This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3233035

Resource Optimization of MAB-based Reputation Management for Data Trading in Vehicular Edge Computing

Huizi Xiao, Lin Cai, *Fellow, IEEE*, Jie Feng, Qingqi Pei, and Weisong Shi, *Fellow, IEEE*

***Abstract*—Vehicles are hesitant to upload data to edge servers in vehicle edge computing (VEC) as many vehicle data collected and perceived by various on-board sensors contain sensitive and personal information and lack economic incentive. Instead of free access to shared data, encrypted data trading will alleviate security and privacy concerns and provide an incentive for vehicle owners to share their data. The edge server needs to pay the price in data trading, and reputation management is a great method to help it trade with reliable and available vehicles. In this paper, we propose a multi-armed bandit (MAB)-based reputation management scheme, so the edge servers can select the high reputation vehicles for data trading, which can ensure the credibility and reliability of the data. The encryption scheme is applied to achieve the required transmission security level and defend the rights and interests of the edge server. On the other hand, implementing security measures will consume the computation and communication resources of the vehicles. We formulate an optimization problem that maximizes the revenue of vehicles in data trading under the constraints of time delay, energy consumption, and security level. Simulation results demonstrate that the proposed scheme is effective and efficient for vehicle reputation management, data trading selection, and resource allocation.**

***Index Terms*—Data trading, multi-armed bandit algorithm, reputation management, resource optimization, vehicular edge computing.**

I. Introduction

Along with the continuing development of on-board sensors, intelligent devices, and wireless communication technology, connected and autonomous vehicles (CAV) can acquire and produce vast amounts of useful data with great potential for new data-driven services [1]. The combination of massive data and machine learning makes the vehicle data valuable to many new services, such as assisted driving, simultaneous localization and mapping, objective tracking, and entertainment recommendation, etc. Edge servers may request data from vehicles in vehicular edge computing (VEC), and vehicles should upload the data to the edge server to support various intelligent applications. However, economic challenges and privacy concerns hinder the integration of data sharing in edge servers. Instead of free access to shared data, encrypted data trading will alleviate the security and privacy concerns of users and provide an incentive to vehicle owners to share their data.

Vehicles as data sellers are able to connect with the edge server, sharing and further increasing the data utility via pushing data as a kind of commodity in a digital market [2]. Ensuring the credibility and availability of trading data is a critical problem for edge servers. Trust and reputation systems can provide protection based on social control norms, which are considered as part of the next-generation security mechanisms named soft security [3]. This contrasts with traditional security mechanisms (hard security) such as authentication and access control. Moreover, it has been proposed for various applications, for example, to validate the trustworthiness of sellers and buyers in online transactions. Reputation management plays a vital role for data trading in vehicular networks and should be adaptive to the changing behaviours and the types of data [4]. Frequent changes in vehicle location bring difficulties to reputation management, and variations in the environment will also change the quality and correctness of vehicle data. It is necessary to maintain a globally updated reputation value to share the vehicle's reputation knowledge. In traditional trading systems, introducing reputation (or credit) can make economic transactions more efficient to grow the economy. Similarly, in data trading, we also apply the reputation concept. Each vehicle can be associated with a reputation score that is saved in the cloud server and can be accessed by all edge servers. After each data trading transaction, the edge server involved can report the satisfaction level to update the reputation score of the vehicle.

For data sellers, only if they get enough revenue from the data trading are they willing to consume computation and communication resources and compromise privacy to a certain degree to share data. In addition, encryption is needed to protect data security and privacy, so edge servers are willing

This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB1807500, in part by the National Natural Science Foundation of China under Grants 62102297, 62001357, 62202005, and 62132013, in part by the Key Research and Development Program of Shaanxi under Grant 2022GY-437, in part by the Guangdong Basic and Applied Basic Research Foundation under Grants 2020A1515110496 and 2020A1515110079, in part by the China Postdoctoral Science Foundation under Grant 2021M692501, in part by the Open Project of Shaanxi Key Laboratory of Information Communication Network and Security under Grant ICNS202005, and in part by the China Scholarship Council.
(*Corresponding authors: Jie Feng; Qingqi Pei*)

H. Xiao and Q. Pei are with State Key Laboratory of ISN, School of Telecommunication Engineering, Xidian University, Xi'an, Shaanxi 710071, China, and also with Engineering Research Center of Trusted Digital Economy, Universities of Shaanxi Province (email: huizi_xiao@stu.xidian.edu.cn, qqpei@mail.xidian.edu.cn).

L. Cai is with the Department of Electrical and Computer Engineering, University of Victoria, BC, Canada (email: cai@ece.uvic.ca).

J. Feng is with the State Key Laboratory of ISN, School of Telecommunication Engineering, Xidian University, Xian, Shaanxi 710071, China, and also with the Shaanxi Key Laboratory of Information Communication Network and Security, Xian University of Posts & Telecommunications, Xian, Shaanxi 710121, China (email: jiefengcl@163.com).

W. Shi is with the Department of Computer and Information Sciences, University of Delaware, Newark, DE 19716, USA (email: weisong@udel.edu).

to pay for the uploaded data. The data quality, including the security level, and the response time of data transactions, affect the vehicle's reputation value evaluated by the edge server. However, implementing security measures also consumes more computing and communication resources of the vehicles. There is a tradeoff between security and efficiency, and vehicles need to make resource allocation decisions to maximize revenue, save energy, and reduce latency [5], [6]. Due to the avalanche effect [7] of the encryption scheme and data availability, edge servers need accurate encrypted data from vehicles with high reputation value. Vehicles aim to obtain high revenue and a benign reputation with the lowest energy consumption.

To solve the above problems, in this paper, a multi-armed bandit (MAB)-based reputation management scheme is proposed to select the high reputation vehicle to trade data. Thus, vehicles have the incentive to maintain a high reputation. The original MAB problem is finding the most profitable slot machine among many machines [8], in which flexibility and adaptivity are suitable for dynamically changing vehicular networks. Encrypt trading data with a lightweight and fast block cipher can protect the security and privacy of the vehicle data and defend the rights and interests of the edge server. Each vehicle aims to jointly optimize transmit power, computation frequency, and block length to maximize revenue while meeting latency and security requirements. The main contributions of this paper are summarized as follows

- We propose a MAB-based reputation management scheme to select a high reputation vehicle for trading data, which can dynamically update the reputation value based on changing environment and vehicle status.
- We formulate an optimization problem that maximizes the revenue of vehicles in data trading under the constraints of time delay, energy consumption, and security level requirement. Also, the edge servers can securely acquire high-quality and low-error data traffic.
- We obtain the closed-form solutions for setting computation frequency and block length, and a minorization-maximization (MM) algorithm is developed to obtain optimal transmit power.
- Simulation results demonstrate that the proposed scheme is effective and efficient for vehicle reputation management, data trading selection, and resource allocation.

The remainder of this paper is organized as follows. The related works are described in Section II. Section III depicts the system model and problem formulation. Section IV presents the solution to the optimization problem. Section V presents the simulation results. The conclusion and future research issues are given in Section VI.

II. Related Work

The revenue model is a critical problem in data trading [2]. There are different measures to price the data by applying economic factors to auction mechanisms, game theory, the optimization problem, and blockchain technology. X. Cao et al. [9] formulated an iterative auction mechanism to coordinate the trading between multiple data owners, collectors, and users, which also avoids access to the agents' private information directly. In [10] and [11], the non-cooperative games are performed among multiple data consumers, and Nash equilibrium is analyzed and solved under different conditions to obtain the optimal transfer payment. The authors in [12] applied consortium blockchain technologies to ensure secure and trustful data trading and proposed a general blockchain-based data trading framework for the internet of vehicles. Also, an iterative double auction mechanism is used to achieve social welfare maximization to decide the amount of traded data and the price among buyers and sellers. A family of data pricing mechanisms for revenue maximization under different market settings is proposed in [13]. Their MGneral mechanism optimally solves the problem of revenue maximization by formulating it as a polynomial size convex program when multiple types of buyers coexist. For the problem of maximizing revenue in data trading, the above schemes rarely consider the encryption scheme for data security, and the iterative process, such as the auction and game theory method, will significantly increase the communication cost. Blockchain often leads to efficiency and throughput problems.

