WILEY | Hindawi

Research Article

Refined Path Planning for Emergency Rescue Vehicles on Congested Urban Arterial Roads via Reinforcement Learning Approach

Longhao Yan (ID),[1] **Ping Wang** (ID),[1] **Jingwen Yang** (ID),[1] **Yu Hu** (ID),[1] **Yu Han,**[2,3] **and Junfeng Yao**[4,5]

[1]*School of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China*
[2]*School of Intelligent Systems Engineering, Sun Yat-Sen University, Guangzhou 510275, China*
[3]*Guangdong Province Key Laboratory of Fire Science and Technology, Guangzhou 510006, China*
[4]*School of Information Engineering, Chang'an University, Xi'an 710064, China*
[5]*China Communications Information Technology Group Co., Ltd, Beijing 100088, China*

Correspondence should be addressed to Ping Wang; pingwang@chd.edu.cn

Fast road emergency response can minimize the losses caused by traffic accidents. However, emergency rescue on urban arterial roads is faced with the high probability of congestion caused by accidents, which makes the planning of rescue path complicated. This paper proposes a refined path planning method for emergency rescue vehicles on congested urban arterial roads during traffic accidents. Firstly, a rescue path planning environment for emergency vehicles on congested urban arterial roads based on the Markov decision process is established, which focuses on the architecture of arterial roads, taking the traffic efficiency and vehicle queue length into consideration of path planning; then, the prioritized experience replay deep Q-network (PERDQN) reinforcement learning algorithm is used for path planning under different traffic control schemes. The proposed method is tested on the section of East Youyi Road in Xi'an, Shaanxi Province, China. The results show that compared with the traditional shortest path method, the rescue route planned by PERDQN reduces the arrival time to the accident site by 67.1%, and the queue length at upstream of the accident point is shortened by 16.3%, which shows that the proposed method is capable to plan the rescue path for emergency vehicles in urban arterial roads with congestion, shorten the arrival time, and reduce the vehicle queue length caused by accidents.

1. Introduction

Urban arterial roads are important components of the urban traffic system, traffic accidents occurring on them produce a major impact on the entire urban road network and result in huge casualties and large economic losses [1–3]. It is common to see that traffic accidents would lead to congestions due to the high volume of urban road networks, which greatly change the capacity of road network and influence the optimal path for emergency rescuing [4–7]. Thus, being a typical phenomenon in urban traffic accident rescuing, it is of great significance to conduct studies on path planning for emergency vehicles on congested urban arterial roads.

The complex composition of urban arterial roads, multiple intersections, and large traffic flow have brought many difficulties to urban road vehicle path planning, which make it more complicated when compared with route planning on freeways [8–12]. Furthermore, the rescue path planning of emergency vehicles is different from the general route planning problem. Firstly, the rescue path planning of emergency vehicles will encounter various road conditions; because of the large traffic volume of urban roads, traffic accidents result in road congestion easily, which greatly reduces the capacity of the road upstream of the accident site and makes arrival to the accident site time-consuming. Moreover, the architecture of many urban arterial roads is composed of main road and side road, containing lots of

intersections, which gives a number of options for path planning and deserves a refined plan.

A great number of research studies focus on vehicle path planning problems. Yang et al. proposed a path planning method for emergency vehicles, where the road network is divided into a weighted grid and then the rescue path is planned by vector grid map method [13]. The green ant method was proposed by Jabbarpour et al. to provide path planning for unmanned ground vehicles which leads to low power consumption [14]. Wu et al. used an improved ant colony algorithm in congested urban area for dynamic path planning, introducing a designed road evaluating factor instead of road distance to combine with particle swarm optimization [15]. A system was established by Zhang et al. to plan path in real time by efficiently caching and conducting an experiment on a real road network [16]. Karouri et al. proposed an efficient path planning method over large-scale traffic scenario using Dijkstra greedy algorithm and Green Light Optimal Speed Advisory service [17]. Most of the research studies mentioned above are towards large-scale road networks, yet the path planning for the refined urban road needs to be improved, and more attention should be paid to the research on the optimal rescue path under different traffic control schemes on congested roads.

With the development of artificial intelligence, algorithms like deep Q-network have been widely used for decision making in various practical problems [18–20]. The use of reinforcement learning in path planning is increasing and has provided different goal-oriented path planning for various types of vehicles due to its strong performance and high applicability in decision making of path selection. Liu et al. designed a best path selection method to help different types of intelligent driving vehicles based on the prior knowledge applied reinforcement learning strategy [21]. Combining hierarchical reinforcement learning with neural networks, Yu proposed a path planning algorithm for mobile robots and tested it in different scenarios [22]. Chen proposed a path planning scheme using deep reinforcement learning for autonomous vehicles to reduce transport costs and to increase traffic efficiency [23]. While being a value-based deep reinforcement learning algorithm with stability and high qualified experience leading to strong performance in strategy optimization [24], PERDQN can provide path planning for emergency rescue vehicles in a complex road network with reliability. Therefore, we established a refined road network model for congested urban arterial roads and used PERDQN to provide path planning for emergency vehicles during traffic accidents, which aims at reaching the accident site with shortest time and least loss of road capacity on road network.

The contents of this paper are organized as follows. Section 2 explains how to build a Markov decision process for urban arterial roads while considering the traffic efficiency and impact on road network, and also describes how rescue paths for emergency vehicles are planned based on PERDQN algorithm under different traffic control schemes. Section 3 presents the result of the proposed method in the case of real urban arterial road. Section 4 summarizes the contributions of this paper and gives an outlook on our future study.

2. Methodology

2.1. Path Planning on Congested Urban Arterial Roads Based on PERDQN.
The rescue path of the emergency vehicle refers to the driving track from its departure to the arrival at the scene of the traffic accident, different rescue paths lead to diverse arrival time and road network impact. And from the perspective of emergency vehicle, the path planning problem can be regarded as a problem about how to drive to the accident site with the shortest time and with the least impact on the road network, while the driving decisions emergency vehicle makes at each road node determine the rescue path it drives. Therefore, a rescue path planning environment for congested urban arterial roads based on the Markov decision process is constructed, and PERDQN is used to plan path for emergency vehicles on basis of the environment. Through interacting with the MDP environment, the emergency vehicles learn from experience and improve the path planning capabilities and eventually find the optimal rescue path under various control schemes. Figure 1 shows the framework of rescue path planning for emergency vehicles based on PERDQN.

2.2. Establishment of Congested Urban Arterial Road Environment.
Urban arterial roads have diverse functions and complex compositions. Most of them consist of main roads and side roads, and there are often many intersections interspersed in them. Moreover, it is a common phenomenon that the main roads in the opposite directions on the arterial road are separated by continuous road fences, while numbers of joint roads make mainroad and side road in the same direction connected. This provides multipleoptions for emergency vehicles path planning during accidents rescuing, especially when the road is congested due to traffic accidents. Thus, it is necessary to establish a model that extracts the key issue of problem while maintaining the characteristics of the arterial roads.

