Research on Vehicle Networking Resource Management Based on Trust Model in Intersection Scene

Yanyi Li, Zhaonian Li, Wenhui Wang, and Zhenjiang Zhang(✉)

Beijing Jiaotong University, No. 3, Shangyuan Village, Haidian District, Beijing, China
zhangzhenjiang@bjtu.edu.cn

Abstract. Recently, the traditional cloud computing network of vehicle networking has some problems to be solved: 1) the security trust between the vehicle and the subgrade unit; 2) The vehicle may be attacked by potentially malicious edge servers during task unloading. In this paper, in order to solve the above problems, aiming at the security problem of the edge computing network in the vehicle task unloading and resource allocation problems of multi-vehicle and multi-subgrade units in the urban intersection scene, the vehicle task unloading and resource management optimization algorithm and trust model based on the vehicle edge computing network are constructed, and the approximate optimal simulation is carried out for the urban intersection scene. Simulation results show that the proposed algorithm can effectively improve the overall efficiency of the system.

Keywords: Internet of Vehicles · Edge Computing · Resource Allocation · Trust Model · Reinforcement Learning

1 Introduction

Nowadays, the introduction of edge computing unit in vehicle networking has some problems to be solved in practical application scene [1]. The first is the security and trust issues between the vehicle and the subgrade unit. When the malicious subgrade unit maliciously interferes with the vehicle, it will greatly affect the service quality of the vehicle edge computing network. On the other hand, the computing resources and energy resources in the subgrade processing unit are relatively limited and dynamic, so when the vehicle is attacked by malicious edge servers during the unloading process, the resource competition between vehicles will be more intense and unstable.

Based on the analysis of the trust problem and resource allocation problem between vehicles and subgrade units in road intersection scene, this paper proposes a distributed resource scheduling algorithm TBMADDPG based on trust model and deep reinforcement learning. Under the premise of ensuring the constraints of vehicle task unloading and resource allocation scheme, Minimize the time and energy cost consumed by the vehicle to process the task, and ensure that the vehicle in the system can complete the task calculation or unloading in a trusted environment.

D.-J. Deng and J.-C. Chen (Eds.): SGIoT 2023, LNICST 557, pp. 45–50, 2024.
https://doi.org/10.1007/978-3-031-55976-1_5

2 Related Works

In recent years, the global trust model analysis based on direct trust and indirect trust in the scene of vehicle networking is diverse. The distributed trust mechanism model proposed by Ke et al. comprehensively considers the information interaction between vehicles and subgrade units and the decay of trust over time and other factors [2]. The study [3] introduced timestamp mechanisms and blockchain concepts into trust management systems. Feng et al. used the satisfaction function of the service and the degradation factor function of time to improve the direct trust value described by the Bayes equation and improve the accuracy of the indirect trust model through an improved gray correlation method [4].

With the popularity of machine learning and its potential in the field of IoT applications, more and more researchers are applying this technology in the strategy optimization of task offloading and resource scheduling. For example, in the study [5, 6], LSTM algorithm is used to dynamically predict the edge communication and computing resources of mobile users, and make decisions based on the predicted future data. In the face of the challenges of vehicle computing intensive applications, multi-agent training reduces the instability in the environment and ensures that the strategy is updated [7].

3 System Model

Assumed that the intersection has two types of subgrade service units, including normal and malicious, and are connected to each other through optical fibers. To improve the model's ability to identify malicious subgrade units and balance the allocation of edge resources, this paper proposes a distributed resource scheduling algorithm TBMADDPG based on trusted model and deep reinforcement learning, so that the system can identify malicious subgrade units while reducing the overall time and energy consumption.

For all vehicles and subgrade units, the global trust matrix can be expressed as: $TrustMatrix(t) = [Trust_{i,j}(t)]$. In the description of the trust degree of the subgrade unit, the trust coefficient of the subgrade unit in a certain time slot is defined as the average of the global trust value of all vehicles, which can be expressed as: $Trust_j(t) = \sum_{i=1}^{n} Trust_{i,j}(t)/n$.

In the model, the subgrade unit to provide vehicles also dynamic change of computing resources, with the allocation rate denoted by T_R^t. The global trust matrix of the model is updated according to the interaction between the vehicle and the subgrade unit in each time slot. The vehicle's own state information is first defined as: $s_i(t) = \left[\tau_{max}, E_{r,j}, t_{w,j}\right]$. And The system state space of the current timeslot t can be defined as: $S_t = \left(s_1(t), s_2(t), \cdots, s_n(t), t_{w,m}, T_R^t, TrustMatrix\right)$.