Vehicles within a road segment generate amounts of repetitive information in VEC. The buyers can select more trusted vehicles for data trading through reputation management. In [14], a centralized reputation management scheme is proposed to detect malicious nodes in the vehicular network, and they evaluate the performance under malicious attacks. However, there are no specific items and basis for reputation scoring in the paper. RLE [15] is a reputation-based leader election system for an opportunistic autonomous vehicle platoon, where reputation value is recorded on the blockchain. Z. Tian et al. [16] propose a reputation framework for identifying denial of traffic service to resolve the trustworthiness problem in the application level of the internet of connected vehicles. A consortium blockchain-based resource sharing paradigm in the internet of vehicles is provided in paper [17], in which the interactions are encapsulated as transactions and recorded in RSUs. A lightweight consensus mechanism is also proposed to reduce computational power consumption and motivate vehicles to participate in resource sharing. The above reputation management solutions are still too heavy for dynamically changing vehicles on the road.

It is necessary to meet latency requirements and energy consumption constraints through resource optimization in VEC. An optimization problem is formulated in [18] to maximize the long-term utility of the vehicle edge computing network. They use the Markov process, Q-learning, and deep reinforcement learning to obtain the optimal resource allocation policies. In [19], the authors formulate the problem of joint node selection and resource allocation to minimize the total computation overhead in terms of the weighted sum of task completion time and monetary cost for computation resources. Based on Stackelberg dynamic game, H. Xu et al. [20] propose a resource pricing and trading scheme to allocate edge computing resources, and blockchain technology is applied to record the entire resource trading process to protect security and privacy. Optimizing resources and improving security levels are contradictory. Using deep learning to optimize resources

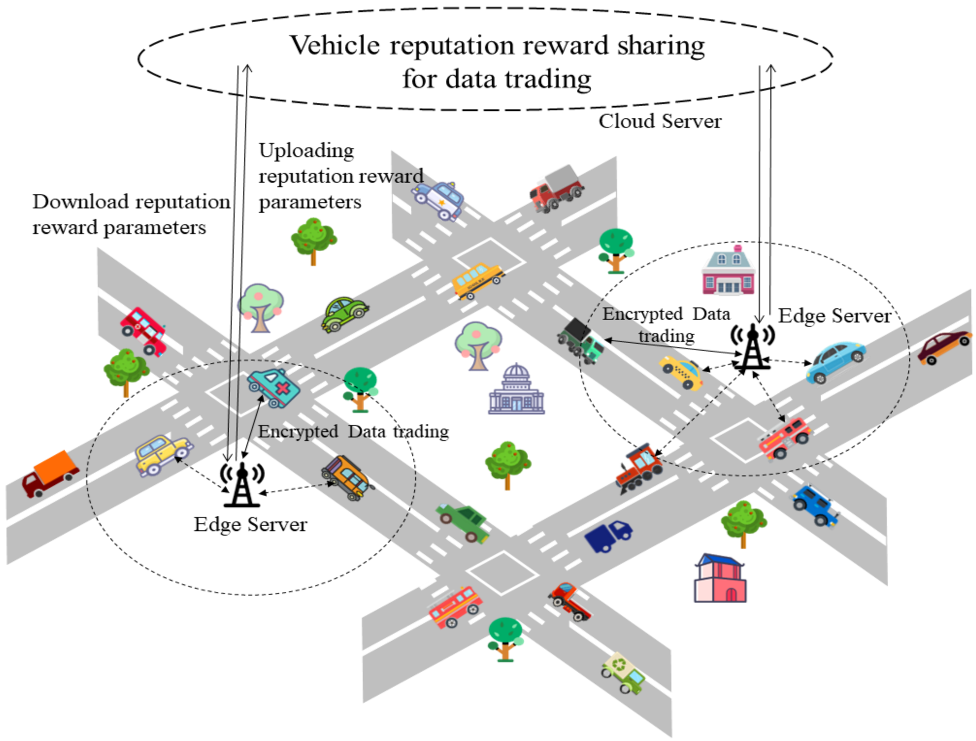


Fig. 1: The system scenario.

or blockchain to protect security and privacy will sacrifice enormous computing and communication resources.

Many previous data-trading proposals based on auction or blockchain technologies are quite heavy and cannot quickly adapt to the highly dynamic vehicle network environment. How to leverage the reputation and optimize encryption and resource allocation to support fast and low-overhead data trading for vehicular edge computing applications remains an open issue, which motivated this work.

III. System Description And Problem Formulation

A. System Scenario

Future smart vehicles on the road can sense and acquire valuable data. Selling the data to edge servers will maximize the value of the data and benefit intelligent transportation and smart city applications. Appropriate compensation for the data can provide the incentive to vehicles to sacrifice some energy and a certain degree of privacy to share their data. To protect the security and privacy of the vehicle data and the rights and interests of edge servers, who purchase the data should be encrypted. Furthermore, the edge server hopes to trade data with vehicles with high reputation values to ensure the credibility and reliability of the data.

In the system scenario in Fig. 1, data trading is between the vehicles and edge servers (e.g., located at the roadside unit, RSU). Meanwhile, all edge servers can communicate with the cloud server, which maintains the reputation values for all vehicles. The flexibility and adaptivity of the MAB algorithm are suitable for updating reputation values for dynamic changing vehicular networks. Denote the vehicle group as $\mathcal{I} = \{1, 2, 3, \cdots, I\}$. Once the vehicle is within the coverage of an edge server, the server can obtain the vehicle's reputation value and MAB learning parameters from the cloud. The edge server can choose a vehicle $i \in \mathcal{I}$ to obtain data according to its reputation value. When the server uses the reputation value to process encrypted data trading with the chosen one, the MAB algorithm can be utilized to update the reputation of the vehicles online. By setting the parameters, the MAB algorithm can tradeoff exploration and exploitation, so vehicles other than the one with the highest reputation can be selected when desirable considering other factors such as communication cost. When the vehicle leaves this coverage area, the edge server can update its reputation value and MAB learning parameters to the cloud to be shared by other edge servers. Once the edge server selects a vehicle for data trading, the vehicle can allocate computation and communication resources, and decide the security settings to maintain its reputation and maximize revenues.

B. Reputation and Security Level Measurement

The reputation management in the edge server for the vehicles in the data trading requires a trust and reputation calculation model. The vehicle needs to encrypt trading data to alleviate the vehicle's privacy concerns and ensure the rights and interests of edge servers.

1) Trust and Reputation Modeling: Trust and reputation systems generally prepare scores to evaluate the behaviors of the vehicles in VEC. These scores help the edge servers to improve the quality of their decision by deploying incentive mechanisms and ensuring the trading data has a higher credibility level.