The node-segment model is a common method that simplifies the path planning problem while preserving the characteristics of the road networks [13–15]. The model is composed of nodes and road segments, which especially focuses on the joint point of roads, where vehicle can change path, such as the connection between the main road and side road in the same direction, and the intersections that connect opposite roads are regarded as nodes in the model. The node model $N = \left\{N_{ij}\right\}_{i=1,2,\ldots,m;j=1,2,\ldots,n}$, where N_{ij} denotes the jth node on the ith road, m is the number of main roads and side roads in the road network, and n denotes the amount of joint points on separated roads. The segment model $\mathrm{Seg} = \left\{\mathrm{Seg}_{ij}\right\}_{i=1,2,\ldots,m-1;j=1,2,\ldots,n-1}$ represents the road segment between nodes, where Seg_{ij} is the road segment between N_{ij} and N_{ij+1}, denoted as $\mathrm{Seg}_{ij} = |N_{ij+1} - N_{ij}|$. Seg_{ij} is the length of road segment. Moreover, different lanes on the single main road or side road are regarded as indistinguishable for sharing the same limited speed and traffic control requirements, leading to the same travel efficiency and road impact in path planning.
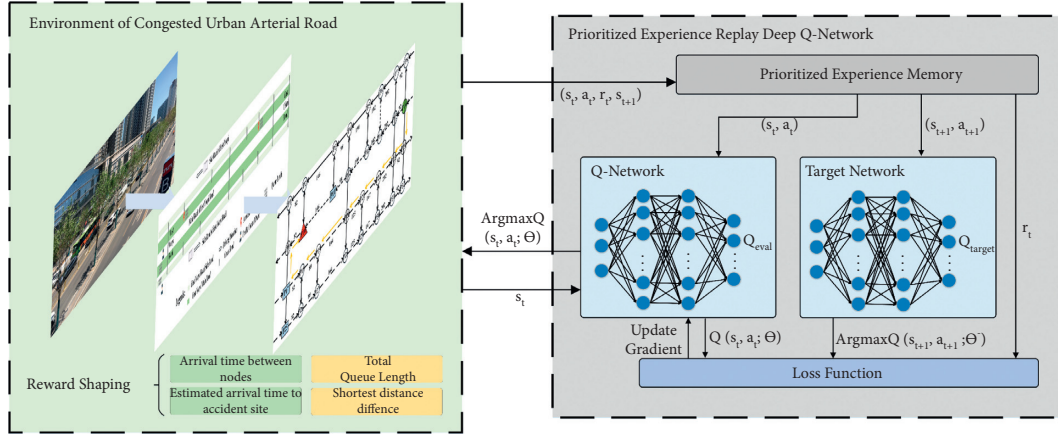
FIGURE 1: Framework of rescue path planning for emergency vehicles on congested urban arterial road using PERDQN.

On the basis of the established node-segment model of road networks, the path planning problem needs to be further adapted into a decision-making problem; being a dynamic, stochastic mathematical framework that describes sequential decision making, the Markov decision process (MDP) is introduced to establish a decision-making environment for the path planning. A Markov decision process consists of a tuple $\langle S, A, T, R \rangle$, where state $S = (s_k, k = 1, 2, \ldots,)$ represents the position that emergency vehicle is on, denoted as road node in road network, the mapping from node N_{ij} to state s can be represented as $s_k = \sum_{l=1}^{i-1} n_l + j|_{\text{node}=N_{ij}}$, and n_l is the mount of node at l th lane. Action $A = (1, 2, 3, \ldots,)$ indicates the next road node that emergency vehicle is going to. T is the transition probability of the vehicle moving from the current node to the next node, which also represents the mechanism of MDP; the factors that affect vehicle driving between nodes include the speed of the vehicle on the road section, the length of the road section, and the capacity of the road segment, and they are included in reward shaping. The reward function $R(S, A, S')$ defines the reward value obtained from the current node to the next node through driving. Great attention is paid to the traffic efficiency and road network impact in traffic research studies [25–27], and with the traffic wave theory suggested in [28, 29] and the Bureau of Public Roads (BPR) function proposed in [30], the vehicle queuing length (VQL) during traffic accidents and arrival time in congested road networks can be calculated. Further, to let emergency vehicles take the road network impact and travel efficiency into consideration, the reward function is designed to be

$$R_{ij} = \begin{cases} -\left[\alpha \left(\alpha_1 t_a^{ij} + \alpha_2 t_{ea}^i \right) + \beta \left(\beta_1 q_l + \beta_2 l_{dif}^{ij} \right) \right], & s_j \neq s_{\text{terminal}}, \\ -\left[\alpha \left(\alpha_1 t_a^{ij} + \alpha_2 t_{ea}^i \right) + \beta \left(\beta_1 q_l + \beta_2 l_{dif}^{ij} \right) \right] + r_{\text{terminal}}, & s_j = s_{\text{terminal}}, \end{cases} \tag{1}$$

where R_{ij} denotes the reward that emergency vehicle received when driving from state s_i to state s_j, t_a^{ij} is the time taken from s_i to state s_j, and it is calculated by BPR function(5), t_{ea}^i denotes the estimated arrival time to accident site from s_i, which is the sum of the arrival time of each section on the shortest path (the shortest path to accident site from s_i is generated by Dijkstra's algorithm based on segment matrix that has been mentioned in part of modelling, Dijkstra's algorithm is a method widely used for shortest route calculation [31]), and q_l refers to the sum of the expected queue lengths under different control schemes and is calculated by traffic wave theory(1). $l_{dif}^{ij} = l_j - l_i$, where l_j indicates the shortest distance from s_j to the accident point that also calculated by Dijkstra's algorithm. t_a^{ij} and l_{dif}^{ij} focus on the difference between nodes and provide local information for emergency vehicle, while t_{ea}^i offers estimated time to accident site and q_l indicates the impact of the road

network under accidents and traffic control schemes. Moreover, to make the time factors t_a^{ij}, t_{ea}^i and distance factors q_l, l_{dif}^{ij} share the same magnitude in value, time factors are preprocessed by rescaling normalization to the range of mean of distance in road segments $L = \{l_i\}_{i \in \text{road segment}}$. $\alpha, \alpha_1, \alpha_2, \beta, \beta_1, \beta_2$ are factors that control ratio of corresponding indicators, and they are set to be empirical values as 0.9, 0.9, 0.1, 0.1, 0.2, 0.8, respectively. s_{terminal} is the goal of route and refers to the location of the accident site. r_{terminal} is a constant that motivates vehicle to get to the accident point and is 150 in our established experiment.

In the path planning problem, the emergency vehicle needs to choose its driving behaviour to reach the next location according to current location, which means in MDP, the emergency vehicle selects its driving action according to the driving strategy $\pi = \pi(a|s)$ on the basis of

current state s ($\pi(a|s)$ indicates the probability of choosing action a at state s). Therefore, the goal of MDP is to find an optimal strategy that drives emergency vehicle to accident site as soon as possible. During the decision-making process of emergency vehicle, each action should be aimed at maximizing long-term return on effort, that is, to reach the accident point with the shortest time and the smallest road network impact, and the contribution of route selection decisions to reaching the accident point is quantitatively evaluated by state-action value function:

$$Q_\pi(s, a) = \mathbb{E}_\pi(G_t | s_t = s, a_t = a)$$
$$= \mathbb{E}_\pi(r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \cdots | s_t = s, a_t = a), \quad (2)$$

where $q_\pi(s, a)$ is the value of state-action pair of (s, a), $\mathbb{E}_\pi$ denotes the expectation value of policy π, G_t denotes the cumulative reward value, r_{t+1} refers to the reward at $t + 1$, and γ is the reward decay factor. The optimal strategy for emergency vehicles should have the largest state-action value function under any circumstances to ensure that the decisions it made own largest reward; supposing π_* is the optimal policy, then

$$Q_{\pi_*}(s, a) = \max_\pi Q_\pi(s, a). \quad (3)$$

Based on the optimal policy, the optimal action at current state according to π_* can be selected as follows:

$$\pi_*(a|s) = \begin{cases} 1, & \text{if } a = \arg\max Q_*(s, a), \\ 0, & \text{else.} \end{cases} \quad (4)$$

2.3. Path Planning Based on Prioritized Experience Replay Deep Q-Network. By constructing the MDP environment, we transformed the path planning problem into a decision-optimization problem, to be more specific, it is now a problem about how to make emergency vehicle obtain the optimal policy in path planning on the established urban arterial road model. Based on the optimal policy, the emergency vehicle decides where to go according to current location and reaches the selected position at the next step in path planning dimension. Continuously, emergency vehicle plans its driving route based on the current location until it reaches the accident site. By then, the driving trajectory of emergency vehicle forms its rescue path.

