-axis is the average of CO_2 emission in 60-minutes simulation for each vehicle.



(a) (b)

Fig. 10: Volume-to-Capacity Ratio (V/C) and Delay Time (DT) under congestion attacks after 60-minutes simulation. The x-axis of (a) and (b) is the intersection denoted as A, B, C, D. The y-axis of (a) is the Volume-to-Capacity Ratio (V/C), y-axis of (b) is the Delay Time (DT)(seconds).The attack vehicle is put on the Intersection C.

3) Safety Analysis: The goal of the congestion attack is to disturb the signal plan and increase the total delay of all vehicles in the intersection. It causes a series of environmental, economic, and social problems, such as the increasing of CO_2 emission and fuel consumption, the raising of travel costs, and the occurrence of traffic congestion and accidents. Recent work has shown that the congestion attack called *the last vehicle attack* may cause the failure of COP algorithm, and increase the total delay by as high as 68.1%, which completely reverses the benefit of using the I-SIG system (26.6% decrease). Fig. 9 shows the attack strategy of *the last vehicle attack*.

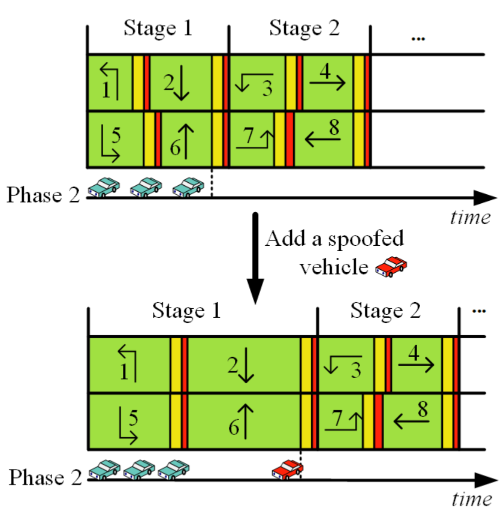
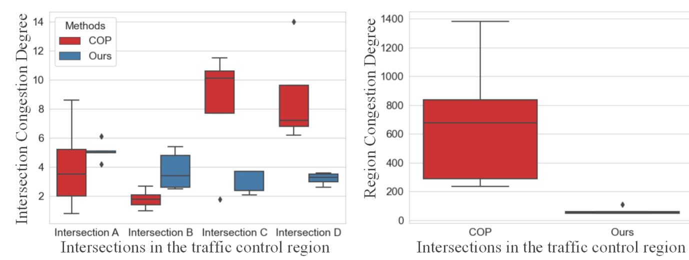


Fig. 9: Attack strategy of the 'last vehicle'. As shown, the spoofing vehicle was set on the phase 2, and in the end prolonged the green light duration of phase 2.



(a) (b)

Fig. 11: Intersection congestion degree (ICD) and Region Congestion Degree (RCD) under congestion attacks after 60-minutes simulation. The x-axis of (a) and (b) is the intersection denoted as A, B, C, D. The y-axis of (a) is the Intersection congestion degree (ICD), y-axis of (b) is the Region Congestion Degree (RCD). The attack vehicle is put on the Intersection C.

In the attack model, a spoofing vehicle is set to stop at the end of a phase. It will extend the green light duration of the attacked phase (phase 2), and thus delay the green light start time of all the following phases (phase 3,4,7,8). After a period of continuous attack, these phases will be growingly congested with more queuing vehicles. To maximize the impact of the attack, the attacker tends to add spoofing vehicles when the number of vehicles in the target phase is large. The behind intuition is that, under the last vehicle attack, the cumulative time delay will be greater when there are more vehicles in the road, causing more massive congestion.