The trust and reputation evaluation of the edge server to the vehicle i is denoted as $\varrho_i := \{b_i, d_i, u_i, \vartheta_i\}$, where $b_i, d_i, u_i, \vartheta_i \in [0,1]$ and $b_i + d_i + u_i = 1$. Here, b_i and d_i refer to the belief and unbelief of the vehicle i for the edge server, respectively. Belief is the probability that the data information is true, whereas disbelief is the probability that the data information is false. u_i is the uncertainty of vehicle i, which is the individual confidence in the knowledge of an edge server on vehicles. ϑ_i is the base rate [21], which is a predefined constant formed from an existing impression without solid evidence. The communication quality of a link between the vehicle i and the edge server is q_i, i.e., the successful transmission probability of the data packet, which determines the uncertainty u_i. So b_i, d_i, and u_i can be calculated from the statistics of past communication experience between the edge servers and the vehicle i as follows

$$\begin{aligned} b_i &= (1-u_i)\frac{\gamma_i}{\gamma_i+\rho_i}, \\ d_i &= (1-u_i)\frac{\rho_i}{\gamma_i+\rho_i}, \\ u_i &= 1-q_i, \end{aligned} \tag{1}$$

where γ_i indicates the number of positive interaction records, and ρ_i is the number of negative interaction records [22]. Eventually, we obtain trust and reputation score r_i of vehicle i as follows

$$r_i = b_i + u_i\vartheta_i. \tag{2}$$

2) Transmission Security Level: It is insufficient to consider the correctness and availability of information and the quality of communication links in data trading only. Most of the data collected by the vehicles involve private information of users, such as images, locations, trajectories, and so on. To maximize the utility of data for more revenues, users may be willing to sacrifice a certain degree of privacy to share data with edge servers. However, if the trading data is transmitted directly, the privacy will be leaked to eavesdroppers, which also violates the rights and interests of the edge server who paid to purchase the data. Therefore, encryption schemes should be implemented to ensure confidentiality and realize a certain transmission security level.

Symmetric encryption means encrypting and decrypting data using the same key. Its convenience, speed, and flexibility are ideal for dynamically moving VEC scenarios. Under determining the secret key, the block cipher algorithm uses the symmetric key and algorithm to encrypt and decrypt the N-bit blocks of the original data [23]. Considering the brute force attack to break the cipher, the transmission security level is proportional to the block length of the encrypted message. Assuming the block length of vehicle i is denoted as N_i, there are 2^{N_i} combinations. The transmission security level is defined as

$$\Xi_i = \log_2(N_i), \tag{3}$$

which should be considered in calculating reputation value for a data trading transaction, and it will also affect the reputation of the vehicles.

C. MAB-based Online Reputation Management Algorithm

The flexibility and adaptivity of the MAB algorithm are suitable for updating reputation values for dynamically changing vehicular networks. MAB is a framework to model a sequential decision-making problem, which has an agent repeatedly choosing one of the multiple actions, also called arms [24]. The agent draws an arm and observes a reward from the environment. Before actually selecting that arm, an instantaneous reward is unknown to the agent. The objective of each action is to maximize the total reward of the iterative process. The optimality of action and the selection of algorithms mainly depend on the environment. There are generally three types of environments [25]: stationary stochastic environment, switching stochastic environment, and adversarial environment. The reward distributions cause the difference in the environment, and there are various mature MAB algorithms suitable for the above environments, such as upper confidence bound (UCB) [26], Thompson Sampling (TS) [27], and exponential-weight algorithm for exploration and exploitation (EXP3) [28]. The condition of the vehicles not only shows regularity due to hardware, weather, configuration and so on, but also will affect its data trading settings change over time because of their mobility on the road. Therefore, the reputation value of the vehicles needs a dynamic observation process, and a certain observation opportunity is given to the vehicles whose state changes. MAB online learning tools can be well adapted to these problems.

Low latency is critical in the VEC. Assuming the latency requirement of the edge server for data trading is $T^{\max}$, and the time delay of vehicle i is represented as T_i. When the vehicle takes more than $T^{\max}$ to process and transmit trading data, we reduce its reputation value to zero for this transaction. Therefore, the delay indicator function is given by

$$\Phi(T_i, T^{\max}) = \begin{cases} 1, & \text{if } T_i \le T^{\max} \\ 0, & \text{if } T_i > T^{\max} \end{cases}, \forall i \in \mathcal{I}. \tag{4}$$

We apply the following method to calculate the reputation value for the vehicle i in transaction m,

$$\Upsilon_i^m = \frac{\Xi_i}{\Xi_{\max}} \cdot r_i \cdot \Phi(T_i, T^{\max}), \tag{5}$$

where $\Xi_{\max}$ is the maximum transmission security level according to the maximum encryption block length.

Taking Υ_i^m as the reward of the vehicle i in the mth iteration of the MAB algorithm, the goal is to select the vehicle with a large reward value (reputation value) each time the edge server conducts data tradings to ensure the credibility and availability of the acquired data. We achieve the target by choosing the largest value of A_i^m to realize action A^m in the vehicle group for each transaction. The specific representations are as follows

$$A^m = \text{argmax}\left[A_i^m\right], \tag{6}$$

and

$$A_i^m = \overline{\Upsilon_i^{m-1}} + \sqrt{\frac{2\ln m}{Q_i^{m-1}}}, \tag{7}$$

where $\overline{\Upsilon_i^{m-1}}$ is the average reputation value of vehicle i in the

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3233035

previous $m-1$ transactions, and Q_i^{m-1} is the number of times that the edge servers select vehicle i till the transaction $m-1$. The transaction count of each vehicle is dynamically updated by the edge servers and uploaded to the cloud for experience sharing. Therefore, action A^m tends to give a new vehicle in the reputation management system more opportunity to record its truest reputation status. In order to save storage resources by not storing the reputation value Υ_i^m of each transaction m for each vehicle i, we use the following exponential weighted moving average method to obtain the average reputation value $\overline{\Upsilon_i^m}$ of each vehicle i in transaction m,

$$\overline{\Upsilon_i^m} = (1 - \frac{1}{Q_i^m})\overline{\Upsilon_i^{m-1}} + \frac{1}{Q_i^m}\Upsilon_i^m, \tag{8}$$

where Q_i^m is the number of times that the edge servers select vehicle i till the transaction m, so we only store the average reputation value and compute the new reputation value in the current transaction to update the new average one.

Algorithm 1 ε-greedy Multi-Armed Bandit (MAB)-based Online Reputation Management Algorithm

Initialization:

- Initialize the communication quality $\mathbf{q} = \{q_i\}$ and the probability of positive interactions $\mathbf{Prob} = \{P(pos_i)\}$ from the distribution.
- Set $m = 1$ and the specified greedy parameter ε.
- Let $\boldsymbol{\gamma} = \{\gamma_i\}$, $\boldsymbol{\rho} = \{\rho_i\}$, $\boldsymbol{Q}^m = \{Q_i^m\}$ and $\overline{\boldsymbol{\Upsilon}^m} = \{\overline{\Upsilon_i^m}\}$ to zero.
- Iteratively select each vehicle and optimize resources for the data trading process. Increase the m, $\boldsymbol{Q}^m$ and record the $\overline{\boldsymbol{\Upsilon}^m}$.

Iteration:

```
while m > I do
    if A rand value ∈ [0,1] < ε then
        Choose random vehicle i to trade data.
    else
        Compute A_i^m = \overline{Υ_i^{m-1}} + sqrt(2 ln m / Q_i^{m-1}).
        Choose vehicle i with A^m = argmax [A_i^m].
    end if
    Update Q_i^m = Q_i^m + 1.
    Vehicle i jointly optimize resource to get optimal P_i, f_i and N_i to maximize utility.
    Update γ_i = γ_i + 1 or ρ_i = ρ_i + 1.
    Compute Φ(T_i, T^max), r_i and security level to get reputation value Υ_i^m = (Ξ_i / Ξ_max) · r_i · Φ(T_i, T^max).
    Record \overline{Υ_i^m} = (1 - 1/Q_i^m) Υ_i^{m-1} + (1/Q_i^m) Υ_i^m.
    Update m = m + 1.
end while
```

Output: Share reputation value $\overline{\Upsilon_i^m}$ and parameters ε, m, Q_i^m, γ_i, ρ_i to the cloud.

We employ the ε-greedy algorithm to give the state-changing vehicles ample opportunity to trade data. The specific procedure of the ε-greedy MAB-based online reputation management algorithm is described in Algorithm 1. Properly increasing ε can enhance the probability of the edge server randomly selecting vehicles with fewer transaction experiences, and reducing ε can stick out the high reputation vehicles.