In reinforcement learning, the optimal policy is learned from experience of the interaction with MDP and achieved by estimating optimal action-value function because the action-value function is the basis for the vehicle to determine how to drive at the current node. Only when the values of different actions in different states are known can the vehicle choose the action with largest value as the optimal one.

In PERDQN [24], deep neural network is used to obtain the estimated optimal action-value function as nonlinear function approximator, formulated as $Q(s, a; \theta) \approx Q_{\pi_*}(s, a)$, where θ refers to the parameters of neural network. To be more specific, two neural networks with the same structure including Q-network and target network are constructed,

the Q-network produces $Q(s, a; \theta)$ and evaluates current state-action pair, the target network is used to generate target value as $y = r + \gamma\max_{a'} Q(s', a'; \theta^-)$, θ^- is the parameter of target network, and s' and a' refer to sampled state-action pair. The loss of Q-network is calculated by

$$L(\theta) = \mathbb{E}_{s_t, a_t, r_t}\left[\left(\mathbb{E}_{s_{t+1}}[y | s_t, a_t] - Q(s_t, a_t; \theta)\right)^2\right], \quad (5)$$

which is the difference between estimated value and target value whose gradient is

$$\nabla_\theta L(\theta) = \mathbb{E}_{s_t, a_t, r_t, s_{t+1}}\left[\left(r + \gamma\max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; \theta^-)\right. \right.$$
$$\left.\left. - Q(s_t, a_t; \theta)\right)\nabla_\theta Q(s_t, a_t; \theta)\right]. \quad (6)$$

Through iterative updates, the Q-network approximates the state-action value and gradually learns to be the optimal state-action value function $Q_{\pi_*}(s, a)$. By then, the decision made according to equation (4) is the optimal strategy, which is the optimal path for emergency vehicles.

Moreover, PERDQN introduces the prioritized replay method to improve the performance of experience replay, which is a stochastic priority sampling method. Experience with more information will have a higher probability of being sampled, and the probability of each experience being sampled is

$$P(t) = \frac{p_t^\alpha}{\sum_k p_k^\alpha}, \quad (7)$$

where p_t is the priority of the i th experience and a is used to control the amplitude of priority. Proportional prioritization variant is calculated by

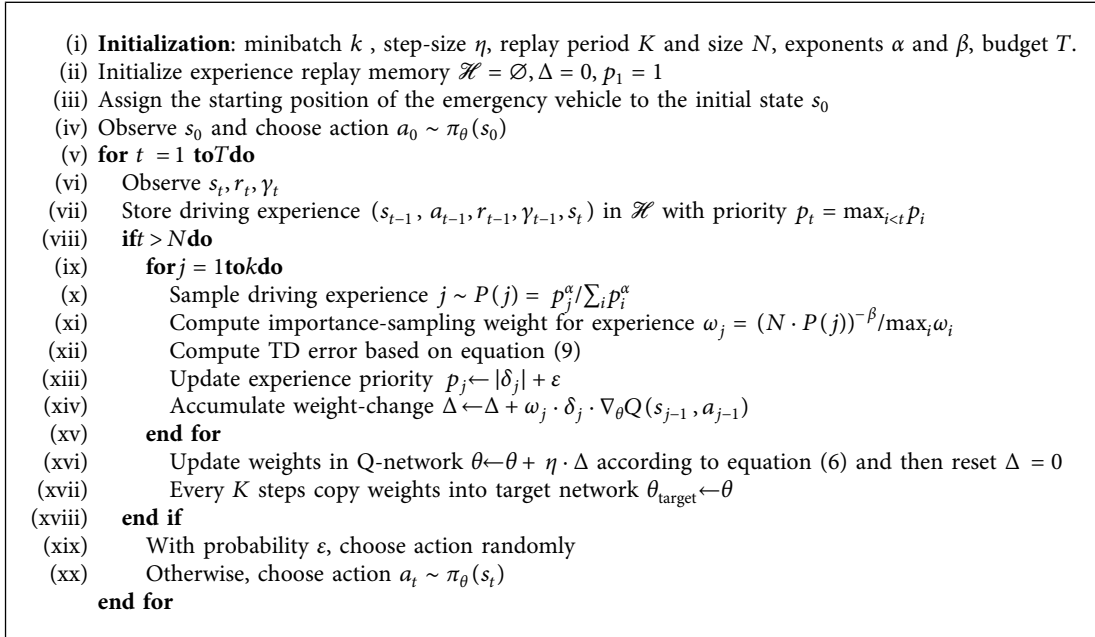$$p_t = |\delta_t| + \varepsilon, \quad (8)$$

where

$$\delta_t = r_t + \gamma_t Q_{\text{target}}(s_t, \arg\max_\alpha Q(s_t, a)) - Q(s_{t-1}, a_{t-1}), \quad (9)$$

where δ_t denotes the TD error in t th experience and ε is a small positive constant that ensures that the edge case of experiences whose TD error is zero can be sampled. Moreover, ε-greedy policy is introduced to PERDQN to ensure the exploration and to avoid local optimal in training.

The pseudocode of PERDQN-based path planning is shown in Algorithm 1.

3. Results and Discussion

3.1. Experiment Setup. East Youyi Road is an east-west urban arterial road located in Xi'an, Shaanxi Province, China. It is a typical urban arterial road with main roads and side roads in both directions, fences existing between the opposite main roads, and intermittent joint point connecting the main roads and side roads in the same direction. Emergency vehicles can change lanes between the main road and the side road in the same direction on the joint point or change to any lane at the intersection. Figure 2 presents the architecture and surrounding environment of East Youyi Road in the section from West Wenyi Road to East Cehui Road.

(i) **Initialization**: minibatch k, step-size η, replay period K and size N, exponents α and β, budget T.
(ii) Initialize experience replay memory $\mathcal{H} = \varnothing, \Delta = 0, p_1 = 1$
(iii) Assign the starting position of the emergency vehicle to the initial state s_0
(iv) Observe s_0 and choose action $a_0 \sim \pi_\theta(s_0)$
(v) **for** $t = 1$ **to** T **do**
(vi)   Observe s_t, r_t, γ_t
(vii)   Store driving experience $(s_{t-1}, a_{t-1}, r_{t-1}, \gamma_{t-1}, s_t)$ in $\mathcal{H}$ with priority $p_t = \max_{i<t} p_i$
(viii)   **if** $t > N$ **do**
(ix)   **for** $j = 1$ **to** k **do**
(x)   Sample driving experience $j \sim P(j) = p_j^\alpha / \sum_i p_i^\alpha$
(xi)   Compute importance-sampling weight for experience $\omega_j = (N \cdot P(j))^{-\beta} / \max_i \omega_i$
(xii)   Compute TD error based on equation (9)
(xiii)   Update experience priority $p_j \leftarrow |\delta_j| + \varepsilon$
(xiv)   Accumulate weight-change $\Delta \leftarrow \Delta + \omega_j \cdot \delta_j \cdot \nabla_\theta Q(s_{j-1}, a_{j-1})$
(xv)   **end for**
(xvi)   Update weights in Q-network $\theta \leftarrow \theta + \eta \cdot \Delta$ according to equation (6) and then reset $\Delta = 0$
(xvii)   Every K steps copy weights into target network $\theta_{\text{target}} \leftarrow \theta$
(xviii)   **end if**
(xix)   With probability ε, choose action randomly
(xx)   Otherwise, choose action $a_t \sim \pi_\theta(s_t)$
    end for

ALGORITHM 1: Rescue path planning for emergency vehicles based on PERDQN.
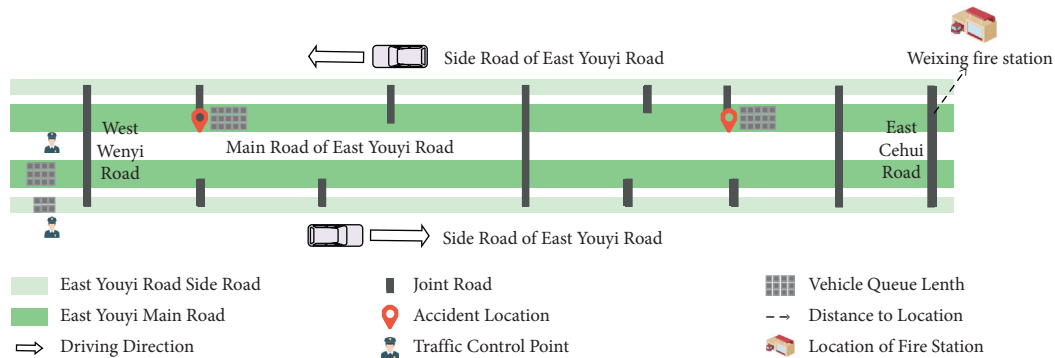

FIGURE 2: Location of East Youyi Road and surrounding environment.

