


Article

Collaborative Decision-Making Method of Emergency Response for Highway Incidents

Junfeng Yao ^{1,2,†}, Longhao Yan ^{3,†}, Zhuohang Xu ³ , Ping Wang ^{4,*} and Xiangmo Zhao ^{1,5}¹ School of Information Engineering, Chang'an University, Xi'an 710064, China² China Communications Information & Technology Group Co., Ltd., Beijing 100088, China³ School of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China⁴ School of Intelligent Systems Engineering, Sun Yat-sen University, Guangzhou 510006, China⁵ School of Electronic Information Engineering, Xi'an Technological University, Xi'an 710021, China

* Correspondence: wang0372@e.ntu.edu.sg; Tel.: +86-15353622304

† These authors contributed equally to this work.

Abstract: With the continuous increase in highway mileage and vehicles in China, highway accidents are also increasing year by year. However, the on-site disposal procedures of highway accidents are complex, which makes it difficult for the emergency department to fully observe the accident scene, resulting in the lack of sufficient communication and cooperation between multiple emergency departments, making the rescue efficiency low and wasting valuable rescue time, and causing unnecessary injury or loss of life due to the lack of timely assistance. Thus, this paper proposes a multi-agent-based collaborative emergency-decision-making algorithm for traffic accident on-site disposal. Firstly, based on the analysis and abstraction of highway surveillance videos obtained from the Shaanxi Provincial Highway Administration, this paper constructs an emergency disposal model based on Petri net to simulate the emergency on-site disposal procedures. After transforming the emergency disposal model into a Markov game model and applying it to the multi-agent deep deterministic strategy gradient (MADDPG) algorithm proposed in this paper, the multiple agents can optimize the emergency-decision-making and on-site disposal procedures through interactive learning with the environment. Finally, the proposed algorithm is compared with the typical algorithm and the actual processing procedure in the simulation experiment of an actual Shaanxi highway traffic accident. The results show that the proposed emergency-decision-making method could greatly improve collaboration efficiency among emergency departments and effectively reduce emergency response time. This algorithm is not only superior to other decision-making algorithms such as genetic algorithm (EA), evolutionary strategy (ES), and deep Q network (DQN), but also reduces the disposal processes by 28%, 28%, and 42%, respectively, compared with the actual disposal process in three emergency disposal cases. In summary, with the continuous development of information technology and highway management systems, the multi-agent-based collaborative emergency-decision-making algorithm will contribute to the actual emergency response process and emergency disposal in the future, improving rescue efficiency and ensuring the safety of individuals.

Keywords: traffic engineering; emergency decision making; multi-agent deep reinforcement learning; traffic accident; Petri net; Markov game



Citation: Yao, J.; Yan, L.; Xu, Z.; Wang, P.; Zhao, X. Collaborative Decision-Making Method of Emergency Response for Highway Incidents. *Sustainability* **2023**, *15*, 2099. <https://doi.org/10.3390/su15032099>

Academic Editor: Shaojian Qu

Received: 18 December 2022

Revised: 13 January 2023

Accepted: 14 January 2023

Published: 22 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the rapid development an increasing scale of freeways has brought huge economic benefits to society and also provided great convenience for individuals. However, with the increase in the number of vehicles and the driving distance, freeway traffic accidents occur frequently [1,2]. Research has shown that there is a significant spatial autocorrelation between traffic accidents in various provinces and cities in China, which means that the number of traffic accidents and deaths is high with high aggregation, which

leads to traffic congestion in populated areas, resulting in the waste of resources and inability to reflect sustainable development. Traffic accidents will not only threaten people's safety; improper emergency response will increase rescue time, cause secondary damage to the accident site, waste resources, and hinder sustainable development [3]. In the actual emergency disposal process of highway emergencies, the accident site is often complex, the response time of the emergency department is urgent, and there is a lack of adequate communication between multiple emergency departments, which makes the emergency departments unable to fully understand the progress of the accident site, resulting in poor cooperation between the emergency departments and low emergency rescue efficiency. Therefore, the emergency response process can be optimized to improve the emergency rescue efficiency, and to better protect people's lives [4,5]. Thus, an emergency response decision-making method is crucial to improve inter-department cooperation and optimize the emergency response process. On this basis, how to obtain timely and accurate emergency decision making, as well as reduce casualties and economic losses effectively, is also an urgent subject to study [6].

As one of the typical means to improve traffic safety, emergency decision making (EDM) in road traffic accident emergency response has been highly valued by road safety researchers. A large number of researchers have tried to use various methods to formulate emergency-decision-making procedures to improve the safety of road traffic: Khan conducted research on sustainable traffic safety in Hong Kong based on the Bayesian network method, in which accident risk factors are introduced into the analysis of real accident reports [7]; Ding proposed a zero-sum game method containing Pythagorean fuzzy uncertain linguistic variables, which improved the fuzziness of decision makers' evaluation data in emergency situations for road emergency decision making [8]; J. Pérez González has developed an analytical network platform to conduct time-space distribution statistics and display of the accident distribution information, thus providing information assistance and support for the decision making of emergency personnel [9]; Wang's interval dynamic reference point method based on prospect theory introduces psychological factors into the consideration of emergency decision making, effectively improving the rationality of emergency decision making [10]. The above research has made unique contributions to the field of emergency decision making, but there are few studies on the emergency disposal processes of road accident scenes, especially regarding the team cooperation of various departments and the emergency disposal process, which need to be further improved [11–13].

The mathematical modeling of road emergency scene disposal is the basis of emergency decision making research. How to retain the characteristics of highway traffic accident emergency disposal while simulating the evolution of accident scenes is the focus of modeling research. Nowadays, a large number of researchers at home and abroad have put forward their own solutions to the modeling of emergency accidents: Song et al. have considered the actual situation and features of emergency traffic evacuation, and developed maximum covering models [14]; Qi L et al. used hybrid Petri nets to comprehensively and systematically simulate the situation and response of emergency accidents, discrete events and the evolution of emergency scenarios being clearly depicted, and with the help of Markov chain and fuzzy mathematics, they analyzed the performance of the traffic emergency command system, which greatly improved the rescue efficiency of the system [15,16]; based on Markov chains and the evolution mechanism of emergencies, Sun et al. built a situation assessment model for emergency responses to highway traffic accidents, and used it to infer the probability of traffic accidents under different conditions [17]. The above research often regards the uncertainty and dynamic of the emergency scenario evolution at the accident site as a dynamic sequential decision-making problem based on "scenario disposal". However, considering that accident disposal often involves multiple emergency response departments, the on-site disposal of highway accidents should be considered as a multi-agent cooperation problem. Therefore, as a sequential decision-making model involving multiple objects, the Markov game method provides a reliable theoretical basis

and potential solutions for building a decision-making model for the emergency disposal of traffic accidents.

On the basis of building a model that can simulate the evolution law of the accident scene, the design of supporting collaborative emergency response algorithms has become the focus of further research. To date, a variety of intelligent decision-making algorithms, including genetic algorithms, and evolutionary strategies have been applied to various decision-making problems [18,19]. With the rapid development of artificial intelligence algorithms, decision-making algorithms represented by reinforcement learning have provided solutions for emergency-decision-making and other problems [20,21]. Among them, the multi-agent reinforcement learning algorithm is particularly suitable for solving the emergency disposal problem at an accident site. In fact, the multi-agent reinforcement learning algorithm has been widely used in the control and decision-making problems of intelligent transportation systems. Wang et al. proposed a new multi-agent reinforcement learning framework based on a collaboration group, which can effectively control the large-scale road network through the collaborative vehicle infrastructure system and effectively alleviate the congestion at multiple intersections [22]. On the basis of a highly scalable independent double Q-learning method based on double estimators and the upper confidence limit strategy, Wang et al. provided a solution for finding the optimal signal timing strategy in large-scale traffic signal control [23]. Guo et al. proposed an improved near-end strategy optimization algorithm to improve vehicle congestion in the traffic system. By adaptively adjusting the algorithm's super parameters and limiting the update range of the strategy, the algorithm's decision optimization capability is improved as much as possible while ensuring the robustness of the algorithm [24]. The above research shows the wide applicability and decision-making ability of a multi-agent in-depth reinforcement learning algorithm in the transportation system. Considering the strong demand for stability in highway emergencies, due to its high stability and strong performance, the multi-agent in-depth deterministic strategy gradient (MADDPG) algorithm shows natural advantages [25–27]. The introduction of this algorithm induces the emergency-decision-making algorithm to improve the efficiency of emergency disposal decision making under a high degree of stability, and contributes towards coordinating the emergency behavior of various departments.

The rest of this paper is organized as follows: The second section describes how to use Petri nets to establish a traffic emergency response decision-making model to simulate the process of highway emergency disposal. The third section proposes an emergency-decision-making method based on a multi-agent reinforcement learning framework, and explains how to use this method to coordinate the emergency responses of different departments in detail. The fourth section shows how to use the proposed modeling and decision-making methods to optimize decision making in real traffic accident disposal cases, and analyzes the results. The fifth section summarizes the contributions of this paper and the conclusions derived from the experimental results, pointing out the limitations of existing research and potential future research directions.

2. Modeling of a Highway Emergency Scene Disposal Process Based on Petri Net

Based on the analysis of emergency response monitoring videos and a large number of written reports of highway emergencies from February to September 2019 obtained from the Shaanxi Highway Toll Center, as well as discussions with emergency experts from the Qinling Highway Management Office, this paper establishes a mathematical model based on Petri net that can subdivide the functions of various emergency departments and simulate the on-site emergency disposal process. The emergency disposal process at the accident site is regarded as the interactive process between the emergency response department and the emergency disposal at the accident site. The accident site is constantly changing under the influence of the response behavior exhibited by the emergency department, and the complete process from the occurrence of the accident to the completion of on-site disposal can be described as the emergency task. With the model construction method utilizing the flow chart style, the disposal process at the traffic accident site can be simulated.

Petri net is a mathematical representation of discrete parallel systems that is suitable for describing asynchronous and concurrent computer system models. Petri net has both strict mathematical representation and intuitive graphical representation; it also has rich system description means and system behavior analysis techniques, providing a solid conceptual basis for the simulation of multiple systems in the physical world. The place, transition, flow relationship, and token are the basic components of Petri net. In the mathematical model established in this paper, these elements represent the on-site accident disposal situation, the emergency disposal behavior of various departments, and the evolution principle of the emergency scenario, where the flow relationship specifies the impact relationship between emergency tasks and various response behaviors, and the token indicates the current state of the accident disposal process.