D. Resource Optimization Problem of Vehicle

While the edge server performs the MAB-based online reputation management algorithm, it is accompanied by the actual encrypted transmission and data trading from the vehicle to the edge server, which consumes computation and communication resources. So, after the edge server chooses a vehicle for data trading according to the reputation value, the vehicle needs to optimize and allocate its computation, transmission, and security settings to maximize its revenue under the constraints of data trading delay and energy consumption.

1) Encryption Computation: Before starting to transmit data, the vehicle should encrypt the trading data at first. Denote the data size to be transferred as S_i, and let the processing density of the on-board CPU be l_i (cycles/bit). We can obtain the needed CPU cycles, $S_i l_i$. Assuming the encryption computation frequency is f_i, the time delay of the encryption computation is given by $T_i^{en} = \frac{S_i l_i}{f_i}$. We model the computation power of the on-board CPU as $k_i f_i^3$, where k_i represents the effective switched capacitance relying on the chip architecture [29]. Therefore, the energy consumption for encryption computation is $E_i^{en} = k_i S_i l_i f_i^2$.

2) Data Transmission: Let B be the bandwidth, and N_0 is the power spectral density of the additive white Gaussian noise (AWGN). The capacity of an AWGN channel can be expressed as follows

$$R_i = B\log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER})BN_0}\right), \tag{9}$$

where P_i and h_i are the transmit power and channel fading coefficient between vehicle i and the edge server, respectively. d_i is the distance from i to the edge server and β is the path loss exponent. Last, $\Gamma(\mathbb{BER}) = -\frac{2\ln(5\mathbb{BER})}{3}$ denotes the signal to noise ratio (SNR) gap, which presents how far the system is from the achieving capacity to realize the desired target bit error rate (BER) [30]. The transmission time delay can be obtained as $T_i^c = \frac{S_i}{R_i}$. The energy consumption to transmit encrypted data is $E_i^c = P_i T_i^c$.

3) Revenue: Before data trading, the edge server aims to select a vehicle with a higher reputation value according to the MAB algorithm. At the same time, a higher price should be paid for a vehicle with a high reputation value to incentivize the vehicle to increase its reputation in each transaction. After implementing the selection decision, the edge server expects low BER data traffic acquisition due to the avalanche effect of the encryption scheme. Otherwise, the ciphertext may not be decrypted correctly. The utility of the vehicles in data trading can be represented as follows

$$\mathcal{U}_i = W_R \overline{\Upsilon_i^{m-1}} \frac{R_i}{N_i}(1 - \mathbb{BER})^{N_i}, \tag{10}$$

where W_R is the utility price, $\frac{R_i}{N_i}$ is the number of transmitted encryption blocks per second, and $(1 - \mathbb{BER})^{N_i}$ is the probability that N_i bits per block are transmitted without errors.

On the vehicle side, to maximize revenues and reduce its own energy consumption costs for both encryption and communication, the revenue is given by

$$\mathcal{G}_i = \mathcal{U}_i - W_E(E_i^{en} + E_i^c), \tag{11}$$

where W_E is the energy price.

4) Problem Formulation: For vehicle i to trade data, it can optimize transmit power P_i, computation frequency f_i, and block length N_i to maximize its profit. Eventually, we can formulate the following resource optimization problem,

$$\begin{aligned} \max_{P_i, f_i, N_i} \ & \mathcal{G}_i \\ \text{s.t. } & (a): T_i^{en} + T_i^c \leq T^{\max}, \\ & (b): E_i^{en} + E_i^c \leq E_i^{\max}, \\ & (c): \Xi_i \geq \Xi^{\min}, \\ & (d): 0 \leq f_i \leq F_i, \ P_i \geq 0, \\ & (e): N_i \in \{N^{\min}, \cdots, N^{\max}\}, \end{aligned} \tag{12}$$

where $E_i^{\max}$ and F_i are the maximum energy consumption and the maximum computation frequency, respectively. $\Xi^{\min}$ is the allowable lowest transmission security level for the data trading task requirement. $N^{\min}$ and $N^{\max}$ are the minimum and maximum block lengths. In (12), (a) and (b) are the time delay and energy consumption constraints, respectively. (c) restricts the transmission security level. (d) is the constraint of the on-board CPU computation frequency and transmit power of the vehicle i. (e) represents the block length as a discrete variable with the minimum and maximum constraints.

IV. Solution of the Formulated Problem

The problem (12) is a mixed-integer non-linear programming problem (MINLP), as shown in the previous section, which is challenging to solve. We apply the block coordinate descent (BCD) method here to decompose it into three sub-problems to solve separately.

A. Block Length

To solve the problem of obtaining optimal block length, which is a discrete variable, we relax constraint (e) into continuous variable constraint $(1e)$. So the relaxed problem is as follows

$$\begin{aligned} \max_{N_i} \ & W_R \overline{\Upsilon_i^{m-1}} \frac{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)}{N_i} (1 - \mathbb{BER})^{N_i} \\ & (1c): \log_2(N_i) \geq \Xi^{\min}, \\ & (1e): N^{\min} \leq N_i \leq N^{\max}. \end{aligned} \tag{13}$$

We can merge $(1c)$ and $(1e)$ into the following inequality

$$\max\left\{\lceil 2^{\Xi^{\min}} \rceil, N^{\min}\right\} \leq N_i \leq N^{\max}, \tag{14}$$

where $\lceil \cdot \rceil$ means round up that can acquire an integer. Due to $0 < \mathbb{BER} \ll 1$, the objective function in problem (13) decreases monotonically as N_i increases. Therefore, the optimal block length of the vehicle i can be calculated as follows

$$N_i^* = \max\left\{\lceil 2^{\Xi^{\min}} \rceil, N^{\min}\right\}. \tag{15}$$

B. Computation Frequency

We transform the original problem (12) into the following problem to obtain computation frequency.

$$\begin{aligned} \min_{f_i} \ & W_E k_i S_i l_i f_i^2 \\ \text{s.t. } & (2a): \frac{S_i l_i}{f_i} + \frac{S_i}{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)} \leq T^{\max}, \\ & (2b): k_i S_i l_i f_i^2 + \frac{P_i S_i}{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)} \leq E_i^{\max}, \\ & (2d): 0 \leq f_i \leq F_i. \end{aligned} \tag{16}$$

Constraints $(2a)$, $(2b)$ and $(2d)$ can be merged into the following new feasible interval

$$f_i \in \left[\max\left\{0, \frac{S_i l_i R_i}{T^{\max} R_i - S_i}\right\}, \min\left\{F_i, \sqrt{\frac{E_i^{\max} R_i - P_i S_i}{R_i k_i S_i l_i}}\right\}\right]. \tag{17}$$

The objective function in problem (16) is a parabola, and its minimum value can be obtained at the lower bound of the feasible interval. Therefore, the optimal computation frequency of the vehicle i can be calculated as follows

$$f_i^* = \max\left\{0, \frac{S_i l_i R_i}{T^{\max} R_i - S_i}\right\}. \tag{18}$$

C. Transmit Power

Under given block length N_i and computation frequency f_i, the original problem (12) can be simplified to

$$\begin{aligned} \max_{P_i} \ & W_R \overline{\Upsilon_i^{m-1}} \frac{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)}{N_i} (1 - \mathbb{BER})^{N_i} \\ & - W_E \left(k_i S_i l_i f_i^2 + \frac{P_i S_i}{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)}\right) \\ \text{s.t. } & (3a): \frac{S_i l_i}{f_i} + \frac{S_i}{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)} \leq T^{\max}, \\ & (3b): k_i S_i l_i f_i^2 + \frac{P_i S_i}{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)} \leq E_i^{\max}, \\ & (3d): P_i \geq 0, \end{aligned} \tag{19}$$

which is hard to be proved as a convex problem about P_i. But the constraints $(3a)$, $(3b)$ and $(3d)$ in (19) are convex sets, respectively, as proved in Appendix A. We set

$$f(P_i) = W_R \overline{\Upsilon_i^{m-1}} \frac{B \log_2\left(1 + \frac{P_i |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)}{N_i} (1 - \mathbb{BER})^{N_i}, \tag{20}$$

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3233035

and

$$h(P_i) = W_E\left(k_i S_i l_i f_i^2 + \frac{P_i S_i}{B\log_2\left(1+\frac{P_i|h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER})BN_0}\right)}\right). \tag{21}$$

$f(P_i)$ is a logarithmic function with respect to the variable Pi, which is a concave function. $h(P_i)$ is also a concave function [31], [32]. The following inequality holds for function $h(P_i)$,

$$h(P_i) \le h(P_i^l) + \nabla h^T(P_i^l)(P_i - P_i^l), \tag{22}$$

where $\nabla h^T(P_i^l)$ is the value of the derivative function of function $h(P_i)$ at P_i shown in (23).