As shown in Figure 2, two traffic accidents were assumed on the main road of East Youyi Road in the east-west direction. Accident case 1 is set to be near West Wenyi Road and is on the left side of the figure, and accident case 2 is near East Cehui Road and is shown on the right side of the figure, and both accidents are placed on the main road in the east-west direction. The nearest fire station to the road section is the Weixing Fire Station located 1.9 km northeast. Therefore, the starting point of emergency rescue vehicles is set to be near the intersection of the east-west main road of East Youyi Road and the intersection of East Cehui Road.

Further, to verify the performance of proposed method, we assume that two traffic accidents occurred at accident point 1 and accident point 2, respectively, causing upstream congestion and vehicle queuing, and the on-site emergency disposal durations are both 15 minutes. Additionally, four typical traffic control schemes conducted by traffic police were applied to the case study to vary the optimal route for different accident points under different traffic control

schemes: scheme 1—reverse main road and side road control (RMSC), scheme 2—reverse main road control (RMC), scheme 3—prograde main road control (PMC), and scheme 4—prograde main road and side road control (PMSC), where prograde is regarded to be driving from east to west because of the starting point of emergency vehicles. Furthermore, the traffic police department is the prioritized department that gets informed after occurrence of traffic accidents and can take responsibilities earlier than any other emergency departments, guaranteeing the road segment at key point being controlled before emergency rescue vehicles from other emergency departments arrive at the accident area. Therefore, under the traffic control at the junction of the East Youyi Road and West Wenyi Road in west-east direction, it is essential for emergency rescue vehicles including firefighters and the ambulances to drive retrograde in the west-east direction of East Youyi Road and reach the accident site faster than driving in the congested section caused by accidents. The upstream section of the accident

site is always under control under any scheme to prevent vehicles from entering the accident area and causing secondary accidents.

The traffic flow parameters on urban arterial roads suggested in [32, 33] are shown in Table 1.

Table 2 demonstrates the traffic density in East Youyi Road from 11:00 am to 13:00 pm under different control measures, which is obtained by collecting a week's data in the actual road sections combined with detailed consultation with the traffic police from Beilin Brigade, Traffic Police Division, Xi'an Public Security Bureau.

3.2. Path Planning Using Reinforcement Learning Approach.

The experiment is conducted on a computer with Intel i5-8300H 2.30 GHz CPU, 8 GB memory, and NVIDIA GTX1060 GPU. We use the same network architecture and hyperparameter settings in both deep reinforcement learning algorithms, whose optimizer is RMSProp with batch size 64 and ε annealed linearly from 1 to 0.0001 with decrease of 0.00005 in each step; both algorithms run 150k steps, and a replay buffer of 10k capacity is used in each algorithm. The structure of neural networks in DQN and PERDQN is also the same, the input layer is consistent with the dimension of state, the first hidden layer is fully connected with 30 neural units, the second fully connected hidden layer owns 15 neural units, and the dimension of the output layer is the same as the shape of action.

Under four typical traffic control schemes, PERDQN and DQN are used for rescue path planning; moreover, the rescue route based on the shortest path (SP) method, which is the most common path plan method in actual rescue path planning, is used for comparison. Figure 2 demonstrates the results of path planning using PERDQN and SP method on the refined road schematic diagram of the East Youyi Road.

Figure 3 shows the node model diagram of the studied road section and optimal path planned by corresponding algorithms. The joint road points, the traffic control places, and the location of accident site are represented by circles, rectangles, and triangles, respectively. The yellow solid lines in separated figures indicate the rescue path planned by corresponding method under different traffic control schemes. (1a)–(1d) in Figure 3 show the optimal path for emergency vehicles under four typical traffic control schemes of case 1 decided by PERDQN, and (1e) in Figure 3 shows the optimal path selected by the shortest route method in case 1. (2a)–(2d) in Figure 3 show the optimal path determined by PERDQN under traffic control schemes in case 2, and (2e) in Figure 3 shows the rescue path of case 2 using the SP method. Figure 3 shows that the path planned by the PERDQN algorithm is different under different traffic control schemes while the path planned by SP remains the same, for the disparity of road traffic efficiency and queue length under different schemes are taken into consideration when using PERDQN. Nevertheless, the SP method always chooses the path with shortest path to the accident site, regardless of which traffic scheme is on, so only a single figure in each case is used for demonstrating the rescue path planned by the SP method. Moreover, the optimal paths of

TABLE 1: Traffic flow parameters for East Youyi Road.

Parameters of traffic flow	Main road	Side road
Critical density	$k_m = 48$ pcu/km	$k_m = 42$ pcu/km
Critical speed	$u_m = 36$ km/h	$u_m = 25$ km/h
Free flow speed	$u_j = 60$ km/h	$u_j = 40$ km/h
Blocking density	$k_j = 100$ pcu/km	$k_j = 95$ pcu/km

TABLE 2: Traffic density on East Youyi Road under different traffic control schemes.

Parameters of traffic flow		Prograde main road (pcu/km)	Prograde side road (pcu/km)	Reverse main road (pcu/km)	Reverse side road (pcu/km)
Normal traffic density		$K_c = 28$	$K_c = 23$	$K_c = 31$	$K_c = 25$
Traffic density in schemes	Scheme 1	$K_s = 82$	$K_s = 75$	$K_s = 98$	$K_s = 79$
	Scheme 2	$K_s = 82$	$K_s = 75$	$K_s = 78$	$K_s = 76$
	Scheme 3	$K_s = 82$	$K_s = 75$	$K_s = 31$	$K_s = 25$
	Scheme 4	$K_s = 82$	$K_s = 77$	$K_s = 31$	$K_s = 25$

different methods under four traffic control schemes are summarized and listed in Table 3; note that L2 denotes P11 in case 1 and P15 in case 2, while L1 is P8, L3 denotes P17, and L4 represents P25 in both cases. The corresponding reinforcement learning training curves in case 1 are shown in Figure 4.

Table 3 shows the optimal paths planned by the PERDQN and SP method, the routes under different control schemes are distinct, and the trajectory diagram of optimal paths is drawn in Figure 3. Figure 4 shows the training curve of case 1, where Figures 4(a)–(d) show the training curves of PERDQN and DQN under four schemes. The horizontal axis is the number of training steps and the vertical axis represents the average reward. In the first 10k steps, average rewards rarely changed because both algorithms do not update their neural network parameters until replay buffer is first fulfilled. Moreover, the existence of prioritized replay experience enhances the sampling effectiveness of experience and makes PERDQN achieve better decision-making effects than DQN, which also prevents PERDQN falling into local optimal that DQN would do, resulting in paths planned by PERDQN being better than those planned by DQN. Figure 4(e) shows the comparison of the training curves of PERDQN under different schemes. It shows that scheme 2 owns the highest average reward, which indicates that the optimal rescue path for case 1 planned by PERDQN is P16-P24-P23-P22-P21-P20-P19-P18-P17-P9-P10 using the RMC scheme, that is, to retrograde through the opposite main road to reach the accident site.