Since all the vehicle's task is not completed by the vehicle itself is unloaded to the calculation model of a calculated in subgrade unit, so the vehicle i in current time slot t uninstall decisions can be expressed as a $(m+1) \times 1$ matrix. So, the system dynamic space can be finally expressed as a $(m+1) \times n$ matrix.

It can be seen that within a certain time slot in the entire model can have $(m+1)^n$ kind action choice. As the increase of model of unit in the vehicle, the number of action space will be to exponential growth make the space is too big, so as to cause the dimension explosion. Therefore, the multi-agent algorithm is adopted to solve this problem. In the actual multi-agent environment, it is assumed that agent A can obtain the behavioral strategies of other agents and train A's own network according to the corresponding experience generated. Each agent can learn by training several different strategies and choosing a general strategy.

Suppose there are K agents in the centralized training process, and the network parameters are $\theta = \{\theta_1, \theta_2, \cdots, \theta_K\}$. The deterministic strategy of all agents can be expressed as $\mu = \left\{\mu_{\theta_1}, \mu_{\theta_2}, \cdots, \mu_{\theta_K}\right\}$. The deterministic policy μ_k for agent k is:

$$\nabla_{\theta_k} J(\mu_k) = \mathbb{E}_{S,A\sim D}\left[\nabla_{\theta_k}\mu_k(a_k|s_k)\nabla_{a_k}Q_k^{\mu}(S, a_1, a_2, \cdots, a_K)|_{a_k=\mu_k(s_k)}\right]$$

For Critic networks, updates can be made according to the loss function:

$$\begin{aligned} Loss(\theta_k) &= \mathbb{E}_{S,A,S',R}\left[\left(Q_k^{\mu}(S, a_1, a_2, \cdots, a_K) - y\right)^2\right] \\ y &= r_k + \gamma Q_k^{\mu'}(S', a_1', a_2', \cdots, a_K')|_{a_j'=\mu_j'(s_j)} \end{aligned} \tag{1}$$

The Actor network can be updated by minimizing the policy gradient of the agent:

$$\nabla_{\theta_k} J \approx \frac{1}{Z}\sum_j \nabla_{\theta_k}\mu_k(s_k^j)\nabla_{a_k}Q_k^{\mu}(S^j, a_1^j, a_2^j, \cdots, a_K^j)|_{a_k=\mu_k(s_k^j)} \tag{2}$$

In order to ensure the stability of the training process, DDPG algorithm uses "soft update" to update some parameters of the target network. The update relation is (Fig. 1):

$$\theta^{Q'} \leftarrow \tau\theta^{Q} + (1-\tau)\theta^{Q'} \tag{3}$$

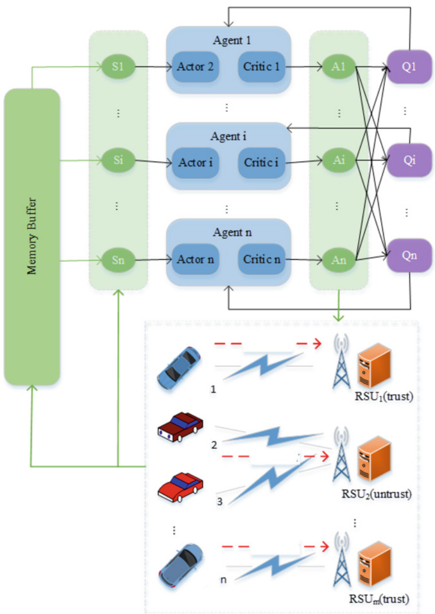

Fig. 1. The algorithm diagram of TBMADDPG

The specific training process of TBMADDPG algorithm is shown in Table 1.