It is known that $f(P_i) - (h(P_i^l) + \nabla h^T(P_i^l)(P_i - P_i^l))$ is always smaller than $f(P_i) - h(P_i)$. Therefore, the MM algorithm can be applied here to iteratively obtain the maximum values of the minorization functions to approximate the objective value. We can convert these maximization problems into minimization forms as follows to solve them conveniently,

$$\begin{aligned}&\min_{P_i}\ \ h(P_i^l) + \nabla h^T(P_i^l)(P_i - P_i^l) - f(P_i)\\ &\text{s.t. } (3a), (3b), (3d).\end{aligned} \tag{24}$$

Let $\chi_i = -\frac{S_i\Gamma(\mathbb{BER})N_0\ln 2}{(E_i^{max}-k_i S_i l_i f_i^2)|h_i|^2 d^{-\beta}}$ and $\zeta_i = \frac{\Gamma(\mathbb{BER})BN_0}{|h_i|^2 d_i^{-\beta}}$, we obtain the following equations from $(3a)$ and $(3b)$ in the lth iteration

$$\begin{aligned}P_{i,\min}^l &= \left(\max\left\{2^{\frac{S_i}{\left(T^{\max}-\frac{S_i l_i}{f_i}\right)B}},\ \frac{1}{\chi_i}\mathrm{W}_0(\chi_i e^{\chi_i})\right\} - 1\right)\zeta_i,\\ P_{i,\max}^l &= \left(\frac{1}{\chi_i}\mathrm{W}_{-1}(\chi_i e^{\chi_i}) - 1\right)\zeta_i.\end{aligned} \tag{25}$$

So, the newly merged constraint is as follows

$$P_{i,\min}^l \le P_i \le P_{i,\max}^l. \tag{26}$$

The objective function in problem (24) is a linear function that minuses the concave function, so it is a convex function that can be differentiated to obtain a minimal value. We represent the variable in problem (24) that minimizes the objective function as $\hat{P}_i$,

$$\hat{P}_i = \left(\frac{W_R\overline{\Upsilon_i^{m-1}}B(1-\mathbb{BER})^{N_i}|h_i|^2 d_i^{-\beta}}{\nabla h^T(P_i^l)N_i\ln 2\ \Gamma(\mathbb{BER})BN_0} - 1\right)\zeta_i. \tag{27}$$

Combining (25), (26) and (27), we can obtain the variable that minimizes the lth problem as follows, and it also becomes the $(l+1)$th iteration parameter,

$$P_i^{l+1} = \begin{cases} P_{i,\min}^l, & \text{if } \hat{P}_i \le P_{i,\min}^l,\\ \hat{P}_i, & \text{if } P_{i,\min}^l < \hat{P}_i \le P_{i,\max}^l,\\ P_{i,\max}^l, & \text{if } \hat{P}_i > P_{i,\max}^l.\end{cases} \tag{28}$$

The specific iterative process of the MM algorithm to obtain optimal transmit power is described in Algorithm 2.

Algorithm 2 Minorization-Maximization (MM)-based Transmit Power Allocation Algorithm

Initialization:
- Initialize the maximum number of iterations $l_{\max}$ and the specified precision ϵ.
- Set the initial P_i^0.
- Let $l = 0$, and compute $\mathcal{G}^0(P_i) = f(P_i^0) - h(P_i^0)$.

Iteration:

while $l \le l_{\max}$ **do**
 According to the equation (28), solving convex optimization problem (24) to obtain the optimal solution P_i^{l+1}.
 Compute $\mathcal{G}^{l+1}(P_i) = f(P_i^{l+1}) - h(P_i^{l+1})$.
 if $|\mathcal{G}(P_i^{l+1}) - \mathcal{G}(P_i^l)| < \epsilon$ **then**
  break.
 end if
 $l = l + 1$.
end while

Output: the optimal transmission power $P_i^* = P_i^{l+1}$.

D. Complexity Analysis

For the resource optimization of vehicle data sellers, since the closed-form solution (28) is derived from the minimum optimization problem (24), let the iteration number of Algorithm 2 be M_1, the algorithm complexity of the MM-based transmit power allocation algorithm is $O(M_1)$ [33]. We also obtain the closed-form solution (15) for the block length and the closed-form solution (18) for the computation frequency. Therefore, the computational complexity of the BCD method is $O(M_1 M_2)$, where M_2 is the iterative number of the BCD.

For the MAB-based online reputation management algorithms of edge server data buyers, except for the maximum value operation in Step 6, the rest are simple numerical operations. The algorithm complexity of the maximum operation is $O(I)$ [34]. So the computational complexity of edge servers to manage reputation is $O(I)$.

The computational complexity of the overall algorithm to select a vehicle for one trading is $O(IM_1M_2)$, which is low thanks to the derivation of the closed-form solutions.

V. Experiment Result

In this section, we evaluate the proposed MAB-based reputation management scheme and resource allocation algorithm for vehicles. We first verify the convergence and effectiveness of the algorithms. Then, the performance indicators such as revenue, utility and energy consumption are compared and analyzed.

A. Simulation Parameters

The simulation is implemented on a laptop with 16-GB RAM, where the CPU is Intel Core i7-10510U with 1.8GHz. All experiments were repeated 10000 times and averaged. We consider that there are $I = 10$ vehicles driving on the road segment, and edge servers running the MAB-based reputation

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3233035

$$\nabla h^T(P_i^l) = \frac{W_E S_i}{B \log_2\left(1 + \frac{P_i^l |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right)} \left[1 - \frac{P_i^l |h_i|^2 d_i^{-\beta}}{\left(1 + \frac{P_i^l |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right) \log_2\left(1 + \frac{P_i^l |h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER}) B N_0}\right) \ln 2 \, \Gamma(\mathbb{BER}) B N_0}\right] \tag{23}$$

TABLE I: Parameter settings in the simulation.

Parameter	Meaning	Value
B	Bandwidth	180 KHz [35]
N_0	Noise power density	-174 dBm/Hz [36]
$\mathbb{BER}$	Desired target bit error rate	10^{-4} [37]
β	Path loss exponent	2
k_i	Capacitance coefficient	10^{-27}
W_R	Utility price	$[1, 9] \times 10^{-4}$
W_E	Energy price	$[1, 9]$
$N_{\min}/N_{\max}$	Minimum/Maximum block length	64 bits / 512 bits
$E_i^{\max}$	Maximum energy consumption	0.6 J
$T^{\max}$	Maximum time delay	[0.5, 4] s
F_i	Maximum computation frequency	2 GHz
S_i	Data size	[20, 200] KB
l_i	Processing Density	[700, 800] cycles/bit
d_i	Communication distance	[10m, 1km]