Petri nets are often expressed as $N_e = [P, T, R_{pre}, R_{post}, C]$, where P is the collection of places and represents the emergency disposal task at the accident site; T is the transition set, representing the disposal behavior of the corresponding department; C is the token set of the model, and different token locations distinguish the situation of the accident; R_{pre} and R_{post} are the forward and backward correlation matrices of Petri nets, respectively. The combination of the two defines the flow relationship of Petri nets, that is, the evolution rules of accidents. In the Petri net constructed in this paper, the place is used to represent emergency tasks in an emergency response, including a series of typical emergency task processes in accident disposal such as casualty rescue, fire rescue, road cleaning, etc., which are abstracted. In addition, the completion of each emergency task is distinguished by numerical value. Taking the fire suppression task as an example, the on-site fire can be divided into three conditions: no fire, fire existed, and fire spread, respectively corresponding to the repository in the Petri net p_{14} , p_{15} , and p_{13} . According to the different description objects, the emergency disposal at the accident site is abstracted as a collection of six dimensions, including people, vehicle, road, environment, road facilities, and accident information, and multiple emergency tasks are extracted from the actual on-site emergency disposal process; their corresponding situation with place is shown in Table 1.

Table 1. Mapping between places and emergency tasks.

Emergency Tasks	Place	Description
People		
Emergency site evacuation	p_1	Crowd is chaotic
	p_2	Crowd to be evacuated
	p_3	Crowd has been evacuated
Rescue of accident victims	p_4	The victims need rescue
	p_5	The victims are rescued
Care of accident victims	p_6	Rescue workers injured
	p_7	The victims are dead or injured
	p_8	The victims were treated
Vehicle		
Explosion risk	p_9	Risk of explosion
	p_{10}	No explosion risk
Cause of fire	p_{11}	Fire fighting plan is required
	p_{12}	Fire fighting plan existed
Fire conditions	p_{13}	Fire spread at the accident site
	p_{14}	Fire existed at accident site
	p_{15}	No fire at the emergency site
Vehicle damage	p_{16}	Damaged vehicles are not handled
	p_{17}	Damaged vehicles are handled
Traffic		
Traffic control	p_{18}	No traffic control
	p_{19}	Temporary traffic control

Table 1. Cont.

Emergency Tasks	Place	Description
Traffic control	p_{20}	Blockage at the accident site
	p_{21}	No blockage at the accident site
	Environment	
Road cleaning	p_{22}	Sundries on site need to be cleaned
	p_{23}	No sundries on site
Site smoke	p_{24}	Smoke spread
	p_{25}	Smoke existed
	p_{26}	No smoke
	Road facilities	
Facility damage	p_{27}	Road facilities are damaged
	p_{28}	Road facilities have been repaired
Escape routes	p_{29}	Escape routes closed
	p_{30}	Escape routes opened
Information board	p_{31}	Information not published
	p_{32}	Information published
	Accident information	
Accident report	p_{33}	Site conditions not reported
	p_{34}	Site situation has been reported
Accident record	p_{35}	The accident is not recorded
	p_{36}	The accident is recorded

In the disposal process of highway traffic accidents, different emergency response departments have unique emergency response responsibilities. In order to improve the efficiency of emergency disposal and avoid secondary accidents caused by misoperation, various emergency disposal tasks at the accident site need to be completed by a special emergency response crew. For example, facing a traffic accident with fire, the fire can only be extinguished by the fire department at the accident site using fire-fighting equipment, while other departments' hasty attempts to extinguish the fire may exacerbate the fire due to improper extinguishing methods, thus hindering the disposal process of the accident. In order to effectively simulate the response responsibilities of different departments, this paper uses the transition function in Petri net to describe the specific response responsibilities of the emergency department. By sector, Table 2 shows the corresponding relationship between a variety of typical emergency site response behaviors and transitions.

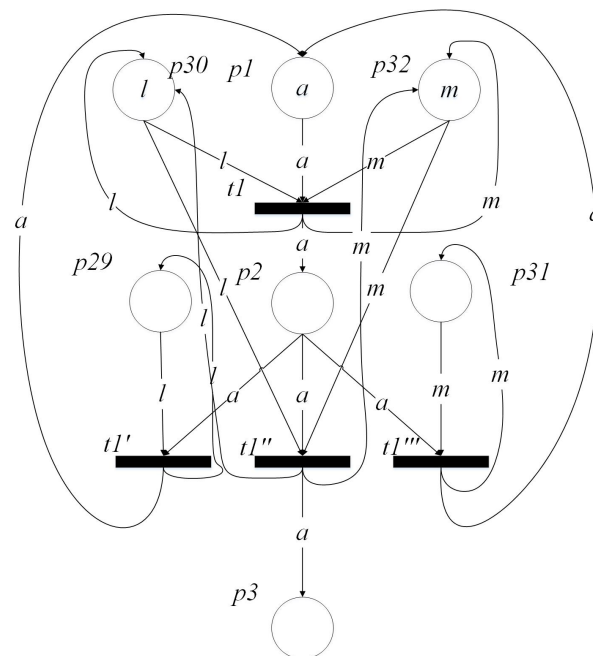
Table 2. Mapping between transitions and emergency actions.

Transition	Description
Road Administration Department	
t_1	Evacuate the crowd
t_2	Report the accident
t_3	Early fire fighting
Traffic Police Department	
t_4	Block the accident site
t_5	Record the accident site
t_6	Deregulate the accident site
Fire department	
t_7	Develop fire fighting plan
t_8	Fire fighting
t_9	Dismantle damaged vehicle
Emergency medical department	
t_{10}	Rescue the injured
t_{11}	Emergency treatment
Road clearance department	
t_{12}	Tow the damaged vehicle
Road maintenance department	
t_{13}	Clean up the accident site

Table 2. Cont.

Transition	Description
t_{14} Traffic Management Center	Maintenance of road facilities
t_{15}	Turn on the on-site fan
t_{16}	Open the escape route
t_{17}	Change variable signal notification

The evolution of the accident scene can be seen as a process in which emergency disposal tasks change constantly under the influence of emergency response behavior. In this paper, the rule of accident site evolution is represented by flow relationships and tokens in Petri nets. Taking the emergency task of on-site evacuation as an example, Figure 1 shows the situation in which on-site evacuation described by Petri nets changes under the influence of response behavior.



Place set:
 $P = \{p1, p2, p3, p29, p30, p31, p32\}$
Transition set:
 $T = \{t1, t1', t1'', t1'''\}$
Token set:
 Site evacuation= $\{a\}$, Escape routes= $\{l\}$, Information board= $\{m\}$
Token value:
 a, l, m

Figure 1. Interaction rules between emergency tasks and response behaviors in the case of evacuation.

As shown in Figure 1, p_1 , p_2 , and p_3 indicate that the on-site crowd is chaotic, the on-site crowd is to be evacuated, and the on-site crowd has been evacuated, respectively; p_{29} and p_{30} describe whether escape routes are opened, p_{31} and p_{32} show whether the information on the information board is released. The meanings of different places are listed in Table 1. The evolution law of on-site evacuation in Figure 1 is designed according to the following principles: when there is a crowd at the scene of a traffic accident that needs to be evacuated, and if the evacuation channel is opened and the variable information signs indicate the correct direction, the on-site crowd will be successfully evacuated under the

guidance of the response behavior of the traffic police department to evacuate the crowd. However, when the evacuation route is closed or the variable information signs give the wrong instructions, if the behavior of evacuating people is still carried out, the crowd will be confused because the escape route is closed or the guidance information is wrong, thus worsening the on-site disposal situation and hindering the emergency response process.

In order to better illustrate how the constructed model simulates the complete accident handling process, Figure 2 shows the video captured during the fire accident monitoring of the K1161+300m super-long tunnel No.1 in the Qinling Mountains of the Western Han Dynasty in Shaanxi on 1 November 2019. From Figure 3, we can see the emergency accident scene of vehicle collision and fire in the tunnel. Meanwhile, we can deduce that vehicle fire fighting, victim rescue, and traffic control have been carried out, and the traffic accident has been recorded, which provides a basis to establish the state space of the Markov game (based on Table 3). Figure 3 shows how to use the constructed Petri net model to simulate the fire drill.



Figure 2. Road surveillance video of traffic emergency fire drill in Qinling No.1 Tunnel. (a) Traffic control at accident scene, (b) On-site disposal of accidents.

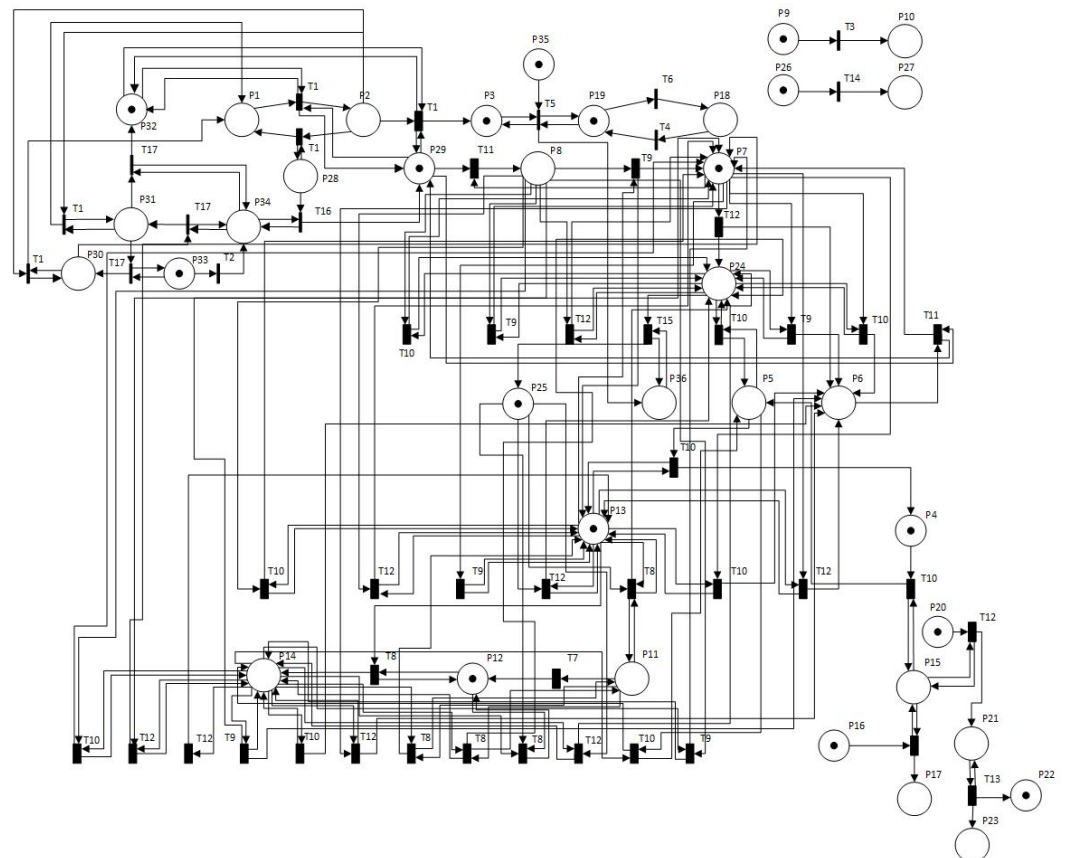


Figure 3. Petri net that simulates the emergency-decision-making process in the case of a traffic emergency fire drill in Qinling No.1 Tunnel.