In this section we analyze the traffic efficiency of our method and the COP algorithm under data spoofing attacks. Fig. 10 and Fig. 11 show the result. It's clear that even under the congestion attack, our method gets lower and more stable CV, DT, ICD and RCD values than the COP algorithm, in which the DT is at the same level of the condition without attacks. However, as the attack vehicle is set at the 1th phase of the intersection C, the COP shows obvious abnormal traffic in both Intersection C and its neighbor D, the DT in intersection C increased 156.62% compared to normal traffic especially. Our method gets higher traffic efficiency and shows more general performance in the whole region. It shows robustness against the congestion attacks and stability under large traffic flow.

V. Related Work

A. Reinforcement Learning for Traffic Light Control

As the widely using of the CV technology in transportation systems, conventional signal planning algorithms are no longer suitable due to its inefficiency and potential risks. More works try to use reinforcement learning algorithms to solve the traffic light control problem [10], [35]. Typically, RL-based algorithms take the traffic on the road as state, and the operation on light as action. Methods in [35] designed the state as discrete values like the location of vehicles or number of waited cars, in which the discrete state-action pair value matrix requires However, the discrete state-action pair value matrix requires huge storage space.

To solve the unmanageable large state space of previous methods, recent studies [11], [36] propose to apply Deep Q-learning methods using continuous state representations. These studies learn a Q-function (e.g. a deep neural network) to map state and action to reward. These works vary in the state representation including hand craft features (e.g., queue length, average delay) and image features) They are also different in reward design, including average delay, the average travel time, and queue length.

However, all these methods assume relatively static traffic environments, and hence far from the real case. Therefor,in this paper, we try to train the model under a novel framework, in which the training process are implemented under the real-time interaction between the RL model and the simulator. What's more, all of these works ignore the traffic flow relationship between intersections, which can be used in traffic forecast.By introduce the attention mechanism, Our method is adjusted to different region scale, which is important in the global signal plan scenario.

B. Traffic Signal Control Algorithm Security

The security problems of the intelligent traffic signal system based on CV technology have been being revealed in several recent works [7], [37]. Such attacks can use message falsification (modification), spoofing (masquerading), or replay attacks to maliciously affect the vehicle stream, leading to rear-end collisions in severe cases.

Prior to our study, Laszka et al. performed a theoretical analysis to estimate the potential congestion an attacker can create assuming that she can arbitrarily compromise multiple signal controllers [38]. A follow-up study was then performed for the same attack goal but with a weak assumption, in which the attacker can only compromise the sensors that collects traffic flow information [39], [40]. In comparison, neither of these works analyzes the CV-based signal control scenario targeted in our work. Compared to existing studies, our work valuates the performance in different region scales. The multi-intersection signal control scenario, in which the number of intersections is adjusted to the RL model, is much more realistic.

For traditional traffic control, our first objective should be safety, followed by mobility and other objectives such as sustainability [41]. Safety is traditionally guaranteed by design: including the dual-diagram design scheme [42] and other associated techniques such as the conflict monitor embedded in traffic controllers. Mobility is usually considered by minimizing the total delays or travel times of all vehicles passing the intersection (e.g., for optimizing the timing of a single intersection), or maximizing the throughput or other related measures (for coordinating multiple intersections). The sustainability objective is often defined as the total energy consumption or emissions of vehicles passing the intersection [8], which is significant in green transportation nowadays. In CV-based systems, traffic data from road side and the infrastructure side are collected together by the net, it is a kind of the Cloud-assisted Internet of Things (IoT), thus, a Privacy-Enhanced Retrieval Technology for Cloud-assisted IoT is proposed to preserves data privacy in CV-based system [43]. Furthermore, as the Industry 5.0 is on the way, recent work also discussed the future of the Industry 5.0.cloud encrypted storage model [44], meanwhile, with lower time cost.