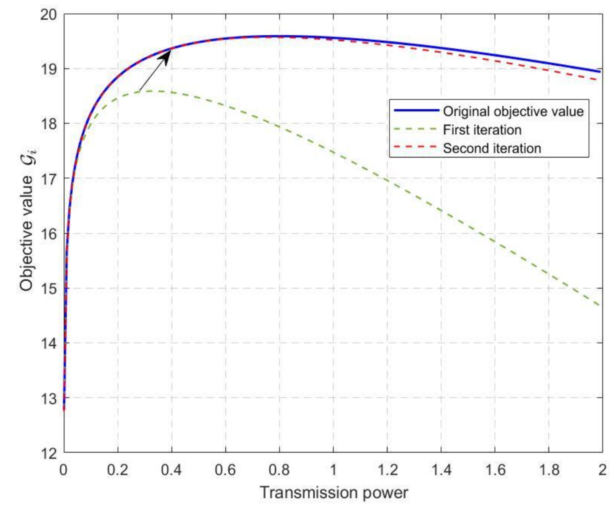
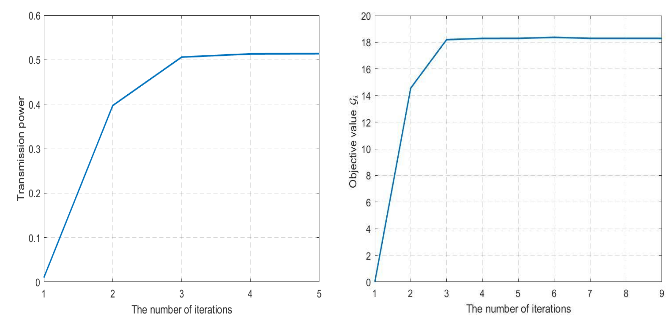


Fig. 2: Principle of Minorization-Maximization Algorithm.



(a) Convergence of Algorithm 2 (b) Convergence of External Loop

Fig. 3: Convergence of Proposed Resource Optimization Algorithms

management for data trading next to the road. They will upload the MAB learning parameters to the cloud to share their reputation experience. The path loss exponent for the edge server to vehicles is 2. Noise power density is -174 dBm/Hz. Detailed parameter settings are listed in Table I. To highlight the advantages of using the MAB algorithm to manage reputation value, we implement the *Original* scheme to select the vehicle with the largest reputation value each time and the *Random* scheme to select the vehicle randomly. We compare the performance with new vehicles added at intervals to highlight the dynamic characteristic of the *Proposed* algorithm. In order to compare the performance of resource optimization, we set up the following three schemes for comparison.

- *RPS* [38]: The scheme randomly sets transmit power and optimizes the remaining variables, such as computation frequency and block length.
- *RFS* [39]: The scheme randomly sets computation frequency and optimizes the remaining variables, such as transmit power and block length.
- *RNS* [40]: The scheme randomly sets block length and optimizes the remaining variables, such as transmit power and computation frequency.

B. Numerical Results

Generally speaking, if resource allocation is involved in the scheme, the solution process will greatly reduce the throughput of data transactions. Commonly used search methods are inefficient, and neural networks need to be trained in advance. But in this paper, since closed-form solutions are obtained, an average decision rate of 3596 data transactions per second can be achieved in the experiment. The processing throughput shows that the proposed scheme is effective and efficient for vehicle reputation management, data trading selection, and resource allocation. Therefore, the proposed method is suitable for large-scale, dynamic, and high-frequency data trading scenarios in VEC.

1) Convergence and Effectiveness of Proposed Algorithms:

The numerical results of this section are to validate the convergence and effectiveness of the proposed algorithms.

Fig. 2 shows the curve changes in the iterations of the MM algorithm in Algorithm 2, which represents the objective $\mathcal{G}_i$ in (11) and iterative approximate curves. It can be seen that as the iteration progresses, the maximum value of the iteration curve is closer to the position of the original target value. This figure also shows the principle process of the MM algorithm. Fig. 3(a) represents that the transmit power will converge into one stable value, which demonstrates the effectiveness of the MM-based transmit power allocation algorithm. Moreover, a stable result can be obtained in just three or four steps at a fast convergence rate. Fig. 3(b) shows the iterative convergence

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3233035

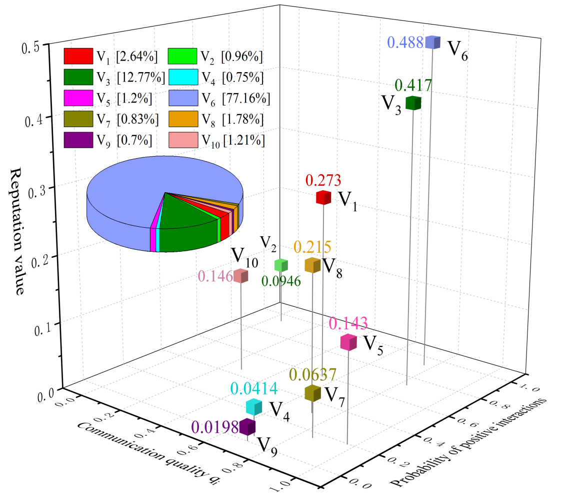


Fig. 4: The effectiveness of the Algorithm 1.

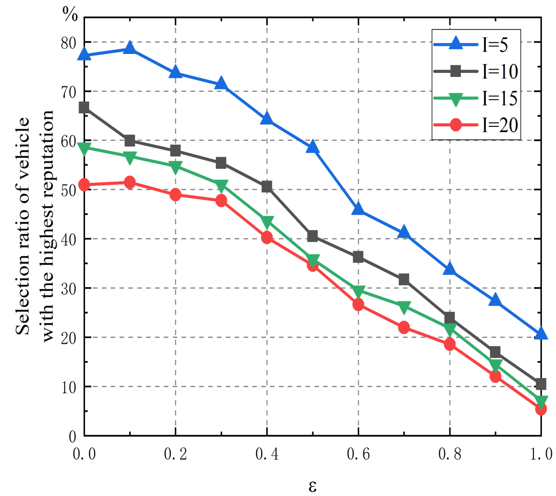


Fig. 5: Selection ratio of the vehicle with the highest reputation under different ε.

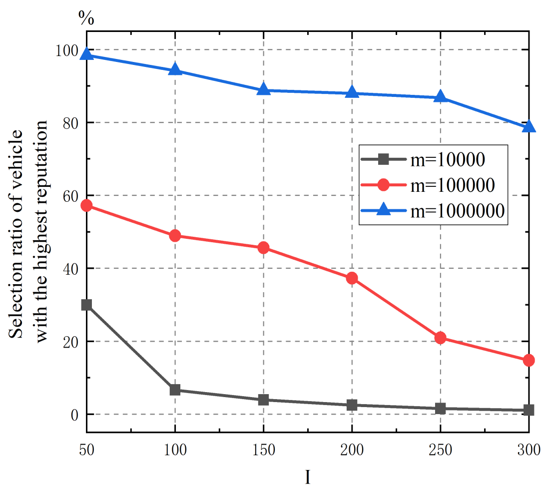


Fig. 6: Selection ratio of the vehicle with the different trading rounds m under different vehicle number I.

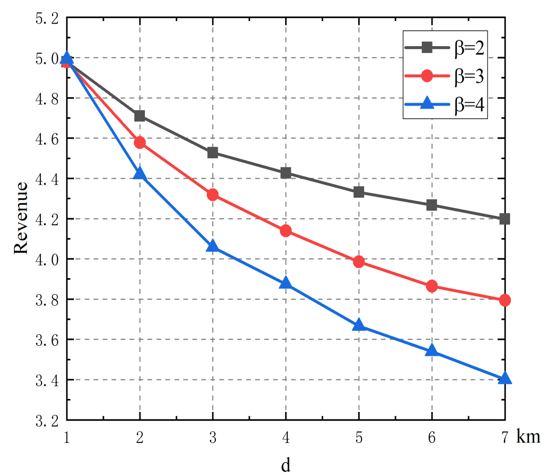


Fig. 7: Revenue of the vehicles with the different path loss exponent β under different communication distances d.

between the transmit power obtained by Algorithm 2 and the other two variables. This external loop validates the fast convergence rate and effectiveness of solving the original problem (12) for resource optimization.