Table 3. State space of Markov game.

Emergency Tasks	Value		
	−1	0	1
People			
Site evacuation s_{11}	p_1	p_2	p_3
Victim rescues s_{12}		p_4	p_5
Victim care s_{13}	p_6	p_7	p_8
Vehicle			
Explosion risk s_{21}		p_9	p_{10}
Fire cause s_{22}		p_{11}	p_{12}
Fire condition s_{23}	p_{13}	p_{14}	p_{15}
Vehicle damage s_{24}		p_{16}	p_{17}
Traffic			
Traffic control s_{31}		p_{18}	p_{19}
Site blocking s_{32}		p_{20}	p_{21}
Environment			
Road clearing s_{41}		p_{22}	p_{23}
Site smoke s_{42}	p_{24}	p_{25}	p_{26}
Road facilities			
Facility damage s_{51}		p_{27}	p_{28}
Escape routes s_{52}		p_{29}	p_{30}
Information board s_{53}		p_{31}	p_{32}
Accident information			
Accident report s_{61}		p_{33}	p_{34}
Accident record s_{62}		p_{35}	p_{36}

As shown in Figure 3, the emergency disposal tasks with 36 states are represented in Table 1, the 17 emergency response behaviors defined in Table 2 belonging to different departments are shown in Figure 3. As shown in Petri net, p_3 , p_7 , and p_9 in the figure correspond to three places, respectively, where the on-site crowd has been evacuated, the incident victim has been injured or has died, and where there is an explosion risk. The token in the place indicates that these states are activated at present, which also indicates that the emergency drill is in the early stage of the accident.

3. Collaborative Emergency Decision-Making Method

3.1. Emergency Decision-Making Framework Based on Multi-Agent Reinforcement Learning Algorithm

On the basis of using Petri nets to model the disposal process of highway accident scenes, this paper proposes a collaborative emergency-decision-making algorithm to formulate strategies for multiple emergency response departments at an accident scene. With the MADDPG algorithm, the multi-agent reinforcement learning agent represents the interaction between emergency response departments, including traffic police and road administration, and the on-site disposal model based on Petri net, and optimizes its decision-making ability by updating parameters. Figure 4 shows the framework of a collaborative emergency-decision-making method for traffic accidents based on the reinforcement learning algorithm:

The proposed method is composed of three parts: database, modeling, and collaborative decision-making. In this established framework, based on the analysis of the on-site emergency monitoring video and a large number of literal reports of highway emergencies from February to September 2019 obtained by the Shaanxi Highway Toll Center, and the discussion and exchange of disposal experience with emergency experts from the Qinling Highway Management Office, the disposal process of typical traffic accidents, including vehicle rear collision, rollover, and vehicle on fire, was summarized and analyzed. We extracted the interaction rules between the accident scene and the emergency crew from the road monitoring video of a real traffic emergency disposal event, and clarified the emergency tasks of each emergency department and the responsibilities of emergency

crew. On this basis, this paper uses the place, transition and flow relationships in Petri net to simulate the evolution of the accident scene, respectively. On this basis, this paper respectively uses the place, transition, and flow relationships in Petri net to simulate the evolution of the accident scene, the disposal work of the different emergency departments, and the information transmission of the different emergency departments. Then, in order to combine it with the reinforcement learning method, Petri net is transformed into a Markov game (MG) model. Since the state, action, and state transition laws of MG correspond to the place, transition, and flow relationships of Petri net, respectively, we can transform the Petri net of accident scene into a MG model to carry out reinforcement learning. Finally, with the help of the multi-agent deep reinforcement learning algorithm, the decision generation of collaborative emergency disposal for multiple emergency response departments is realized. In the multi-agent deep reinforcement learning algorithm, each emergency department corresponds to different agents, and each agent can obtain the MG model, which corresponds to the emergency site, and perform specific actions within each department to realize the optimal emergency-decision-making strategy through continuous learning. Instructions on how to convert the Petri net model into MG, and how to achieve collaborative emergency decision making with the help of the reinforcement learning algorithm, are described in detail in Sections 3.2 and 3.3, respectively.

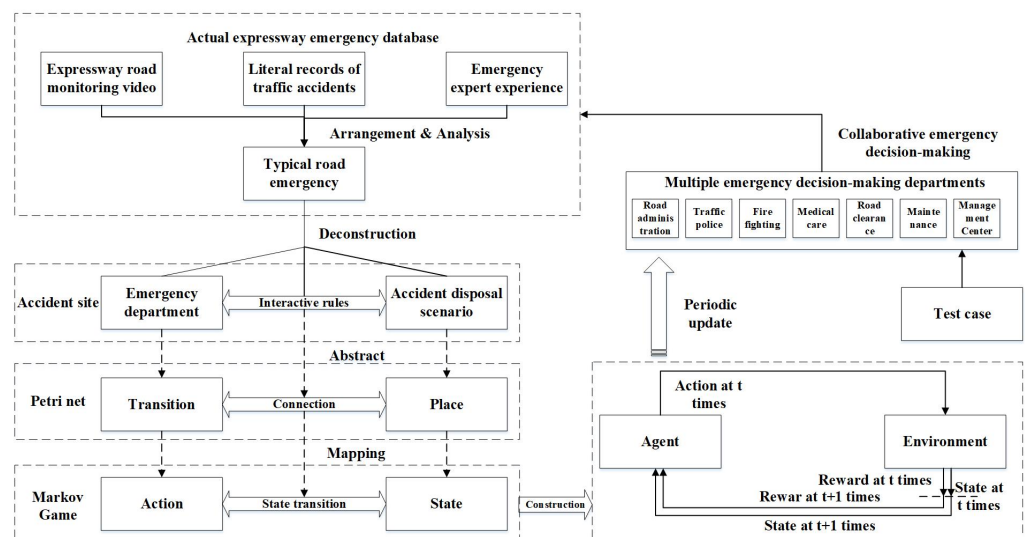


Figure 4. Framework of cooperative decision-making method for traffic incident emergency response.

3.2. Emergency Decision-Making Process Based on Markov Game

In order to provide a collaborative disposal decision for a highway traffic accident, it is necessary to convert the constructed Petri net model into a Markov game model, so as to facilitate further application in the reinforcement learning framework. Markov game is a dynamic and stochastic mathematical model that involves multiple participants, and is often used to describe the sequential decision-making processes of multiple agents. The Markov game model can be divided into tuples $N, S, \{a_i\}_{i \in N}, T,$ and $\{r_i\}_{i \in N}$, where $N = \{1, \dots, n\}$ means different decision makers, representing the emergency response department in the disposal process of traffic emergency decision making; S is the state involved in the game, representing the emergency disposal at the accident site, and the specific composition is shown in Table 1; $\{a_i\}_{i \in N}$ is the action space, which specifies the functional scope of emergency response crew in different emergency departments, $A = \prod a_i$; and $i \in N$ is the cooperative action space, including all agents. The functions of the departments are shown in Table 2. $T : S \times A \times S \rightarrow [0, 1]$ represents the distribution of state transition probability, which is the state transition matrix in the Markov game. It describes the evolution process of emergency scenarios under the influence of emergency actions, and is shown in Table 3. In addition, $r_i : S \times A \times S \rightarrow R, i \in N$ reflects the reward

obtained by each agent in the current state, where R is the reward space, corresponding to the evaluation of emergency response behavior. A typical Markov game is shown in Figure 5.

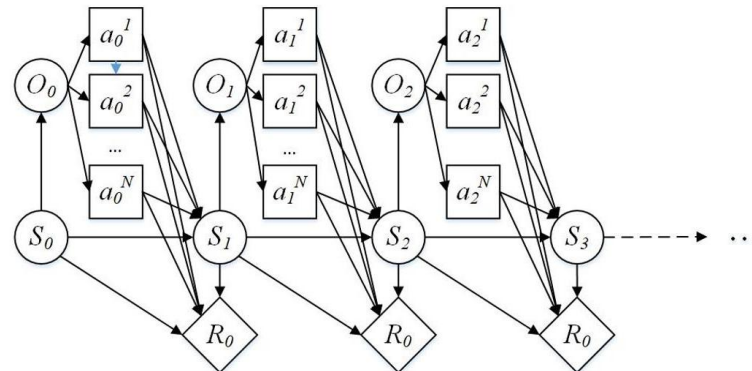


Figure 5. Diagram of Markov game.

As shown in the above figure, S_0 is the initial state of the emergency site, the initial scenario after the accident, and O_0 represents the state observed by the emergency department on the site; we set $O_i = S_i$ in this paper. Each agent according to the observation value O_0 and their current emergency response strategy $\pi_0(o_0)$, generates an action A_0 , and the corresponding reward value can be obtained. This cycle is repeated until the end, so that the evolution of the accident site is simulated iteratively. More specifically, Formula (1) mathematically describes the evolution process of emergency scenarios under the influence of emergency actions:

$$P_{ss'}^a = \mathbb{E}(S_{t+1} = s' | S_t = s, A_t = a) \quad (1)$$

in which S_t and S_{t+1} represent the disposal scenario of the accident scene at the time t and $t + 1$, A_t is the emergency response action at time t , and $P_{ss'}^a$ represents the probability that state s changes to state s' through action a .

In addition, as an important component of MG, the generation of reward value R for emergency decision making follows these principles: when any urgent task is not completed, -1 reward will be sent to each agent, and once all urgent tasks are completed, $\forall S_{ij} = 1$, that is, after the on-site emergency management is completed, each agent will be given a reward of $+100$. This setting can urge the emergency department to understand the urgency of the task, so as to better help the agent optimize its own decision-making ability [28].