C. Congestion Costs Estimation in Green Transportation

In green transportation, traffic congestion increases the delay time and travel time of vehicles, thus leads to environmental problems. As a tool to quantitative analysis the results, estimation of the environmental impact of congestion is performed using different modelling approaches. Important distinction among the various methods applied worldwide is the spatial and temporal resolution they offer. In general, available methods can be divided between microscopic and macroscopic ones [45], [46]. Microscopic emission models can estimate the instantaneous emissions on a second-by-second basis and are more suited to assess interventions on single roads or junctions. In such approaches, fuel consumption and exhaust emissions calculation require the exact speed profile of the vehicles as an input. The speed profiles are provided in most cases by micro traffic simulation models and assumptions on gear changes are made to predict the engine operation [47].

Macroscopic models, on the other hand, have been developed to estimate fleet emissions over a region or on a country-wide scale, and are mostly utilizing average speed as an input. Models of this type, such as COPERT [48], ARTEMIS [49] and NAEI [50], use the average speed to predict fuel consumption and emissions, with the produced factors being expressed in mass of pollutant per unit of distance travelled (e.g. g/km). Two main factors have contributed to the widespread usage of the average speed models: they are easy to use, and they do not require detailed input data, such as second-by-second vehicle trajectories [51].

VI. Conclusion and Future Work

In this paper, we focus on both of the traffic efficiency and the energy efficiency in green transportation. For the traffic efficiency, we model the signal planning of multi-intersection region as a multi-agent reinforcement learning problem; for the energy efficiency, we aim to improve the traffic efficiency meanwhile reduce the fuel consumption and CO_2 emission. In our RL model, we introduce the attention mechanism considering the relationship of traffic flow between intersections. We conduct extensive experiments using a novel framework, in which the training process are implemented under the real-time interaction between the RL model and the simulator. In addition, we compared the performance of our method with the traditional signal plan algorithms under the congestion attacks. To evaluate the performance of RL model in signal plan scenario, we propose metric based on the multi-intersection environment. The results shows that our method is not only fixable to different scale of multi-intersection region, but also robust and stable to the attack conditions.

This work severs as a first step to explore the multi-intersection collaborative signal planning in the next-generation CV-based transportation systems. It is expected to inspire a series of follow-up studies, including but not limited to 1) more extensive evaluation with larger traffic network and more intersections, 2) more energy-efficiency communication strategies among agents of different intersections, 3)safety guarantees and balance of multiple agents.

References

[1] "Connected vehicle applications," https://www.its.dot.gov/pilots/cvpilotapps.htm.

[2] Y. Feng, J. A. Zhang, B. Cheng, X. He, and J. Chen, "Magnetic sensor-based multi-vehicle data association," *IEEE Sensors Journal*, vol. 21, no. 21, pp. 24 709–24 719, 2021.

[3] "Usdot: Multi-modal intelligent traffic safety system (mmitss)," https://www.its.dot.gov/researcharchives/dma/bundle/mmitssplan.html.

[4] "U.s. dot connected vehicle pilot deployment program," https://www.its.dot.gov/pilots/.

[5] "Mmitss final conops: Concept of operations," http://www.cts.virginia.edu/wp-content/uploads/2014/05/Task2.3.CONOPS6FinalRevised.pdf.

[6] S. Sen and K. L. Head, "Controlled optimization of phases at an intersection," *Transportation science*, vol. 31, no. 1, pp. 5–17, 1997.

[7] Q. A. Chen, Y. Yin, Y. Feng, Z. M. Mao, and H. X. Liu, "Exposing congestion attack on emerging connected vehicle based traffic signal control." in *NDSS*, 2018.

[8] H. Jiang, J. Hu, S. An, M. Wang, and B. B. Park, "Eco approaching at an isolated signalized intersection under partially connected and automated vehicles environment," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 290–307, 2017.

[9] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *Machine learning proceedings 1994*. Elsevier, 1994, pp. 157–163.

[10] P. Mannion, J. Duggan, and E. Howley, "An experimental review of reinforcement learning algorithms for adaptive traffic signal control," *Autonomic road transport support systems*, pp. 47–66, 2016.

[11] E. Van der Pol and F. A. Oliehoek, "Coordinated deep reinforcement learners for traffic light control," *Proceedings of Learning, Inference and Control of Multi-Agent Systems (at NIPS 2016)*, 2016.