Table 1. TBMADDPG training algorithm description for N agents

1	Randomly initialize all Actor and Critic networks and their respective weight parameters and their corresponding experience pools
2	Initializes information such as computing resources and network status in the system model
3	Initializing the direct and global trust matrix, and randomly specifying the malicious subgrade unit index
4	**For** episode = 1 to max_ episodes:
5	Initialize the action in the process of exploring the environment parameters: noise variable N_t subgrade unit, vehicle and information;
6	**For** time = 1 to max_ slots:
7	Randomly generate and sort vehicle task information in the current time slot.
8	Update the subgrade unit resource allocation in the model
9	**For** task = 1 to max_ indexes:
10	Each agent outputs actions based on the current policy network and noise perturbations and constraints $a_i = \mu_{\theta_i} + N_t$
11	Perform the action $a = (a_1, a_2, \cdots, a_n)$ to get rewarded r, and the next status s'
12	Waiting time delay queue updates, and the sample data (s, a, r, s') deposit pool R experience
13	**End For**
14	Calculate the global trust of all vehicles to the subgrade unit under the current time slot and update it
15	**For** agent = 1 to N:
16	Randomly sample Z bars of data (s^j, a^j, r^j, s'^j) in experience pool R to form mini batch
17	Update the online network parameters of Critic by formula (1).
18	Update the online network parameters of Actor by formula (2).
19	**End For**
20	Update the Target network parameters of each agent by formula (3)
21	**End For**
22	**End For**

4 Simulation and Result Analysis

In order to verify the validity of resource allocation strategy based on trust model in vehicle edge computing network, subgrade units are divided into normal subgrade units and malicious subgrade units. The malicious subunit not only provides false computing resources but also leaks task information, so the task should be avoided to be unloaded into the malicious subunit. In the simulation, the model randomly selects two subgrade units as malicious subgrade units. Regardless of the type of subgrade unit, all vehicles have an initial trust rating of 0.5 in them.

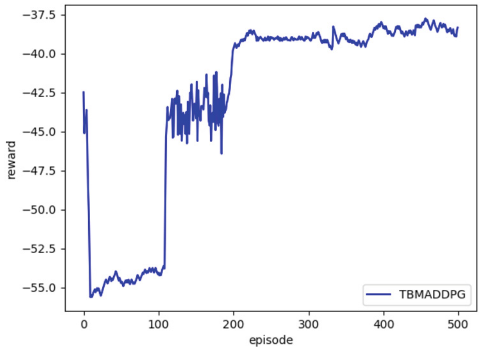

Fig. 2. Algorithm convergence curve

Figure 2 shows the training results of the TBMADDPG algorithm. As can be seen from the figure, after several fluctuations, the algorithm reaches a stable state in 400 rounds. The effect of the model increases rapidly in about 100 rounds because the multi-agent model first stores the experience group into its own experience pool during interactive learning, and then trains the network when the capacity of the experience pool is full. In fact, the model has achieved a good training effect in about 250 rounds, so the training speed of TBMADDPG algorithm is still relatively fast, and a reasonable unloading scheme with minimum delay and energy consumption based on the trust model can be obtained.

5 Future Work

In the simulation of this paper, the proposed algorithms are all based on DDPG algorithm. Considering the shortcomings of DDPG algorithm itself, other reinforcement learning algorithms, such as PPO and A3C, can be adopted in the subsequent research for simulation solutions, which may get better training results or achieve the purpose of simplifying the work.

The actions taken by the agent in this paper are actually from the direct interaction between the vehicle and the subgrade unit, and there is a "heavy unloading" situation in real life, that is, some actions need to be judged by a third party whether to continue to perform. In the face of this problem, the current reinforcement learning is difficult to learn the optimal strategy because of the selfishness of the third-party system nodes.

References

1. Chettri, L., Bera, R.: A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems. IEEE Internet Things J. **7**(1), 16–32 (2020)
2. Ke, X., Zhou, G., Du, Z.: Trust evaluation model for P2P networks based on time and interaction. MATEC Web Conf. **208**, 05005 (2018)
3. El-Sayed, H., Alexander, H., Kulkarni, P., et al.: A novel multifaceted trust management framework for vehicular networks. IEEE Trans. Intell. Transp. Syst. **23**(11), 20084–20097 (2022)

4. Feng, X., Yuan, Z.: A novel trust evaluation mechanism for edge device access of the Internet of things. Wirel. Commun. Mobile Comput. **2022**, 1–12 (2022). https://doi.org/10.1155/2022/3015206
5. Rago, A., Piro, G., Boggia, G., et al.: Antici-patory allocation of communication and computational resources at the edge using spatio-temporal dynamics of mobile users. IEEE Trans. Netw. Serv. Manage. **18**(4), 4548–4562 (2021)
6. Zheng, C., Liu, S., Huang, Y., et al.: Hybrid policy learning for energy-latency trade off in MEC-assisted VR video service. IEEE Trans. Veh. Technol. **70**(9), 9006–9021 (2021)
7. Zhu, X., Luo, Y., Liu, A., et al.: Multiagent deep reinforcement learning for vehicular computation offloading in IOT. IEEE Internet Things J. **8**(12), 9763–9773 (2020)