Using evaluation metrics suggested in [28], the arrival time and vehicle queue length of optimal path under different traffic control schemes in case 1 are shown in Figure 5, where the arrival time, dissipation time, and queue length
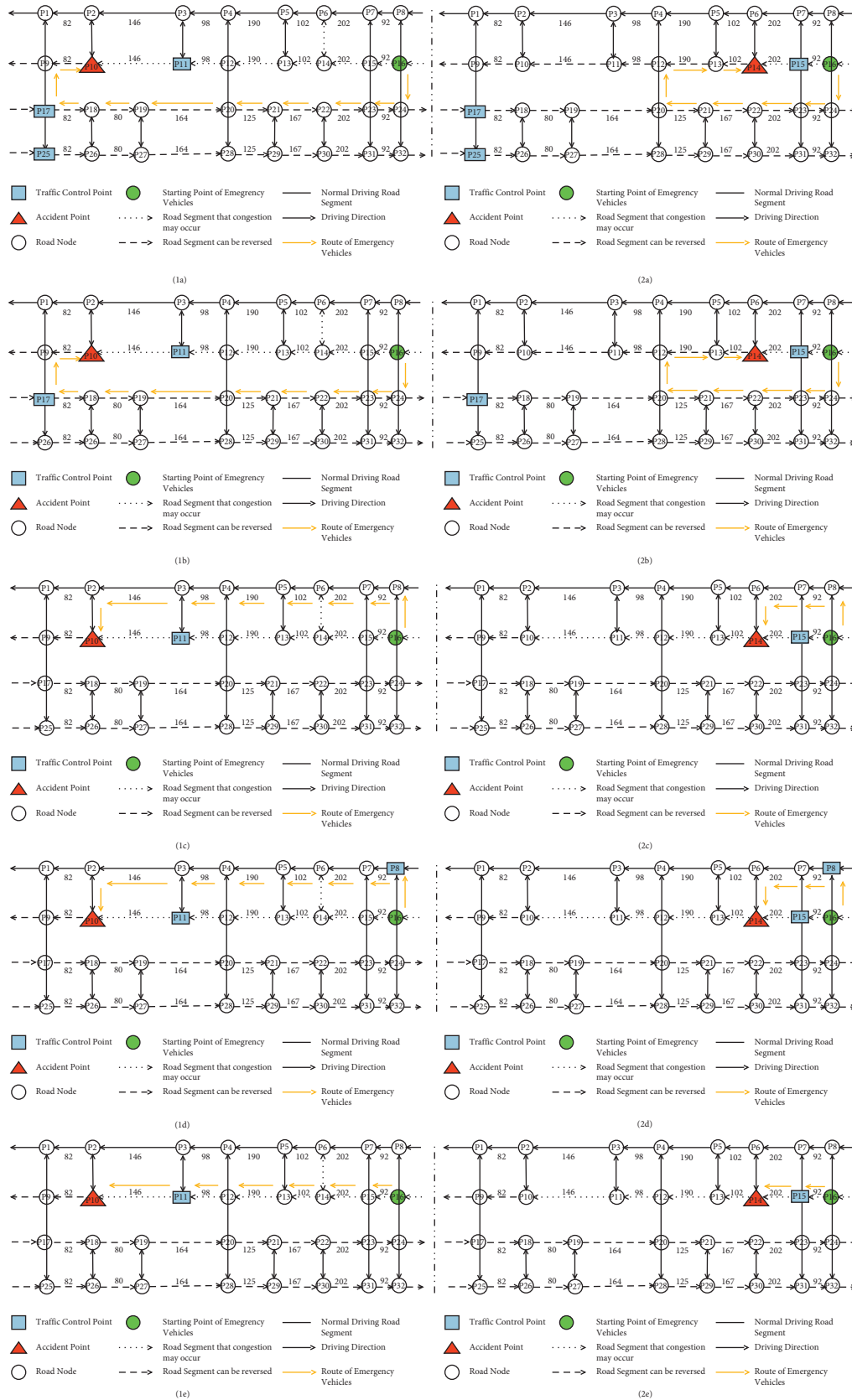
Figure 3: Abstracted schematic diagram of road segment model and optimal rescue path planned by PERDQN and SP method. (1a) Optimal rescue path planned by PERDQN in case 1 under scheme 1. (2a) Optimal rescue path planned by PERDQN in case 2 under scheme 1. (1b) Optimal rescue path planned by PERDQN in case 1 under scheme 2. (2b) Optimal rescue path planned by PERDQN in case 2 under scheme 2. (1c) Optimal rescue path planned by PERDQN in case 1 under scheme 3. (2c) Optimal rescue path planned by PERDQN in case 2 under scheme 3. (1d) Optimal rescue path planned by PERDQN in case 1 under scheme 4. (2d) Optimal rescue path planned by PERDQN in case 2 under scheme 4. (1e) Optimal rescue path planned by SP in case 1. (2e) Optimal rescue path planned by SP in case 2.

TABLE 3: Three typical schemes for emergency vehicle route planning on East Youyi Road.

Scheme	Traffic control schemes	Driving trajectory
Scheme 1 (RMSC + PERDQN)	Traffic control at L2, L3, and L4: emergency vehicles driving in reverse in the opposite lane of the accident site	Case 1: P16-P24-P23-P22-P21-P20-P19-P18-P17-P9-P10 Case 2: P16-P24-P23-P22-P21-P20-P12-P13-P14
Scheme 2 (RMC + PERDQN)	Traffic control at L2 and L3: emergency vehicles driving in reverse in the opposite lane of the accident site	**Case 1: P16-P24-P23-P22-P21-P20-P19-P18-P17-P9-P10** Case 2: P16-P24-P23-P22-P21-P20-P12-P13-P14
Scheme 3 (PMC + PERDQN)	Traffic control at L2: no traffic control at opposite lane of traffic accident	Case 1: P16-P8-P7-P6-P5-P4-P3-P2-P10 Case 2: P16-P8-P7-P6-P14
Scheme 4 (PMSC + PERDQN)	Traffic control at L1 and L2: no traffic control at opposite lane of traffic accident	Case 1: P16-P8-P7-P6-P5-P4-P3-P2-P10 **Case 2: P16-P8-P7-P6-P14**
SP method	Traffic control at L2: no traffic control at opposite lane of traffic accident	Case 1: P16-P15-P14-P13-P12-P11-P10 Case 2: P16-P15-P14

(a)

(b)