When the Petri net model is converted into a Markov game model, Petri's place, transition, and flow relationships correspond to the state, action, and state transition relationships in the Markov game, respectively. In addition to simple element mapping, MG elements also need to be quantified to facilitate later numerical calculation. Table 3 shows the corresponding relationship between the state space in MG and the repository of Petri net. At the same time, the corresponding states and values are defined.

In order to show how to abstract the state, action, and evolution law from the actual accident scene using the MG model more intuitively, Figure 6 shows the process of abstracting the state of the MG model from the accident scene, and Figure 7 shows the method of using the MG model to simulate the evolution of the scene.

Figure 6 shows the state in the Markov game generated from road incident on-site scenarios. We mainly conduct an accident assessment on the incident site through the roadside high-definition camera to preliminarily judge the scene. If there are on-site persons to communicate with the emergency department, we can obtain more precise site state information. It can be seen from Figure 6 that the camera can cover the entire incident site; when the victims are rescued, Markov game state s_{12} is 1, and when the victims are not rescued, the state s_{12} sets to 0 [29]. As shown in Figure 7, if there is smoke at the accident site, the fire status of vehicles emergency task in the Markov model is assigned as 0. If there

is no fire at the site, the corresponding status is assigned as 1. In addition, it can be seen from the figure that the on-site vehicles have changed from the vehicle fire status to the flame extinguishment status after fire-fighting action. Therefore, according to the spatial distribution of the Markov model status and the corresponding relationship between the place and the emergency task, the state of on-site fire (s_{23}) in Markov state space changes from 0 to 1.

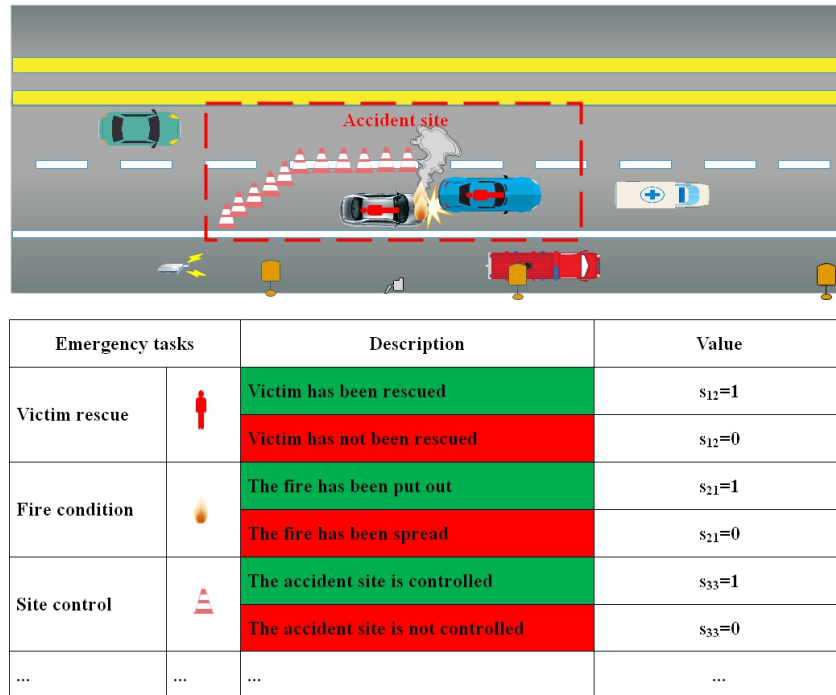


Figure 6. State in Markov game generated from road incident on-site scenarios.

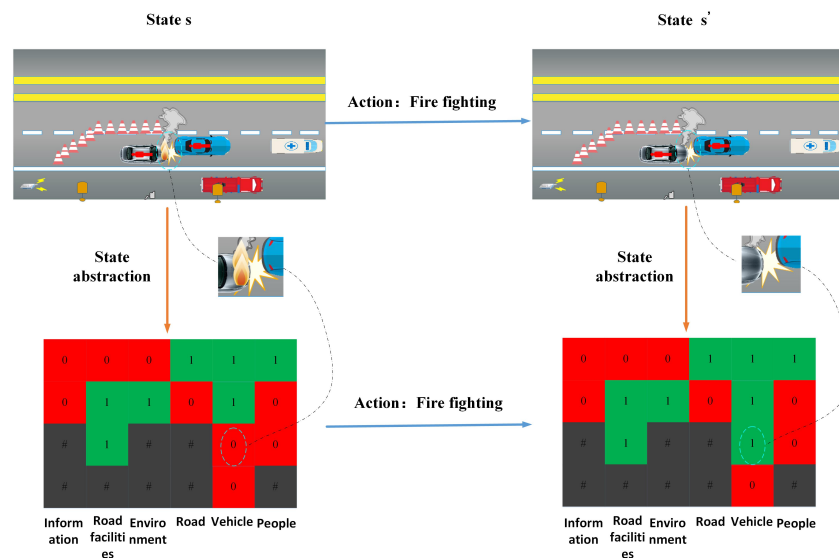


Figure 7. Example of emergency scenario evolution at the accident scene.

3.3. Emergency Decision-Making Method Based on Multi-Agent Reinforcement Learning

The multi-agent reinforcement learning method is an effective method to gradually train and generate agents with intelligent strategies in the process of interaction between agents and tasks. It is a common solution of the Markov game. The multi-agent depth deterministic strategy gradient algorithm (MADDPG) is a classical MADRL algorithm that has been widely used in various fields. MADDPG can use local information and then learn

how to obtain the optimal strategy, so as to achieve optimal action. It is also suitable for the cooperative environment and competitive environment of agents, and can effectively reduce the influence and interference between multiple agents. Figure 8 shows a schematic diagram of the MADDPG algorithm.

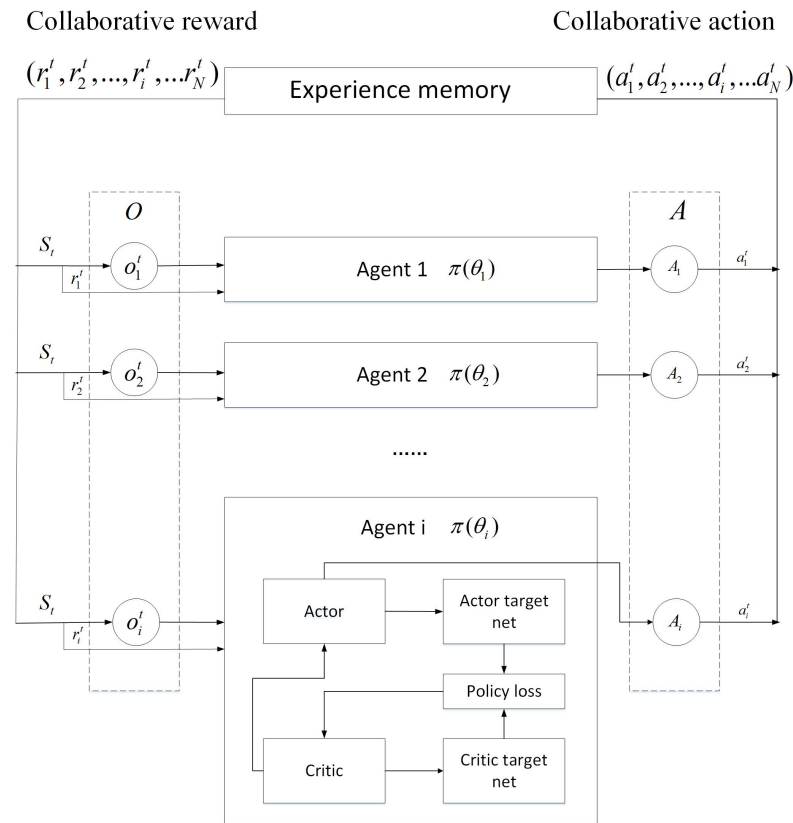


Figure 8. Diagram of the MADDPG algorithm.

MADDPG uses the centralized learning method to train the neural network. Specifically, optimizing the emergency strategy is to maximize the expected cumulative return $J(\theta_i) = \mathbb{E}_{S \sim p^{\mu}, a_i \sim \pi_i} [\sum_{t=0}^{\infty} \gamma^t r_{i,t}]$, where $\pi = [\pi_1, \dots, \pi_n]$ is used to represent the emergency response strategy of n emergency response departments and $\theta = [\theta_1, \dots, \theta_n]$ is the parameter of the neural network used by these generation strategies. Furthermore, the gradient function of strategy update can be obtained by derivation of emergency strategy, as shown in the following Formula (2):

$$\nabla_{\theta_i} J(\theta_i) = \mathbb{E}_{S \sim p^{\mu}, a_i \sim \pi_i} [\nabla_{\theta_i} \log \pi_i(a_i | o_i) Q_i^{\pi}(x, a_1, \dots, a_N)] \tag{2}$$

where o_i is the scene of the accident scene observed by the emergency department of i th and $x = \{o_1, \dots, o_n\}$ consists of the observation values of n emergency departments. $Q_i^{\pi}(x, a_1, \dots, a_N)$ is the action value function of the i th emergency department, which is used to evaluate different emergency response behaviors. Through this function, the emergency department agent can estimate the expected reward for different actions and select the optimal behavior accordingly. When policy μ_{θ_i} is determined, the gradient of expected return will be updated as Formula (3):

$$\nabla_{\theta_i} J(\mu_i) = \mathbb{E}_{x, a \sim D} \left[\nabla_{\theta_i} \mu_i(a_i | o_i) \nabla_{a_i} Q_i^{\mu}(x, a_1, \dots, a_N) \Big|_{a_i = \mu_i(o_i)} \right] \tag{3}$$

In order to learn from the experience of environmental interaction, MADDPG also sets up an experience memory \mathcal{D} to store the past experience of emergency disposal. The tuple $[s, s', \{a_i\}_{i \in N}, \{r_i\}_{i \in N}]$ constitutes the element in \mathcal{D} . In addition, the algorithm also sets a

loss function, which can break the correlation between sampled data through replaying experience and updating the target network, so as to better fit the real action value function $Q_i^\pi(x, a_1, \dots, a_N)$; the parameter update method of the evaluation network of the agent i^{th} is Formula (4):

$$\mathcal{L}(\theta_i) = \mathbb{E}_{x,a,r,x'} \left[\left(Q_i^\mu(x, a_1, \dots, a_N) - y \right)^2 \right] \tag{4}$$

where y is

$$y = r_i + \gamma Q_i^{\mu'}(x', a'_1, \dots, a'_N) \Big|_{a^j = \mu_j(o_i)} \tag{5}$$