[12] W. Chen, X. Qiu, T. Cai, H.-N. Dai, Z. Zheng, and Y. Zhang, "Deep reinforcement learning for internet of things: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 23, no. 3, pp. 1659–1692, 2021.

[13] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.

[14] H. Wei, G. Zheng, H. Yao, and Z. Li, "Intellilight: A reinforcement learning approach for intelligent traffic light control," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018, pp. 2496–2505.

[15] S. I. T. Joseph, S. Velliangiri, and C. S. Chandra, "Investigation of deep learning methodologies in intelligent green transportation system," *Journal of Green Engineering*, vol. 10, pp. 931–950, 2020.

[16] L. Buşoniu, R. Babuška, and B. De Schutter, "Multi-agent reinforcement learning: An overview," *Innovations in multi-agent systems and applications-1*, pp. 183–221, 2010.

[17] M. Niu, B. Cheng, Y. Feng, and J. Chen, "Gmta: A geo-aware multi-agent task allocation approach for scientific workflows in container-based cloud," *IEEE Transactions on Network and Service Management*, vol. 17, no. 3, pp. 1568–1581, 2020.

[18] "Ptv vissim," http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim.

[19] Y. Feng, K. L. Head, S. Khoshmagham, and M. Zamanipour, "A real-time adaptive signal control in a connected vehicle environment," *Transportation Research Part C: Emerging Technologies*, vol. 55, pp. 460–473, 2015.

[20] M. Björklund, "Influence from the business environment on environmental purchasing — drivers and hinders of purchasing green transportation services," *Journal of Purchasing and Supply Management*, vol. 17, pp. 11–22, 03 2011.

[21] J.-P. Rodrigue, C. Comtois, and B. Slack, *The Geography of Transport Systems*, 12 2016.

[22] C. T. Lee, H. Hashim, C. S. Ho, Y. V. Fan, and J. J. Klemeš, "Sustaining the low-carbon emission development in asia and beyond: Sustainable energy, water, transportation and low-carbon emission technology," *Journal of Cleaner Production*, vol. 146, pp. 1–13, 2017.

[23] B. Singh and A. Strømman, "Environmental assessment of electrification of road transport in norway: Scenarios and impacts," *Transportation Research Part D: Transport and Environment*, vol. 25, p. 106–111, 12 2013.

[24] F. Raymand, P. Ahmadi, and S. Mashayekhi, "Evaluating a light duty vehicle fleet against climate change mitigation targets under different scenarios up to 2050 on a national level," *Energy Policy*, vol. 149, p. 111942, 2021.

[25] H. Mostafaei, M. Shojafar, and M. Conti, "Tel: Low-latency failover traffic engineering in data plane," *arXiv preprint arXiv:2009.13640*, 2020.

[26] S. B. Prathiba, G. Raja, K. Dev, N. Kumar, and M. Guizani, "A hybrid deep reinforcement learning for autonomous vehicles smart-platooning," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, pp. 13 340–13 350, 2021.

[27] W. Wang, M. H. Fida, Z. Lian, Z. Yin, Q.-V. Pham, T. R. Gadekallu, K. Dev, and C. Su, "Secure-enhanced federated learning for ai-empowered electric vehicle energy prediction," *IEEE Consumer Electronics Magazine*, pp. 1–1, 2021.

[28] V. R. Konda and J. N. Tsitsiklis, "Actor-critic algorithms," in *Advances in neural information processing systems*, 2000, pp. 1008–1014.

[29] C. Sarraute, O. Buffet, and J. Hoffmann, "Pomdps make better hackers: Accounting for uncertainty in penetration testing," in *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[30] A. Graves, G. Wayne, and I. Danihelka, "Neural turing machines," *arXiv preprint arXiv:1410.5401*, 2014.