(c)
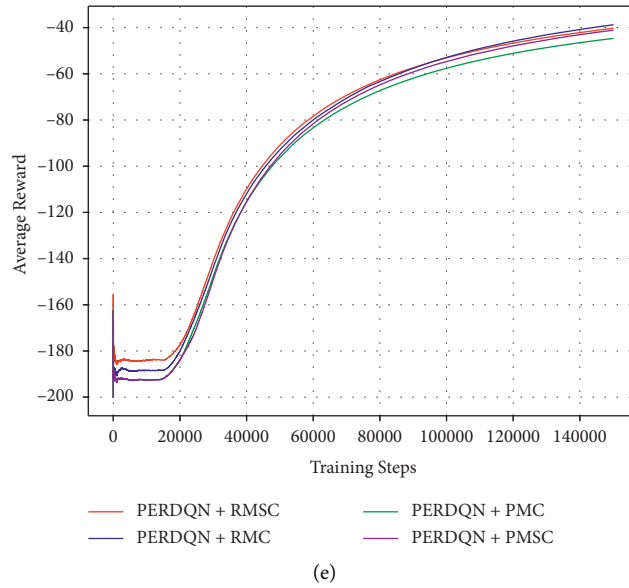

(d)
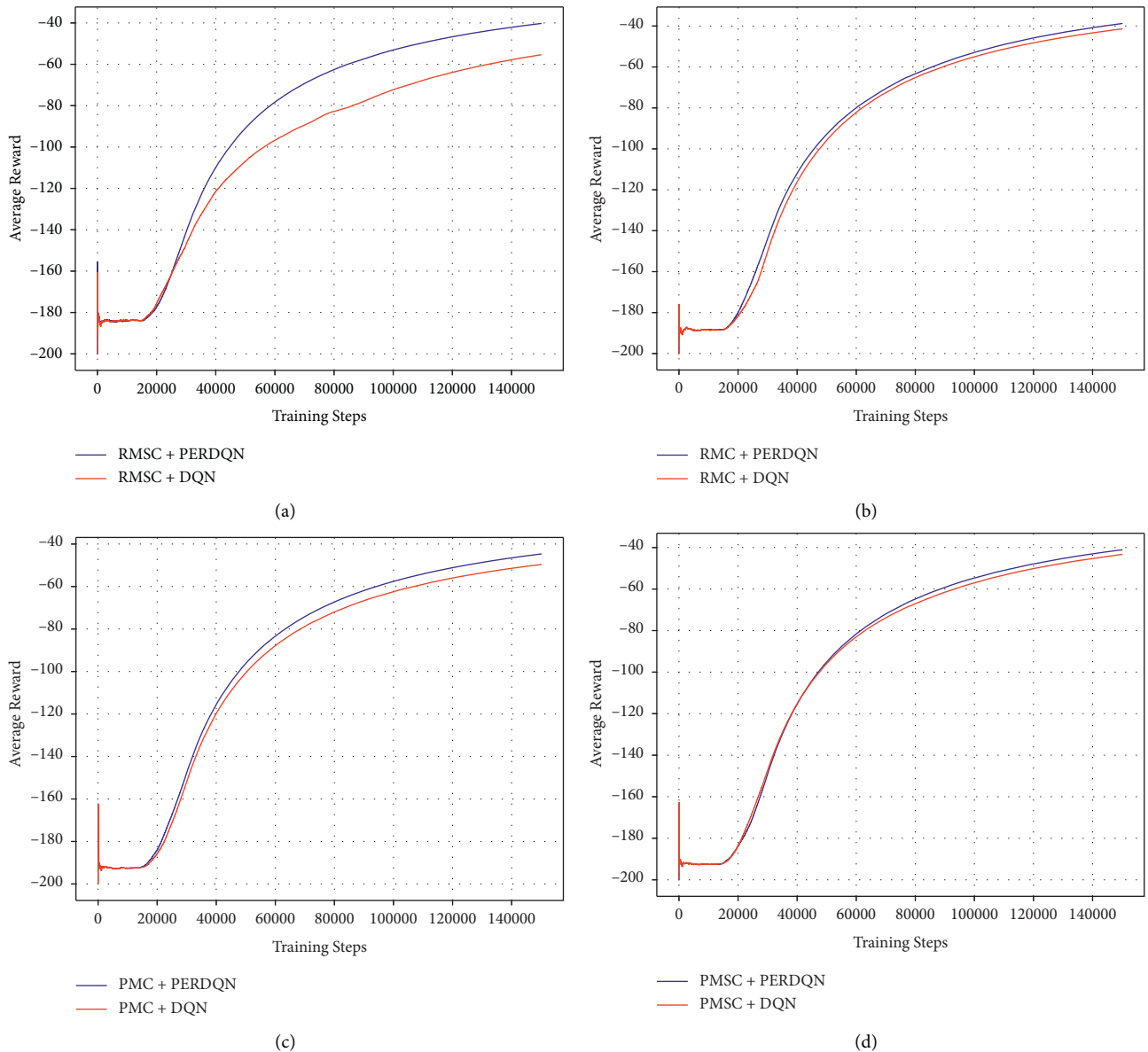
FIGURE 4: Continued.

(e)

Figure 4: Average reward of PERDQN and DQN among training steps under different schemes in case 1. (a) Average reward under scheme 1. (b) Average reward under scheme 2. (c) Average reward under scheme 3. (d) Average reward under scheme 4. (e) Average reward comparison of PERDQN among four schemes.
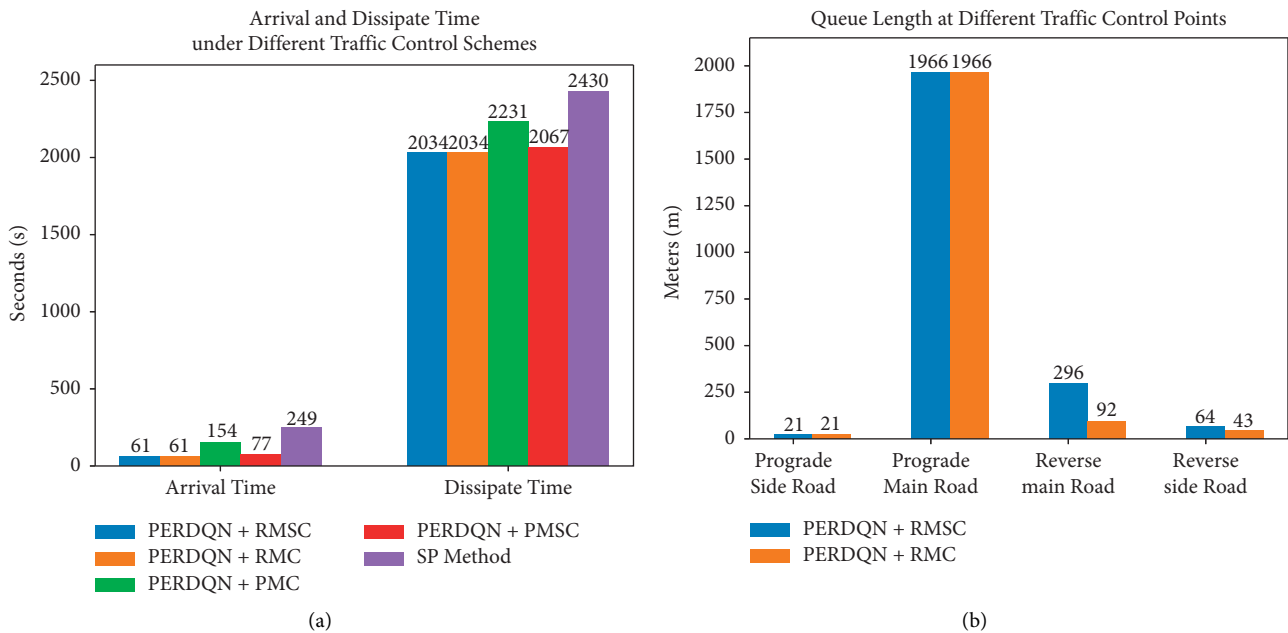

(a)

(b)

Figure 5: Evaluation parameters of different paths planned by PERDQN and SP under different traffic control schemes in case 1. (a) The arrival time and dissipation time of vehicle queue under different traffic control schemes. (b) Queue length of scheme 1 and scheme 2 at different traffic control points.