$Q_i^{\mu'}$ represents the target network of the agent, $\mu' = [\mu'_1, \dots, \mu'_n]$ is the emergency strategy of the target network. In order to strengthen the communication and cooperation between emergency crew, the strategies of other agents are used as a part of the prediction target in the process of agent parameter update. Among them, $\hat{\mu}_{\varnothing_i}$ indicates the prediction $\hat{\mu}_j$ from agent i to agent j , the loss function of agent i is defined as Formula (6):

$$|\mathcal{L}(\varnothing_i^j) = -\mathbb{E}_{o_j, a_j} \left[\log \hat{\mu}_{\varnothing_i^j}(a_j | o_j) + \lambda H(\hat{\mu}_{\varnothing_i^j}) \right] \tag{6}$$

where \hat{y} is

$$\hat{y} = r_i + \gamma Q_i^\mu \left(x', \hat{\mu}_{\varnothing_i^1}(o_1), \dots, \hat{\mu}_{\varnothing_i^i}(o_i), \dots, \hat{\mu}_{\varnothing_i^N}(o_N) \right) \tag{7}$$

With the help of experience memory \mathcal{D} , the parameter update of any agent will be an emergency strategy optimization of all agents, hence the update gradient of each emergency disposal person's strategy is Formula (8)

$$\nabla_{\theta_i}^{(k)} J_e(\mu_i) = \frac{1}{K} \mathbb{E}_{x,a \sim \mathcal{D}_i^k} \left[\nabla_{\theta_i} \mu_i^{(k)}(a_i | o_i) \nabla_{a_i} Q_i^\mu(x, a_1, \dots, a_N) \Big|_{a_i = \mu_i^{(k)}(o_i)} \right] \tag{8}$$

In addition, in order to apply MADDPG to discrete action space, we use the reparameterization method, which is recommended in the literature [30], and introduce ϵ -Greedy, which balances the exploration and utilization in the process of strategy optimization. Figure 9 shows the real intention of using MADDPG to make collaborative emergency decisions. Algorithm 1 presents the pseudocode of the emergency disposal algorithm for a highway accident.

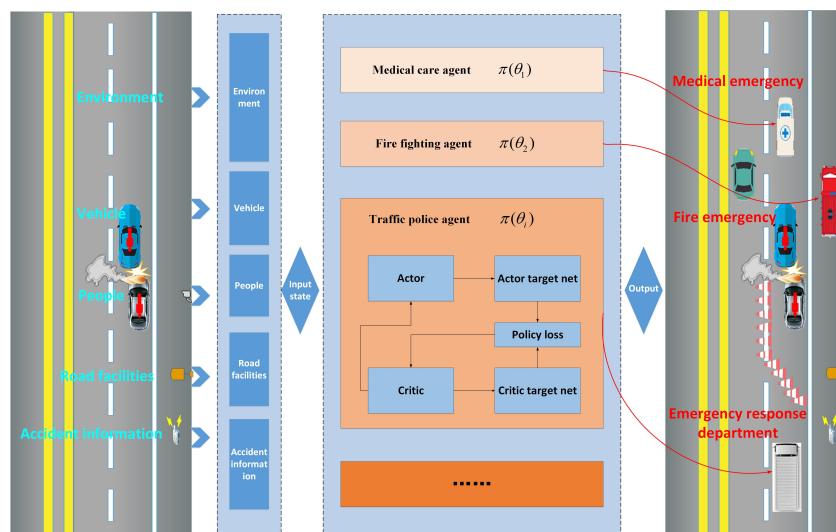


Figure 9. Diagram of cooperative emergency-decision-making method for freeway accident on-site disposal based on MADDPG.

Algorithm 1 Emergency decision-making method for on-site traffic incident disposal based on MADDPG

```

1: for episode= 1,  $E_{max}$  do
2:   Initialize a random noise for emergency action exploration  $N$ 
3:   Observe the initialization emergency of the accident site  $x$ 
4:   Initialize the total time step in the first episode  $t_s = 0$ 
5:   for time step= 1,  $M_{max}$  do
6:     Randomly select the emergency action of each agent  $i$  with a probability of  $\frac{1}{t_s}$ 
7:     Otherwise, for each agent  $i$ , select the emergency response behavior  $a$  according
to the current emergency strategy  $a_i = \operatorname{argmax}_a (\mu_{\theta_i}(o_i) + N_t)$ 
8:     Store  $(x, a, r, x')$  in experience memory  $\mathcal{D}$ 
9:      $x \leftarrow x'$ 
10:     $t_s \leftarrow t_s + 1$ 
11:    for Emergency department person  $i=1, N$  do
12:      Take  $S$  random batch samples  $(x^j, a^j, r^j, x'^j)$  from  $\mathcal{D}$ 
13:      Calculated  $y^j$  according to formula (7)
14:      Update the evaluation by formula (4) to minimize the loss
15:      Update the strategy gradient according to formula (8) according to the sam-
pling results
16:      Update the target network parameters of each agent  $i: \theta'_i \leftarrow \tau \theta_i + (1 - \tau) \theta'_i$ 
17:    end for
18:  end for
19: end for

```

4. Results

4.1. Experiment Initialization

In order to verify the effectiveness of the proposed collaborative emergency-decision-making method in the on-site emergency disposal of highway accidents, three real highway accidents were selected from the research database (represented by cases 1–3 in this paper). These three accidents covered different typical types of highway emergencies, including a fire accident that occurred in the K1161+300m extra-long tunnel No. 1 of Qinling, Western Han, Shaanxi, on 1 November 2019; a vehicle spontaneous combustion accident occurred at K4+000m of Shaanxi Wuding Highway on 7 September 2019; and a vehicle collision with a highway fence that occurred at K13+800m of Xi'an Ring Highway in Shaanxi on 26 February 2019. In order to verify the collaborative decision-making capability of the proposed algorithm, many algorithms, including the comparison baseline, were used for emergency decision making in these cases, and the disposal process steps were used as indicators to measure the length of disposal time and algorithm performance. Meanwhile, the actual emergency disposal process steps extracted from the cases were used to compare with the algorithm evolution results, and further illustrate the potential help of the proposed algorithm in practical applications.

As shown in Figure 9, when a traffic accident occurs, we can obtain the accident information through the information exchange between roadside facilities such as cameras and on-site personnel. The roadside camera can obtain the basic information of the accident scene, such as whether it is on fire, vehicle collision, and the on-site person can communicate with the emergency department by telephone to transmit more detailed on-site accident information. Then, the emergency agents assign corresponding emergency personnel and emergency vehicles to the site for emergency disposal. After determining the cases used in the experiment, it is necessary to further instantiate these cases, that is, convert them into the quantified MG model established in Section 3. Considering that the instantiation of traffic accidents is mainly realized by the adaptive numerical initialization of the MG model state matrix, after converting the state matrix of MG with six rows and four columns into a vector with one row and twenty-four columns, the model conversion method proposed in Section 3 is used. Taking the accident scenario in case 1 as an example, according to the corresponding relationship between the state space of Markov game and

different emergency tasks in Section 3, we can obtain the state attributes of six departments, namely, people, vehicle, traffic, environment, road facilities, and accident information. Among them, the vehicle department has the highest number of emergency tasks (four emergency tasks). Therefore, a 6×4 state matrix is set. However, other departments do not have four emergency tasks, so the spare matrix elements of each department are filled completely by setting 0. Therefore, there are 16 effective state factors and 8 complementary elements to form matrix complements. The formation process of the state matrix is presented in Figure 10. After distinguishing the initial emergency disposal tasks involved in the three cases, their initial statuses are $[0,0,1,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0]$, $[1,1,1,1,0,1,1,0,0,0,0,1,0,0,0,0,1,1,1,0,0,0,0,0]$ and $[1,1,1,1,0,1,0,0,0,0,0,1,0,0,0,0,1,1,1,0,0,0,0,0]$, respectively. In addition, this experiment was carried out on a computer equipped with Intel i5-8300H 2.30 GHz CPU, 8 GB memory, and NVIDIA GTX1060 GPU.

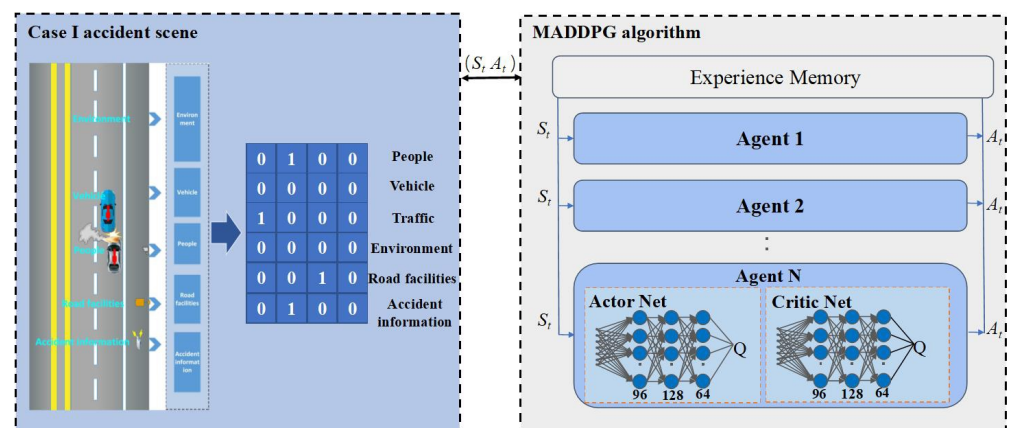


Figure 10. Emergency disposal process from scenario modeling to MADDPG algorithm in case 1.

4.2. Algorithm Comparison Baseline

In addition to verifying the impact of experience transfer on emergency decision making, the effect of the emergency-decision-making algorithm based on the MADDPG algorithm designed for multiple emergency departments on site should also be analyzed and compared. For this reason, this paper purposely selects a number of classic decision optimization algorithms as algorithm comparison baselines to carry out comparative analysis of case results, including the genetic algorithm (GA) [31], and a variety of decision optimization methods, including evolutionary strategy (ES) [32] and the deep Q network (DQN) [33], are used for emergency decision making of highway accident disposal. In addition to the three methods that can directly match the MG model, the genetic algorithm and evolutionary strategy are selected as non-machine learning parametric optimization methods, and the deep Q network is selected because of its deep reinforcement learning background. In addition, the deep Q network is selected as a single-agent algorithm. The results of the decision-making process can highlight the importance of multiple agents in accident scene disposal. The training super parameters of the above algorithms are shown in Table 4. Meanwhile, in the process of emergency rescue, it is necessary to coordinate different departments to form a general group decision. However, there are many uncertain factors in this process, which have brought different degrees of influence on group decision making. In order to improve the robustness of the system under uncertain environments, referring to the two concepts of individual operations and unit adjustment costs in group decision making (GDM), we have made some adjustments to the emergency response department model to improve its robustness in uncertain environments. In the following comparison with the experimental results of various algorithms, we can see that the robustness of the system has been improved [34–36].