[31] J. Oh, V. Chockalingam, H. Lee *et al.*, "Control of memory, active perception, and action in minecraft," in *International Conference on Machine Learning*. PMLR, 2016, pp. 2790–2799.

[32] S. Iqbal and F. Sha, "Actor-attention-critic for multi-agent reinforcement learning," in *International Conference on Machine Learning*. PMLR, 2019, pp. 2961–2970.

[33] M. Rahdar, L. Wang, J. Dong, and G. Hu, "Resilient transportation network design under uncertain link capacity using a trilevel optimization model," *Journal of Advanced Transportation*, vol. 2022, 2022.

[34] I. Mordatch and P. Abbeel, "Emergence of grounded compositional language in multi-agent populations," in *Thirty-second AAAI conference on artificial intelligence*, 2018.

[35] M. A. Wiering, "Multi-agent reinforcement learning for traffic light control," in *Machine Learning: Proceedings of the Seventeenth International Conference (ICML'2000)*, 2000, pp. 1151–1158.

[36] L. Li, Y. Lv, and F.-Y. Wang, "Traffic signal timing via deep reinforcement learning," *IEEE/CAA Journal of Automatica Sinica*, vol. 3, no. 3, pp. 247–254, 2016.

[37] T. Jeske, "Floating car data from smartphones: What google and waze know about you and how hackers can control traffic," *Proc. of the BlackHat Europe*, pp. 1–12, 2013.

[38] A. Laszka, B. Potteiger, Y. Vorobeychik, S. Amin, and X. Koutsoukos, "Vulnerability of transportation networks to traffic-signal tampering," in *2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS)*. IEEE, 2016, pp. 1–10.

[39] A. Ghafouri, W. Abbas, Y. Vorobeychik, and X. Koutsoukos, "Vulnerability of fixed-time control of signalized intersections to cyber-tampering," in *2016 Resilience Week (RWS)*. IEEE, 2016, pp. 130–135.

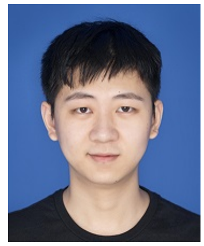
[40] Y. Feng, G. Mao, B. Cheng, C. Li, Y. Hui, Z. Xu, and J. Chen, "Magmonitor: Vehicle speed estimation and vehicle classification through a magnetic sensor," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2020.

[41] M. Li, X. Chen, X. Lin, D. Xu, and Y. Wang, "Connected vehicle-based red-light running prediction for adaptive signalized intersections,"

Journal of Intelligent Transportation Systems, vol. 22, no. 3, pp. 229–243, 2018.
[42] R. P. Roess, E. S. Prassas, and W. R. McShane, *Traffic engineering*. Pearson/Prentice Hall, 2004.
[43] T. Wang, Y. Quan, X. S. Shen, T. R. Gadekallu, W. Wang, and K. Dev, "A privacy-enhanced retrieval technology for the cloud-assisted internet of things," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2021.
[44] P. K. R. Maddikunta, Q.-V. Pham, P. B, N. Deepa, K. Dev, T. R. Gadekallu, R. Ruby, and M. Liyanage, "Industry 5.0: A survey on enabling technologies and potential applications," *Journal of Industrial Information Integration*, p. 100257, 2021.
[45] P. Boulter, I. S. McCrae, and T. J. Barlow, *A review of instantaneous emission models for road vehicles*. TRL Limited Wokingham, United Kingdom, 2007, vol. 64.
[46] A. Alkafoury, M. Bady, M. H. F. Aly, and A. M. Negm, "Emissions modeling for road transportation in urban areas: State-of-art review," in *Proceeding of 23rd International Conference on—Environmental Protection is a Must*, vol. 23, 2013.
[47] L. I. Panis, S. Broekx, and R. Liu, "Modelling instantaneous traffic emission and the influence of traffic speed limits," *Science of the total environment*, vol. 371, no. 1-3, pp. 270–285, 2006.
[48] L. Ntziachristos, D. Gkatzoflias, C. Kouridis, and Z. Samaras, "Copert: a european road transport emission inventory model," in *Information technologies in environmental engineering*. Springer, 2009, pp. 491–504.
[49] P. Boulter and I. McCrae, "Assessment and reliability of transport emission models and inventory systems–final report," *ARTEMIS deliverable no*, vol. 15, 2007.
[50] P. Brown, M. Broomfield, G. Buys, L. Cardenas, E. Kilroy, J. MacCarthy, T. Murrells, Y. Pang, N. Passant, J. Ramirez Garcia *et al.*, "Uk greenhouse gas inventory, 1990 to 2014: annual report for submission under the framework convention on climate change," 2016.
[51] T. J. Barlow and P. Boulter, *Emission factors 2009: Report 2-a review of the average-speed approach for estimating hot exhaust emissions*, 2009.