are calculated by traffic wave theory and BPR function mentioned in Section 2.2; the traffic flow parameters on arterial road are listed in Tables 1 and 2. As demonstrated in Figure 5, the dissipation time is far longer than the arrival time because the duration of accident on-site disposal is included in the dissipation. The queue length on prograde main road is much larger than that on other three traffic control points because traffic control at L1, L3, and L4 is

released as soon as the emergency vehicles reach the accident site while traffic control at L2 sustains until the on-site accident disposal is accomplished. According to Figure 5(a), the paths planned under schemes 1 and 2 using PERDQN have the same arrival time and dissipation time, so the queue lengths of four traffic control points under schemes 1 and 2 are compared in Figure 5(b); through comparison, it can be found that optimal path under scheme 2 has the shortest
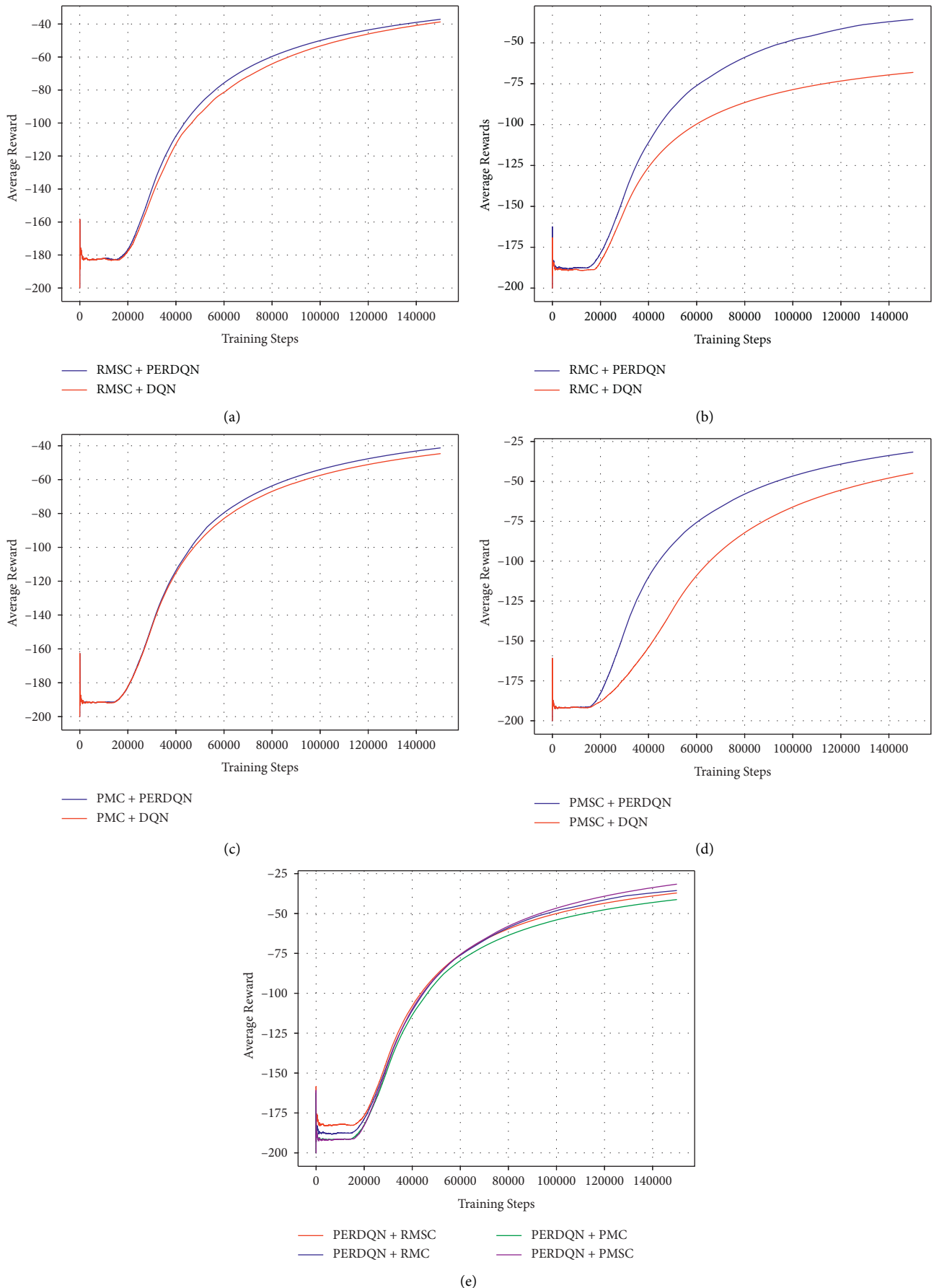
(a)

(b)

(c)

(d)

(e)

Figure 6: Average reward of PERDQN and DQN among training steps under different schemes in case 2. (a) Average reward under scheme 1. (b) Average reward under scheme 2. (c) Average reward under scheme 3. (d) Average reward under scheme 4. (e) Average reward comparison of PERDQN among four schemes.
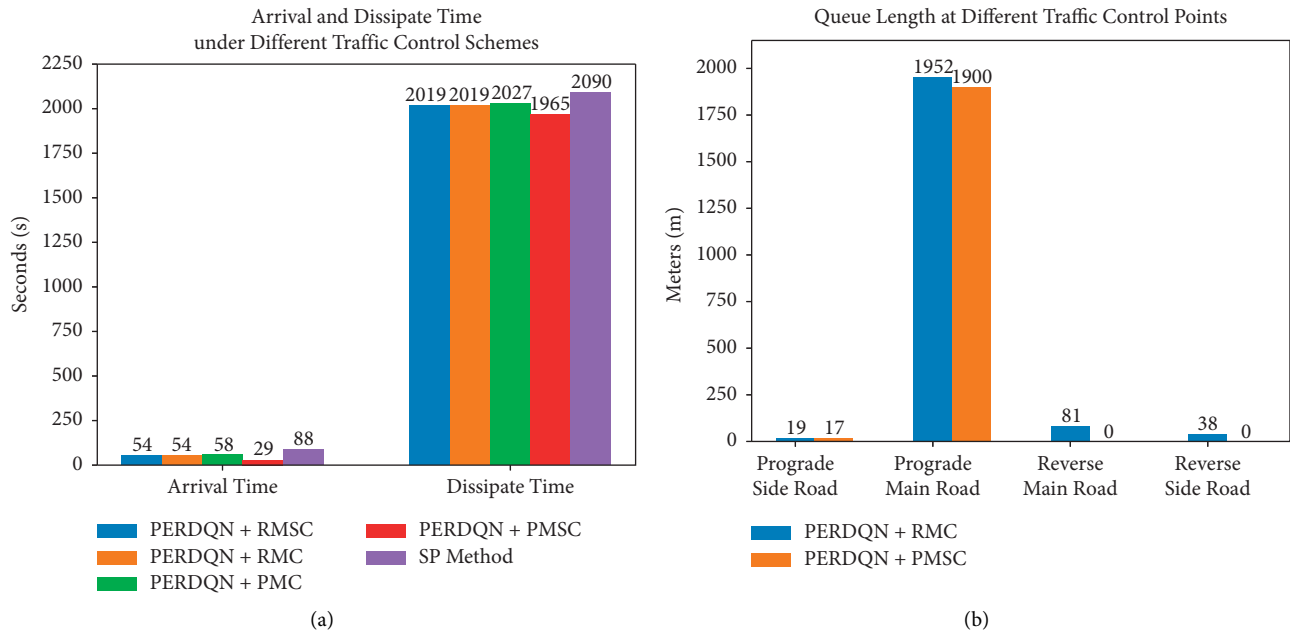
FIGURE 7: Evaluation parameters of different paths planned by PERDQN and SP under different traffic control schemes in case 2. (a) The arrival time and dissipation time of vehicle queue for different traffic control schemes. (b) Queue length of scheme 2 and scheme 4 at different traffic control points.

arrival time, and compared with scheme 1 which has the same arrival time, the length of the vehicle queue is less, which indicates that by driving on the optimal path suggested in scheme 2 by PERDQN, the emergency vehicles can reach the accident site in the shortest time with least queuing length.

Figure 6 shows the training curve of PERDQN and DQN in case 2 under different traffic control schemes, which shows the superiority of PERDQN when compared with DQN, and also shows that scheme 4 leads to the optimal rescue path among four schemes. Figure 7 shows the arrival time, dissipation time of queue, and vehicle queue length at four traffic control points. It is obvious that following the rescue path planned by PERDQN in scheme 4 would lead to a minimum arrival time and least queue length. It is also believed that the trajectory of emergency vehicle passing P16-P8-P7-P6-P14 would be the optimal path for rescuing in case 2.