The following describes the criteria for selecting hyperparameters of each emergency-decision-making process. In the genetic algorithm, if the crossover rate c_r is too high, the structure of individuals with high fitness in the population will be destroyed quickly,

leading to a strong oscillation in strategy optimization; otherwise, the genes of the population changes too slowly, and it is difficult to find an optimized strategy. As for the variability m_r in the genetic algorithm, when it is too large, the process of optimization may turn out to be a random search, when it is too small, the algorithm will fail to produce new genes. Thus, we set crossover rate c_r as 0.2 and variability m_r as 0.005. In the evolutionary strategy, if the variation intensity m_s of ES is too large, the optimization will fluctuate violently, and if it is too small, the proceeding of optimization would be too slow; therefore, we set the variation intensity m_s of ES as 0.2. The explore decay rate γ_ϵ in DQN is the decay factor of ϵ -Greedy strategy, which is an exploratory strategy designed to enhance the exploration of the environment by agents. At each time step, the agent with probability ϵ chooses action randomly, otherwise, it chooses action according to the current strategy. The explore decay rate is exactly the decay factor of ϵ , where $\epsilon_{t+1} = \epsilon_t + \gamma_\epsilon$. The value of ϵ could be relatively large at the beginning of the training process, and gradually become smaller as the training progresses. The explore decay rate γ_ϵ is generally set to range from 0.9 to 0.995; we chose to set it to 0.995 to ensure that the agent has sufficient exploration of the environment. As a member of DRL family, the MADDPG has a large number of sensitive hyperparameters, while most of them are similar to DQN hyperparameters. However, compared with DQN, noise intensity is a specific hyperparameter in MADDPG, and is specifically tuned in our proposed method. The noise intensity N refers to the noise that needs to be added to the action during exploration. The explore noise is added to the action when the agent interacts with the environment and collects training data; such settings could improve the agent's exploration of the environment, so we set the noise intensity of MADDPG as 0.2.

Table 4. Hyperparameters of each emergency-decision-making algorithm.

Algorithm	Symbol	Meaning	Value
GA	c_r	Crossover rate	0.2
	m_r	Variability	0.005
ES	m_s	Variation intensity	0.2
DQN	γ_ϵ	Explore decay rate	0.995
MADDPG	N	Noise intensity	0.2

Figure 10 shows the emergency disposal process from scenario modeling to MADDPG algorithm in case 1 emergency scenario. Initially, the information obtained by cameras and field personnel can be used to display the accident site states, and the corresponding 6×4 status matrix can be generated, in which 6 means 6 different emergency departments, with a total of 24 elements. Then, the generated state matrix is transferred to MADDPG network training; the MADDPG algorithm trains each agent, which represent different emergency departments through the actor and critic networks, and the agent selects the corresponding action accordingly. Finally, the corresponding emergency response process of case 1 can be obtained. More specifically, the internal parameters of each algorithm are set as follows: the structure of the neural network of MADDPG algorithm agents is identical; the main network and target network of each agent in the algorithm are the same; and the structure of actor network and critic network is also the same. They are all composed of fully connected neural networks with three hidden layers. The number of neurons in the three hidden layers is 96, 128, and 64, respectively. The number of neurons in the input layer and output layer of the neural network is determined by the state of the network perception and the shape of the output action. In addition, the population number of the genetic algorithm is 50, the population number of the evolutionary strategy is 50, and the sub algebra is 25. The neural network of the depth Q network also uses the full connection structure. The number of neurons in the four hidden layers is 10, 50, 30, and 15, respectively. Each algorithm will go through 150,000 steps of the decision-making optimization training process, and its final emergency-decision-making effect will be evaluated through the average reward in the training stage and emergency disposal process steps.

4.3. Collaborative Emergency Decision-Making Based on Multi-Agent Deep Reinforcement Learning

In order to elaborate the application of the proposed algorithm in the actual case and its improvement effect on on-site disposal, this paper first analyzes the emergency-decision-making effect of case 1 abstracted from the traffic fire emergency drill at K1161+300m of Qinling Tunnel 1 in the Western Han Dynasty in Shaanxi Province on 1 November 2019. A variety of emergency-decision-making algorithms, including GA, ES, DQN, and MADDPG, have been trained by 1000 episodes, respectively. The maximum step size of each episode is 1000 steps. The time step required for each episode algorithm to complete an emergency response is shown in Figure 11.

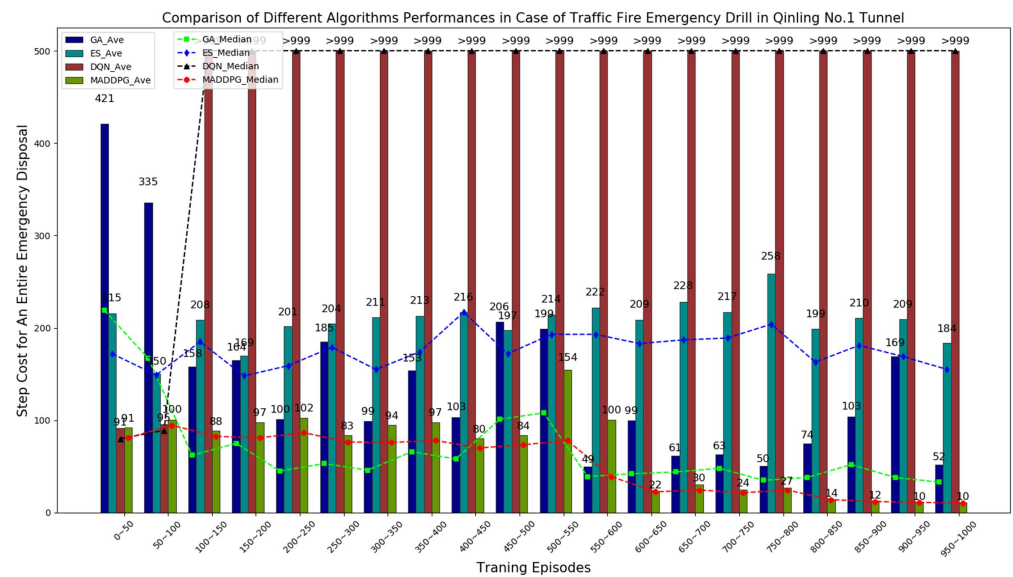


Figure 11. Performance of various algorithms in the case of a fire drill in Qinling No.1 Tunnel.

The horizontal axis in Figure 11 represents the number of training cycles, and the vertical axis is the emergency response time step required for the algorithm to complete all emergency response tasks in one episode. Considering that the disposal time is an important indicator of emergency response evaluation, the time step required for the algorithm to complete all emergency response tasks in a cycle was used to evaluate the quality of emergency decision making. In order to clearly display the results, we divided 1000 training episodes into 20 sections, each of which contained 50 episodes. The average and median values of emergency response time steps of different algorithms in each section are displayed in the form of a histogram and broken line graph, respectively. In addition, this paper selects the time step size of each algorithm in the last section for evaluation, because this represents their emergency-decision-making level after fully training the algorithm convergence. Therefore, a box plot of decision time steps of the last 50 cycles of each algorithm is shown in Figure 12, and the corresponding numerical statistics are shown in Table 5.

Through comparison, it can be found that the time steps of GA and ES decreased to varying degrees after training, and they are respectively stable at 180 steps and 50 steps in the last 50 episodes. The DQN algorithm diverges after training, and the final step cost is more than 2000. The emergency response time step of the MADDPG algorithm decreased significantly after training, and finally stabilized to about 10. It can be seen that MADDPG algorithm performs best among these algorithms. In the last 50 episodes, the time steps for MADDPG to complete an emergency response ranged from 9 to 15, with a standard deviation of 1.69. It shows that the emergency decision making of the trained MADDPG agent is not only fast but also stable.

In order to further evaluate the improvement of the MADDPG algorithm's collaborative emergency decision making on actual emergency disposal, this paper also extracts the actual human disposal process in the fire drill of Qinling Tunnel 1, and compares it with the emergency disposal time step as an indicator. After abstraction, it was found that in the actual emergency drill, the on-site response department spent 14 time steps to complete an emergency response. In order to more intuitively show the difference between MADDPG decision-making algorithm and manual disposal, the difference between the final 300-episode MADDPG disposal time step and manual disposal is shown in Figure 13.

Table 5. Algorithm statistics for the last 50 episodes.

Algorithm	Quantitative Characteristics of Emergency Response			
	Average	Standard Deviation	Maximum	Minimum
GA	52.04	58.45	326	16
ES	184.04	108.86	558	56
DQN	1132.52	682.36	3004	308
MADDPG	10.82	1.69	15	9

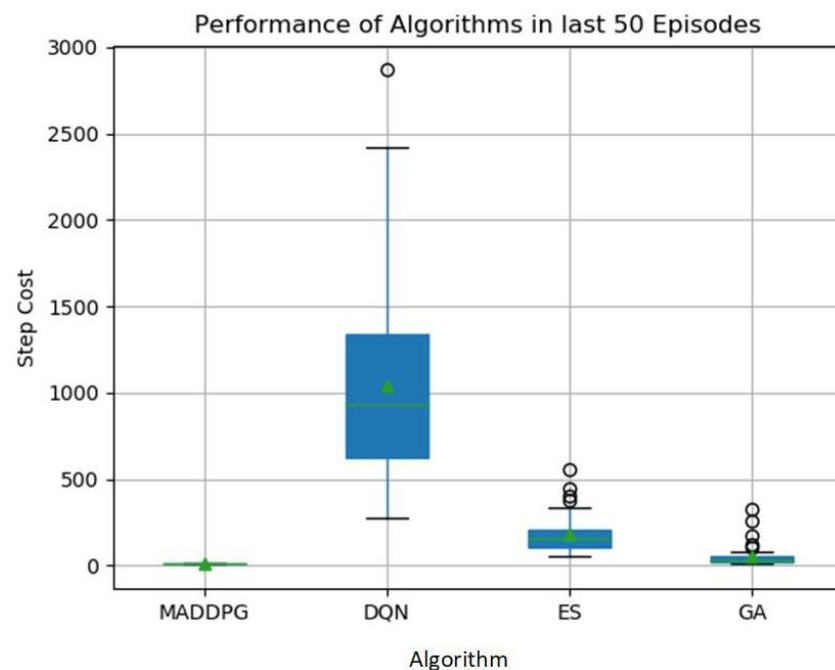


Figure 12. Box plot of step cost for different algorithms in last 50 episodes.