Yike Li received bachelor's degree from Hefei University of Technology in 2018. She is currently a graduate student of Beijing Jiaotong University. Her current research interests include AI security and Image generation.

Wenjia Niu received the bachelor's degree in computer science from Beijing Jiaotong University in 2005, and the Ph.D. degree in computer science from the Chinese Academy of Sciences in 2010. He is currently a Professor with Beijing Jiaotong University. His research interests are AI security, agent, and data mining.

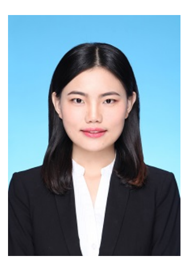


Yunzhe Tian received the bachelor's degree from Beijing Information Science & Technology University in 2020. He is currently pursuing the graduate degree with Beijing Jiaotong University. His research interests are in information security and AI security.

Tong Chen received her M.S. degree in cyber security from Beijing Jiaotong University, Beijing, in 2018. She is currently a Ph.D. candidate of cyber security in Beijing Jiaotong University. Her main research interests are cyber security and reinforcement learning security.

Zhiqiang Xie received bachelor's degree from Beijing Jiaotong University in 2020. He is currently a master student of Cyber security in Beijing Jiaotong University. His research interests are AI security and data privacy preservation.

Yalun Wu received bachelor's degree from Qingdao Agricultural University in 2017. He is currently a graduate student of Beijing Jiaotong University. His current research interests include AI security and information security.



Yingxiao Xiang received the BS degree in computer science and technology from Taiyuan University of Technology in 2016, Master degree from Beijing Jiaotong University in 2019. She is currently a PhD student of information security in Beijing Jiaotong University. Her research interests are in information security and AI security. AI security.

Endong Tong received the Ph.D. degree from the Chinese Academy of Sciences, Beijing, China, in 2013. He is currently an Assistant Professor with Beijing Jiaotong University. He has published more than 30 research papers in refereed international conferences and journals. His current research interests include AI security, services computing, and data mining.

Thar Baker (Senior Member, IEEE) received the Ph.D. degree in autonomic cloud applications from Liverpool John Moores University (LJMU), U.K., in 2010 and became a Senior Fellow of Higher Education Academy in 2018. He is an Associate Professor with the Department of Computer Science, University of Sharjah (UoS), UAE. Before joining UoS, he was a Reader of Cloud Engineering and the Head of the Applied Computing Research Group, Faculty of Engineering and Technology, LJMU. Prior to that, he was a Lecturer of Computer Science with the Department of Computing and Mathematics, Manchester Metropolitan University, U.K. He has published numerous refereed research papers in multidisciplinary research areas, including parallel and distributed computing, algorithm design, green and sustainable computing, and energy routing protocols.

Jiqiang Liu (Senior Member, IEEE) received the B.S. and Ph.D. degrees from Beijing Normal University in 1994 and 1999, respectively. He is currently a Professor with the School of Computer and Information Technology, Beijing Jiaotong University. He has published over 100 scientific papers in various journals and international conferences. His main research interests are trusted computing, cryptographic protocols, privacy preserving, and network security.