4. Conclusions

This paper proposes a refined path planning method for emergency vehicles in congested urban arterial road networks based on reinforcement learning algorithm. By abstracting the position of the road nodes and the length of the road segments, a MDP model that describes decision making for path planning is established, and regarding the travel efficiency and the impact on the road network under different traffic control schemes during accidents, the PERDQN algorithm is introduced to make decision for path planning. Taking traffic efficiency and road network impact into account and paying special attention to the congestions caused by accidents and traffic control schemes, the proposed method is capable to provide optimal path plan for emergency vehicles to reach the traffic accident site on urban arterial roads with shortest time and least road queuing length.

Based on our proposed method, our future works include extending current study to urban road with longer distance, considering the path planning with multiple rescue points, and improving the performance of the path planning algorithms.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Xu, K. An, Z. Ye, Y. Wang, J. Feng, and J. Zhao, "A bi-level model to resolve conflicting transit priority requests at urban arterials," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1353–1364, 2019.

[2] E. Cigu, D. T. Agheorghiesei, S. A. F. Gavriluta, and E. Toader, "Transport infrastructure development, public performance and long-run economic growth: a case study for the Eu-28 countries," *Sustainability*, vol. 11, no. 1, p. 22, 2019.

[3] A. Banerjee, E. Duflo, and N. Qian, "On the road: access to transportation infrastructure and economic growth in China," *Journal of Development Economics*, vol. 145, Article ID 102442, 2020.

[4] X. Chen, Z. Li, Y. Yang, L. Qi, and R. Ke, "High-resolution vehicle trajectory extraction and denoising from aerial videos," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 3190–3202, 2021.

[5] Z. Wang, M. Lu, X. Yuan, J. Zhang, and H. Van De Wetering, "Visual traffic jam analysis based on trajectory data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2159–2168, 2013.

[6] Y. Zhang, Y. Zhang, and R. Su, "Pedestrian-safety-aware traffic light control strategy for urban traffic congestion alleviation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 1, pp. 178–193, 2021.

[7] Z. Zheng, Z. Wang, L. Zhu, and H. Jiang, "Determinants of the congestion caused by a traffic accident in urban road networks," *Accident Analysis and Prevention*, vol. 136, Article ID 105327, 2020.

[8] P. Wang, W. Hao, and Y. Jin, "Fine-grained traffic flow prediction of various vehicle types via fusion of multisource data and deep learning approaches," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2020.

[9] B. L. Ye, W. Wu, and W. Mao, "A two-way arterial signal coordination method with queueing process considered," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3440–3452, 2015.

[10] Z. Kadim, K. M. Johari, D. F. Samaon, Y. S. Li, and H. W. Hon, "Real-time deep-learning based traffic volume count for high-traffic urban arterial roads," in *Proceedings of the 2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, pp. 53–58, Penang, Malaysia, April 2020.

[11] P. Wang, J. Yang, Y. Jin et al., "Research on allocation and dispatching strategies of rescue vehicles in emergency situation on the freeway," in *Proceedings of the 16th IEEE International Conference on Control, Automation, Robotics and Vision*, pp. 130–135, Shenzhen, China, December 2020.

[12] F. Zheng, H. J. van Zuylen, X. Liu, and S. Le Vine, "Reliability-based traffic signal control for urban arterial roads," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 643–655, 2017.

[13] B. Yang, Z. Ding, L. Yuan, J. Yan, L. Guo, and Z. Cai, "A novel urban emergency path planning method based on vector grid map," *IEEE Access*, vol. 8, pp. 154338–154353, 2020.

[14] M. R. Jabbarpour, H. Zarrabi, J. J. Jung, and P. Kim, "A green ant-based method for path planning of unmanned ground vehicles," *IEEE Access*, vol. 5, pp. 1820–1832, 2017.

[15] C. Wu, S. Zhou, and L. Xiao, "Dynamic path planning based on improved ant colony algorithm in traffic congestion," *IEEE Access*, vol. 8, pp. 180773–180783, 2020.

[16] Y. Zhang, Y. L. Hsueh, W. C. Lee, and Y. H. Jhang, "Efficient cache-supported path planning on roads," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 4, pp. 951–964, 2016.

[17] M. Karoui, G. Chalhoub, and A. Freitas, "An efficient path planning GLOSA-based approach over large scale and realistic traffic scenario," *Internet Technology Letters*, vol. 4, no. 4, 2020.

[18] L. Hu, Z. Liu, W. Hu, Y. Wang, J. Tan, and F. Wu, "Petri-net-based dynamic scheduling of flexible manufacturing system via deep reinforcement learning with graph convolutional network," *Journal of Manufacturing Systems*, vol. 55, pp. 1–14, 2020.

[19] Y. Man, Y. Huang, J. Feng, X. Li, and F. Wu, "Deep Q learning driven CT pancreas segmentation with geometry-aware U-net," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1971–1980, 2019.

[20] U. F. Siddiqi, S. M. Sait, and M. Uysal, "Deep Q-learning based optimization of VLC systems with dynamic time-division multiplexing," *IEEE Access*, vol. 8, pp. 120375–120387, 2020.

[21] X. H. Liu, D. G. Zhang, H. R. Yan, Y. Y. Cui, and L. Chen, "A new algorithm of the best path selection based on machine learning," *IEEE Access*, vol. 7, pp. 126913–126928, 2019.

[22] C. Chen, J. Jiang, N. Lv, and S. Li, "An intelligent path planning scheme of autonomous vehicles platoon using deep reinforcement learning on network edge," *IEEE Access*, vol. 8, pp. 99059–99069, 2020.

[23] P. Sun, Z. Guo, J. Lan et al., "ScaleDRL: a scalable deep reinforcement learning approach for traffic Engineering in SDN with pinning control," *Computer Networks*, vol. 190, 2021.

[24] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized experience replay," 2015, http://arxiv.org/abs/1511.05952.

[25] X. Chen, L. Qi, Y. Yang, Q. Luo et al., "Video-based detection infrastructure enhancement for automated ship recognition and behavior analysis," *Journal of Advanced Transportation*, vol. 2020, Article ID 7194342, 12 pages, 2020.

[26] D. Ding, J. Li, P. Tu, H. Wang, T. Cao, and F. Zhang, "Electric vehicle charging warning and path planning method based on spark," *IEEE Access*, vol. 8, pp. 8543–8553, 2020.

[27] X. Liu, X. Qu, and X. Ma, "Improving flex-route transit services with modular autonomous vehicles," *Transportation Research Part E: Logistics and Transportation Review*, vol. 149, Article ID 102331, 2021.

[28] J. Yang, P. Wang, W. Yuan, Y. Ju, W. Han, and J. Zhao, "Automatic generation of optimal road trajectory for the rescue vehicle in case of emergency on mountain freeway using reinforcement learning approach," *IET Intelligent Transport Systems*, vol. 15, no. 9, pp. 1142–1152, 2021.

[29] M. Carey and M. Bowers, "A review of properties of flow-density functions," *Transport Reviews*, vol. 32, no. 1, pp. 49–73, 2012.

[30] B. Li, Z. Xu, and Y. Zhang, "Two-stage multi-sided matching dispatching models based on improved BPR function with probabilistic linguistic term sets," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 2, 2020.

[31] D. D. Zhu and J.-Q. Sun, "A new algorithm based on dijkstra for vehicle path planning considering intersection attribute," *IEEE Access*, vol. 9, pp. 19761–19775, 2021.

[32] National Research Council, Transportation Research Board, *Highway Capacity Manual 2010*, Transportation Research Board of the National Academies, Washington D. C, USA, 2010.

[33] Y. Zhang, *Highway Capacity*, China Building Industry Press, Peking, China, 2016.