Figure 13a shows the change of the time step of MADDPG based emergency disposal decision in the form of a line chart. It can be seen that with the process of agent training, the decision making time step of emergency disposal is significantly reduced. It should be pointed out that the oscillation of the curve in the exploration period is due to the fact that there is no parameter update of the algorithm at this time, but the algorithm really starts to update the parameters from the 572th episode. The time step of the algorithm gradually decreases with the rapid rise, and finally stabilizes to a lower level. This process is the process that MADDPG continues to improve and optimize its emergency response strategy. Figure 13b shows the difference in time step between MADDPG and actual emergency disposal. The horizontal axis of Figure 13b is the number of training cycles, and the vertical axis is the step cost difference between the real emergency disposal decision and the agent. The green and red are used to distinguish the positive and negative step cost differences, respectively. If the difference is greater than zero, the curve is represented by green, indicating that the time required for the emergency decision generated by the agent to complete all tasks in the fire emergency drill is less than 14 steps spent in actual disposal. It can be seen

that with the training of agents, the difference gradually increases, and the decision-making of agents becomes better than the real emergency response process. The reason for this result is that the trained MADDPG based emergency response departments can effectively cooperate with each other, thus improving the efficiency of emergency response.

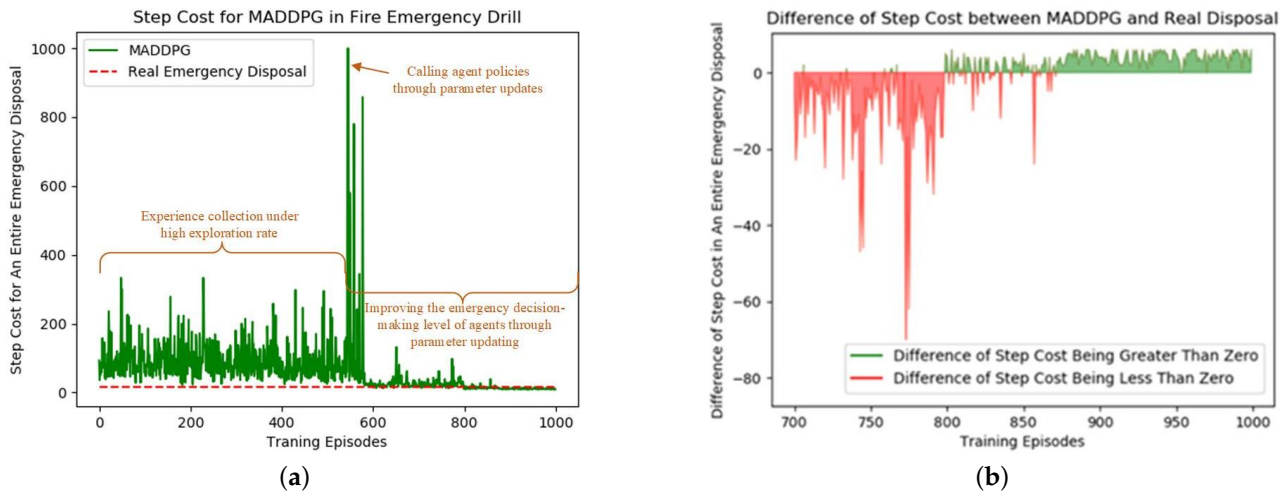


Figure 13. Comparison of step cost between MADDPG emergency decision-making and actual emergency response. (a) Step cost for MADDPG in Fire Emergency Drill, (b) Difference of step cost between MADDPG and real disposal.

Figure 13a shows the change in the time step of MADDPG-based emergency disposal decision in the form of a line chart. With the process of agent training, the decision making time step of emergency disposal is significantly reduced. It should be pointed out that the oscillation of the curve in the exploration period is because there is no parameter update of the algorithm at this time, and the algorithm starts to update the parameters from the 572th episode. The time step of the algorithm gradually decreases after the rapid rise after it starts to update, and finally stabilizes to a lower level, which clearly demonstrates the process of MADDPG continuous improvement and optimization of its emergency response strategy. Figure 13b shows the difference in time step between MADDPG and actual emergency disposal. The horizontal axis of Figure 13b is the number of training episodes, and the vertical axis is the difference in step cost in an entire emergency disposal. The green and red lines are used to distinguish the positive and negative step cost differences, respectively. If the difference is greater than zero, the curve is represented by green, indicating that the time required for the emergency decision generated by the agent to complete all emergency disposal tasks is less than 14 steps spent in actual disposal, and also indicating that the emergency decision generated by the agent is better than the actual disposal, which improves the efficiency of emergency disposal. We can see from Figure 13b, with the continuous training and learning of the agent, the difference between agent and actual disposal gradually increases, and the emergency disposal decision generated by the agent is better than the real emergency disposal process. This illustrates that the emergency response departments trained by MADDPG can effectively cooperate with each other to improve the efficiency of emergency disposal.

In order to prove that the proposed algorithm can provide collaborative decision making for the corresponding departments in the emergency response of different highway emergencies, two additional traffic accident cases, one single-vehicle collision accident that occurred on 2 April 2019 at K20+000m of Xi'an Ring Highway and one two-vehicle rear end collision accident that occurred on 5 April 2019 at K20+070m of Xi'an Ring Highway, were studied as case 2 and case 3. Emergency decision-making algorithms, including actual disposal (AD), GA, ES, DQN, and MADDPG were used in the on-site disposal decisions of case 2 and case 3, and the experimental results of each algorithm were consistent with

the performance of case 1. In order to intuitively demonstrate the optimization ability of different algorithms to respond to emergency decisions, the time steps of emergency disposal generated by different algorithms in the last training cycle of cases 1, 2, and 3 are shown in Table 6.

Table 6. Time step of different decision-making methods in last training episode.

Case	Quantitative Characteristics of Emergency Response				
	AD	GA	ES	DQN	MADDPG
I	14	16	56	308	9
II	7	19	45	158	5
III	7	10	33	163	4

As shown in Table 6, the performances of each algorithm in different cases are consistent: the MADDPG-based emergency-decision-making method is significantly better than GA, ES, and DQN. Meanwhile, in case 2 and case 3, the time step required for actual disposal extracted from the road monitoring video is 7, which is larger than the emergency decision generated by MADDPG, further indicating that the emergency disposal decision based on the MADDPG algorithm can optimize the emergency disposal process by coordinating different emergency departments so as to further improve the efficiency of emergency disposal and reduce the disposal time.

5. Conclusions

With the continuous increase in highway mileage and vehicles in China, highway accidents are also increasing year by year. However, the on-site disposal procedures of highway accidents are complex, which makes it difficult for the emergency department to fully observe the accident scene, resulting in the lack of sufficient communication and cooperation between multiple emergency departments, making the rescue efficiency low and wasting valuable rescue time, seriously endangering the lives of the people involved. This paper proposes a decision-making algorithm for on-site emergency collaborative disposal based on multi-agent deep reinforcement learning algorithm. First of all, with the help of the analysis of Shaanxi provincial highway monitoring videos obtained from the Shaanxi Provincial Highway Administration, this paper constructs an accident scene emergency disposal model based on Petri net to simulate the emergency disposal process at an accident site. After transforming the emergency disposal model into a Markov game model and applying it to the multi-agent deep deterministic strategy gradient (MADDPG) algorithm, which is proposed in this paper, the multi-agent can optimize the emergency-decision-making and on-site disposal procedures through interactive learning with the environment. Finally, in the actual emergency case verification, the difference between the disposal processes of the algorithm simulation and the actual situation are compared, and the proposed algorithm is also compared with typical algorithms such as the genetic algorithm (EA), evolutionary strategy (ES), and deep Q network (DQN). The results show that the emergency response of the proposed algorithm can better mobilize the team cooperation among emergency departments, and effectively reduce the time spent on the emergency disposal process. The algorithm proposed in this paper has been verified on three actual typical highway emergency disposal cases reported by highway road monitoring in Shaanxi Province. The experimental results show the advantages of our algorithm in the collaborative disposal of multiple emergency departments. Furthermore, by comparing the response time steps of MADDPG with actual video recordings of highway traffic emergencies, it can be seen that the response decisions generated by MADDPG can better promote the collaborative response among emergency departments. The response decision proposed by algorithm can reduce the time steps required for emergency response, and shorten the response time by 28%, 28%, and 42%, respectively, in the three actual emergency disposal cases, providing decision support and optimization to the actual response.

With this algorithm, the highway emergency disposal simulation can obtain the real-time emergency disposal situation of the accident scene through high-definition road cameras, make emergency decisions for different emergency disposal departments through the reinforcement learning agent, and finally conduct real-time collaborative disposal among multiple departments through communication terminals, including mobile phones and intercoms. In addition, with the promotion of intelligent highway management and the application of cloud computing technology in actual emergency management, the proposed algorithm will be able to learn the on-site disposal processes of different accidents, and provide decision support to the on-site emergency disposal crew. The multi-agent-based collaborative emergency-decision-making algorithm proposed in this paper can provide technical support for emergency response decisions for emergency departments in the future. In a serious traffic accident, which may cause a large number of casualties and property damage, the emergency decision generated by algorithm can strengthen the coordination of emergency departments, shorten the time of emergency disposal, and improve the efficiency of on-site emergency disposal through the emergency decision proposed by the algorithm according to the situation of the emergency site, which can increase valuable rescue time to ensure people's safety; meanwhile, the traditional traffic accident emergency drill takes a long time and requires a lot of manpower and material resources, which makes the economic cost and time cost of the drill higher, which has a significant impact on sustainability; thus, the number of emergency drills is lower. The introduction of this algorithm can provide a reference for emergency drills, greatly improve the efficiency of emergency drills, reduce the economic cost and time cost loss of emergency drills, and also reduce the waste of resources in the process of emergency response, so as to achieve sustainable development.