Fig. 4 verifies the effectiveness of the Algorithm 1, which exhibits that the MAB-based online reputation management algorithm makes the edge server more inclined to select vehicles with high communication quality and more positive communication records for data trading. Among the 10000 times selections, the edge servers have a proportion of 77.16% to select the vehicle $V6$ with the dynamically updated highest reputation value. At the same time, it also maintains certain trading opportunities for other vehicles with less reputation value. As for the proportion of vehicles with the highest reputation value chosen, it is closely related to the choice of ε in the Algorithm 1. Given the different number of vehicles, Fig. 5 compares the change in the proportion of vehicles with the highest reputation value that the edge server selects for different ε. The probability that the edge server randomly selects a vehicle is greater when ε becomes larger, so the advantage of vehicle reputation is not obvious. When just a few vehicles are on the road, there are not many choices for edge servers, so vehicles with higher reputation values will be selected for data trading with a greater probability.

Fig. 6 shows the selection ratio of the vehicle with the different trading rounds m under different vehicles number I. For a given number of transactions, as the number of vehicles grows from 50 to 300, the proportion of the highest reputation vehicle selected by the edge server gradually decreases. This is because the MAB algorithm will give all vehicles enough chances to present their true reputation condition. As the number of transaction rounds m increases, the edge

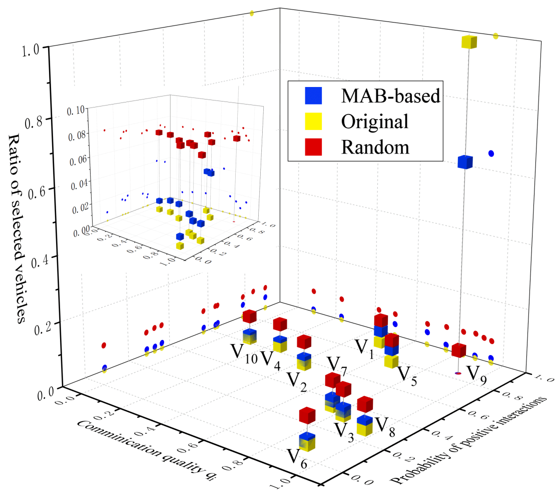


Fig. 8: Selection ratio for initial ten vehicles when dynamically adding vehicles.

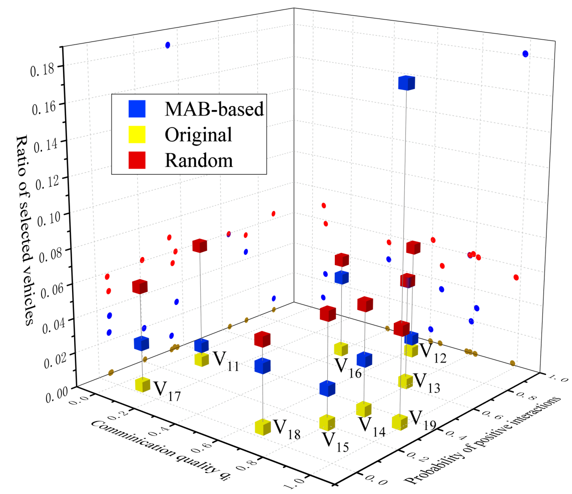


Fig. 9: Selection ratio of newly added vehicles.

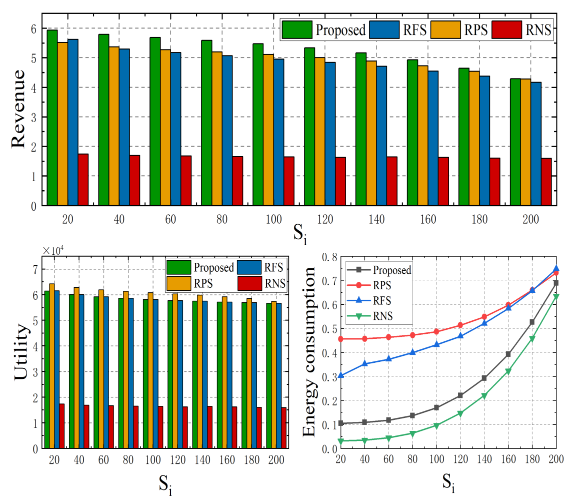


Fig. 10: Revenue, utility and vehicular energy consumption in data trading with different data sizes.

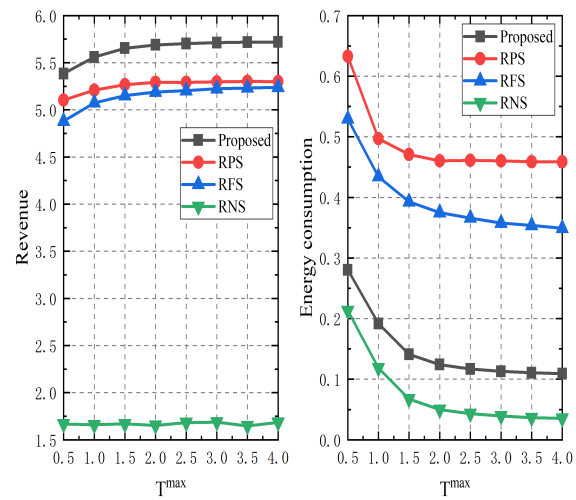


Fig. 11: Revenue and vehicular energy consumption in data trading with different maximum time delay constraints.

server selects the vehicle with the highest reputation in a larger proportion, because the reputation value at this time has accumulated enough experience and transaction attempts. The proposed lightweight algorithm can adapt to the high transaction volume of a large number of vehicles. The revenue of the vehicles with the different path loss exponent β under different communication distances d is shown in Fig. 7. For a given path loss exponent, the farther the vehicle is from the edge server, the lower the revenue, and it will consume more energy and delay. A larger path loss exponent also leads to a reduction in revenue for vehicles.

2) Comparison of Reputation Management Scheme: The numerical results of this section are to compare the performance between *MAB-based*, *Original*, and *Random* reputation management algorithms.

To simulate the characteristics of dynamic vehicle addition in VEC scenarios, we use the initial ten vehicles for 1000 selections and add a new vehicle every 1000 intervals to represent the relationship between the vehicle addition order and the selected ratio. The selection ratio for the initial ten vehicles when dynamically adding vehicles is shown in Fig. 8. The inset zooms in on the portion where the selection ratio is below 0.1. The *Original* scheme always only selects vehicle $V9$ with the highest reputation value and almost ignores other vehicles. The *Random* scheme has an even chance of each vehicle being selected. The proposed *MAB-based* scheme has a high proportion of vehicle $V9$ with the highest reputation value because it has better communication quality and positive transaction records. However, for other vehicles, there is a certain chance of being selected. Fig. 9 shows the ratio at which the next nine vehicles were selected in their respective participations. Since the initial reputation

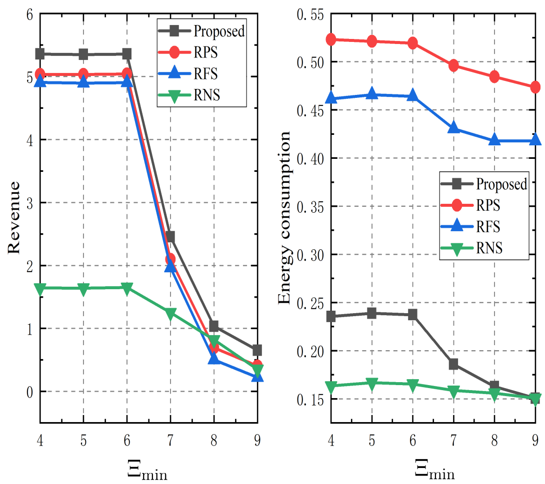


Fig. 12: Revenue and vehicular energy consumption in data trading under different minimum allowable security levels.