However, there are still limitations and space to be improved in this research. For the model building, although major traffic accident scenarios have been taken into account in this accident scene emergency disposal model based on Petri net, there are still some special circumstances that have not been considered, such as whether there are hazardous chemicals and explosion risks on the site, and the reward evaluation of the algorithm is relatively simple, mainly for traffic efficiency and traffic safety; for the the emergency-decision-making algorithm, in this paper, a multi-agent reinforcement learning algorithm is proposed, which mainly optimizes emergency decision making through the agents' interactive learning with the environment. However, these agents lack initiative and cannot directly accept the experience of emergency disposal experts, and they need to constantly learn to obtain optimized emergency disposal processes, which will waste some learning time. Therefore, the further aim of this research is mainly divided into two directions. Firstly, the on-site emergency disposal model can be improved, the emergency disposal process can also be more refined, and the evaluation factor of the emergency disposal site can be added. Secondly, we can select a multi-agent algorithm with better performance, or introduce supervision learning into the optimization decision of emergency disposal to further improve the learning speed of emergency disposal agents and improve the efficiency of emergency disposal. In short, with the continuous development of information technology and highway management system, the multi-agent-based collaborative emergency-decision-making algorithm will help the actual emergency response and emergency disposal in the future, effectively improve the rescue efficiency, ensure people's safety, and also reduce the waste of resources caused by insufficient coordination in the processes of emergency disposal and emergency drills, so as to achieve sustainable development.

Author Contributions: All authors were involved in conceptualization, visualization, investigation, methodology, writing—review and editing; formal analysis, J.Y., L.Y., Z.X. and P.W.; writing—original draft preparation, software, J.Y., L.Y. and Z.X.; validation, resource, supervision, project administration, funding acquisition, P.W. and X.Z. All authors have substantially contributed for the development of this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (No.2020YFB1600400).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Xian, H.; Wang, Y.; Hou, Y.; Dong, S.; Kou, J.; Zeng, H. Research on Influencing Factors of Urban Road Traffic Casualties through Support Vector Machine. *Sustainability* **2022**, *14*, 16203. [[CrossRef](#)]
2. Ma, X.; Ding, C.; Luan, S.; Wang, Y.; Wang, Y. Prioritizing influential factors for freeway incident clearance time prediction using the gradient boosting decision trees method. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 2303–2310. [[CrossRef](#)]
3. Yang, Y.; Jin, L. Visualizing Temporal and Spatial Distribution Characteristic of Traffic Accidents in China. *Sustainability* **2022**, *14*, 13706. [[CrossRef](#)]
4. Sun, X.; Hu, H.; Ma, S.; Lin, K.; Wang, J.; Lu, H. Study on the impact of road traffic accident duration based on statistical analysis and spatial distribution characteristics: An empirical analysis of Houston. *Sustainability* **2022**, *14*, 14982. [[CrossRef](#)]
5. Park, H.; Haghani, A.; Samuel, S.; Knodler, M.A. Real-time prediction and avoidance of secondary crashes under unexpected traffic congestion. *Accid. Anal. Prev.* **2018**, *112*, 39–49. [[CrossRef](#)] [[PubMed](#)]
6. Zeng, Q.; Liu, C.; Duan, H.; Zhou, M. Resource conflict checking and resolution controller design for cross-organization emergency response processes. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *50*, 3685–3700. [[CrossRef](#)]
7. Khan, R.U.; Yin, J.; Mustafa, F.S.; Liu, H. Risk assessment and decision support for sustainable traffic safety in Hong Kong waters. *IEEE Access* **2020**, *8*, 72893–72909. [[CrossRef](#)]
8. Ding, X.F.; Liu, H.C. A new approach for emergency decision-making based on zero-sum game with Pythagorean fuzzy uncertain linguistic variables. *Int. J. Intell. Syst.* **2019**, *34*, 1667–1684. [[CrossRef](#)]
9. Pérez-González, C.J.; Colebrook, M.; Roda-García, J.L.; Rosa-Remedios, C.B. Developing a data analytics platform to support decision making in emergency and security management. *Expert Syst. Appl.* **2019**, *120*, 167–184. [[CrossRef](#)]
10. Wang, L.; Zhang, Z.X.; Wang, Y.M. A prospect theory-based interval dynamic reference point method for emergency decision making. *Expert Syst. Appl.* **2015**, *42*, 9379–9388. [[CrossRef](#)]
11. García, L.A.; Tomás, V.R. A Framework for Enhancing the Operational Phase of Traffic Management Plans. *IEEE Access* **2020**, *8*, 204483–204493. [[CrossRef](#)]
12. Bacon, L.; MacKinnon, L.; Kananda, D. Supporting real-time decision-making under stress in an online training environment. *IEEE Rev. Iberoam. De Tecnol. Del Aprendiz.* **2017**, *12*, 52–61. [[CrossRef](#)]
13. Ma, Z.; Zhu, J.; Chen, Y. A probabilistic linguistic group decision-making method from a reliability perspective based on evidential reasoning. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *50*, 2421–2435. [[CrossRef](#)]
14. Song, S.; Yang, Z.; Yang, W.; Zhang, M. Research on multi-objective decision model based on location-allocation analysis during emergency traffic evacuation. In Proceedings of the 2010 IEEE International Conference on Emergency Management and Management Sciences, Beijing, China, 8–10 August 2010; pp. 106–109.
15. Qi, L.; Zhou, M.C.; Luan, W.J. Emergency traffic-light control system design for intersections subject to accidents. *IEEE Trans. Intell. Transp. Syst.* **2015**, *17*, 170–183. [[CrossRef](#)]
16. Huang, Y.S.; Weng, Y.S.; Zhou, M.C. Design of traffic safety control systems for emergency vehicle preemption using timed Petri nets. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 2113–2120. [[CrossRef](#)]
17. Sun, T.; Huang, Z.; Zhu, H.; Huang, Y.; Zheng, P. Congestion pattern prediction for a busy traffic zone based on the Hidden Markov Model. *IEEE Access* **2020**, *9*, 2390–2400. [[CrossRef](#)]
18. Ospina-Mateus, H.; Quintana, Jiménez, L.A.; Lopez-Valdes, F.J.; Berrio Garcia, S.; Barrero, L.H.; Sana, S.S. Extraction of decision rules using genetic algorithms and simulated annealing for prediction of severity of traffic accidents by motorcyclists. *J. Ambient Intell. Humaniz. Comput.* **2021**, *12*, 10051–10072. [[CrossRef](#)]
19. Zeleskidis, A.; Dokas, I.M.; Papadopoulos, B.K. Knowing the safety level of a system in real-time: An extended mathematical model of the STAMP-based RealTSL methodology. *Saf. Sci.* **2022**, *152*, 105739. [[CrossRef](#)]
20. Do, S.; Lee, C. Multi-agent Reinforcement Learning in a Large Scale Environment via Supervisory Network and Curriculum Learning. In Proceedings of the 2021 21st International Conference on Control, Automation and Systems (ICCAS), Jeju, Republic of Korea, 12–15 October 2021; pp. 207–210.
21. Serrano, W. Deep Reinforcement Learning with the Random Neural Network. *Eng. Appl. Artif. Intell.* **2022**, *110*, 104751. [[CrossRef](#)]
22. Wang, T.; Cao, J.; Hussain, A. Adaptive Traffic Signal Control for large-scale scenario with Cooperative Group-based Multi-agent reinforcement learning. *Transp. Res. Part C Emerg. Technol.* **2021**, *125*, 103046. [[CrossRef](#)]
23. Yang, J.; Zhang, J.; Wang, H. Urban traffic control in software defined internet of things via a multi-agent deep reinforcement learning approach. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 3742–3754. [[CrossRef](#)]

24. Guo, J.; Harmati, I. Evaluating semi-cooperative Nash/Stackelberg Q-learning for traffic routes plan in a single intersection. *Control Eng. Pract.* **2020**, *102*, 104525. [[CrossRef](#)]
25. Qin, Z.; Yao, H.; Mai, T. Traffic optimization in satellites communications: A multi-agent reinforcement learning approach. In Proceedings of the 2020 International Wireless Communications and Mobile Computing (IWCMC), Limassol, Cyprus, 15–19 June 2020; pp. 269–273.
26. Li, S. Multi-agent deep deterministic policy gradient for traffic signal control on urban road network. In Proceedings of the 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA), Dalian, China, 25–27 August 2020; pp. 896–900.
27. Li, Z.; Xu, C.; Zhang, G. A Deep Reinforcement Learning Approach for Traffic Signal Control Optimization. *arXiv* **2021**, arXiv:2107.06115.
28. Brockman, G.; Cheung, V.; Pettersson, L.; Schneider, J.; Schulman, J.; Tang, J.; Zaremba, W. Openai gym. *arXiv* **2016**, arXiv:1606.01540.
29. Balfaqih, M.; Alharbi, S.A.; Alzain, M.; Alqurashi, F.; Almilad, S. An Accident Detection and Classification System Using Internet of Things and Machine Learning towards Smart City. *Sustainability* **2021**, *14*, 210. [[CrossRef](#)]
30. Rezende, D.J.; Mohamed, S.; Wierstra, D. Stochastic backpropagation and approximate inference in deep generative models. *Int. Conf. Mach. Learn.* **2014**, *32*, 1278–1286.
31. Chen, L.; Wang, Y.; Guo, G. An improved genetic algorithm for emergency decision making under resource constraints based on prospect theory. *Algorithms* **2019**, *12*, 43. [[CrossRef](#)]
32. Roy, P.C.; Islam, M.M.; Murase, K.; Yao, X. Evolutionary path control strategy for solving many-objective optimization problem. *IEEE Trans. Cybern.* **2014**, *45*, 702–715. [[CrossRef](#)]
33. Wu, Y.C.; Dinh, T.Q.; Fu, Y.; Lin, C.; Quek, T.Q. A hybrid DQN and optimization approach for strategy and resource allocation in MEC networks. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 4282–4295. [[CrossRef](#)]
34. Ji, Y.; Li, H.; Zhang, H. Risk-averse two-stage stochastic minimum cost consensus models with asymmetric adjustment cost. *Group Decis. Negot.* **2022**, *31*, 261–291. [[CrossRef](#)]
35. Qu, S.; Wei, J.; Wang, Q.; Li, Y.; Jin, X.; Chaib, L. Robust minimum cost consensus models with various individual preference scenarios under unit adjustment cost uncertainty. *Inf. Fusion* **2023**, *89*, 510–526. [[CrossRef](#)]
36. Wei, J.; Qu, S.; Wang, Q.; Luan, D.; Zhao, X. The Novel Data-Driven Robust Maximum Expert Mixed Integer Consensus Models Under Multirole's Opinions Uncertainty by Considering Noncooperators. *IEEE Trans. Comput. Soc. Syst.* **2022**. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.