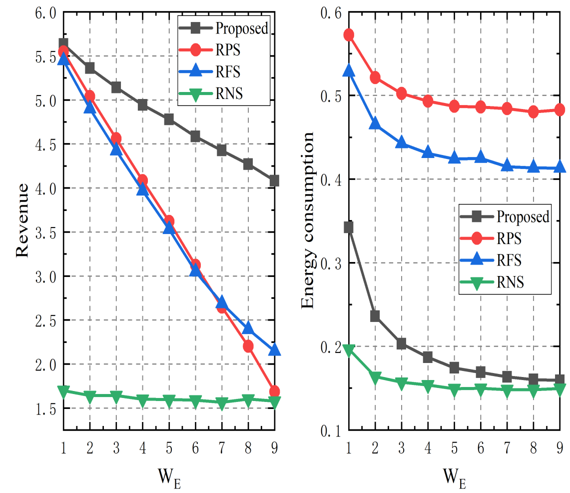


Fig. 13: Revenue and vehicular energy consumption in data trading with different energy consumption price scales.

conditions of the later participating vehicles are not as good as $V9$, the *Original* scheme hardly selects them. The proposed *MAB-based* scheme gives them enough opportunities to update to obtain their real reputation value and does not have less chance than existing vehicles because they are newly added vehicles. *Random* scheme still gives them an even chance to be selected. Obviously, our proposed scheme is more dynamic and adaptive.

3) Comparison of Resource Optimization Scheme: This section is to compare the performance between *Proposed*, *RPS*, *RFS*, and *RNS* resource optimization algorithms. We repeat the optimal allocation process of transmit power, computation frequency, and block length for 10000 vehicles to obtain the average numerical results.

Fig. 10 examines the revenue, utility, and vehicular energy consumption with different data sizes. As the size of the transaction data increases, the utility of the vehicle that comes from the payment of the edge server based on the low-error data acquisition probability will decrease slightly, and the energy consumption will increase accordingly, resulting in a decrease in the unit revenue of the vehicle. Although the proposed scheme is neither the highest utility nor the lowest energy consumption, the combined revenue of the two is the highest. *RNS* has the lowest energy consumption, but its utility is extremely low. *RPS* has the highest utility but causes the heaviest energy consumption. Fig. 11 investigates revenue and vehicular energy consumption in data trading with different maximum time delay constraints $T^{\max}$. As the latency requirement of the data trading task decreases, the revenue of the vehicle will increase, and the energy consumption will be significantly decreased, but the reduction will gradually slow down and stabilize. *RNS* still has the lowest energy consumption and revenue. Therefore, to improve the vehicle's revenue under the constraints of delay and energy consumption, it is essential to optimize the block length N_i.

Fig. 12 demonstrates revenue and vehicular energy consumption in data trading under different allowable lowest security levels $\Xi_{\min}$. With the increase of $\Xi_{\min}$, the revenue and energy consumption of the vehicle will be reduced. So the energy consumption of vehicles must be appropriately sacrificed to increase utility to achieve higher revenue. For vehicles, it is reluctant to see an increase in the security level requirement, because this will lead to a significant reduction in utility. However, when security requirements are particularly low, the vehicle guarantees a certain degree of minimum transmission security level. Fig. 13 shows the revenue and energy consumption in data trading with different energy consumption price scales W_E. To achieve a balance between utility and energy consumption when W_E increases, the vehicle will reduce energy consumption to realize the deceleration of revenue. This means that vehicles want to reduce energy costs as unit energy becomes more expensive in real life.

In summary, the proposed scheme can effectively manage vehicle reputation, select data trading, and allocate resources, and it is suitable for large-scale, dynamic, and high-frequency data trading scenarios in VEC.

VI. Conclusion

In this work, we propose MAB-based reputation management to select the vehicle for data trading to ensure the credibility and reliability of the data in VEC. Resource allocation algorithms are embedded to maximize revenue for vehicles under the constraints of time delay, energy consumption, and transmission security level. Relying on the MAB approach, the vehicle reputation status can be updated online dynamically, and all vehicles have enough opportunities to trade data, while a new vehicle can still have a chance to obtain the actual and variational reputation value. Also, the edge servers can securely acquire high-quality and low-error data traffic. Encrypting trading data can protect the privacy of users and ensure the interests of edge servers. We jointly optimize the transmit power, computation frequency, and block length to higher utility and lower energy consumption. For the block length and computation frequency, we obtain the closed-form solution, and a MM-based algorithm is developed to allocate transmit power. The experiment results verify the convergence and effectiveness of the proposed algorithms and know that the scheme is effective and efficient for vehicle reputation management, data trading selection, and resource allocation. How to optimize the security and resource problems for more complex data trading scenarios in vehicle networking with lightweight and effective methods suitable for dynamic VEC requires further research. There will be more complicated trading models and security relationships.

Appendix A
Proof of Convex Sets in Problem (19).

Constraint $(3a)$ in problem (19) can be transformed as follows

$$B\log_2\left(1+\frac{P_i|h_i|^2 d_i^{-\beta}}{\Gamma(\mathbb{BER})BN_0}\right)\geq\frac{S_i}{T^{\max}-\frac{S_i l_i}{f_i}},\tag{29}$$

where the left formula is a logarithmic function with respect to the variable P_i. The upper level set of a concave function is a convex set, so constraint $(3a)$ as a convex set is proved.

Constraint $(3b)$ in problem (19) can be transformed as follows

$$P_iS_i-(E_i^{\max}-k_iS_il_if_i^2)B\log_2\left(1+\frac{P_i|h_i|^2d_i^{-\beta}}{\Gamma(\mathbb{BER})BN_0}\right)\leq 0,\tag{30}$$

where the twice derivative function of the formula on the left with respect to P_i is

$$(E_i^{\max}-k_iS_il_if_i^2)\frac{\left(\frac{|h_i|^2d_i^{-\beta}}{\Gamma(\mathbb{BER})N_0}\right)^2}{B\ln 2\left(1+\frac{P_i|h_i|^2d_i^{-\beta}}{\Gamma(\mathbb{BER})BN_0}\right)^2}.\tag{31}$$

The above function is greater than zero, so constraint $(3b)$ as a convex set is proved.

Hitherto, the constraints in the problem (19) are all convex sets.

References

[1] C. Kaiser, A. Stocker, G. Viscusi, M. Fellmann, and A. Richter, "Conceptualising value creation in data-driven services: The case of vehicle data," *International journal of information management*, vol. 59, p. 102335, 2021.

[2] F. Liang, W. Yu, D. An, Q. Yang, X. Fu, and W. Zhao, "A survey on big data market: Pricing, trading and protection," *Ieee Access*, vol. 6, pp. 15 132–15 154, 2018.

[3] S. A. Ghasempouri and B. T. Ladani, "Modeling trust and reputation systems in hostile environments," *Future Generation Computer Systems*, vol. 99, pp. 571–592, 2019.

[4] Z. Tian, X. Gao, S. Su, J. Qiu, X. Du, and M. Guizani, "Evaluating reputation management schemes of internet of vehicles based on evolutionary game theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5971–5980, 2019.

[5] J. Feng, L. Liu, Q. Pei, and K. Li, "Min-max cost optimization for efficient hierarchical federated learning in wireless edge networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 11, pp. 2687–2700, 2022.

[6] C. Gao, G. Wang, W. Shi, Z. Wang, and Y. Chen, "Autonomous driving security: State of the art and challenges," *IEEE Internet of Things Journal*, 2021.

[7] S. Ramanujam and M. Karuppiah, "Designing an algorithm with high avalanche effect," *IJCSNS International Journal of Computer Science and Network Security*, vol. 11, no. 1, pp. 106–111, 2011.

[8] S. Takeuchi, M. Hasegawa, K. Kanno, A. Uchida, N. Chauvet, and M. Naruse, "Dynamic channel selection in wireless communications via a multi-armed bandit algorithm using laser chaos time series," *Scientific reports*, vol. 10, no. 1, pp. 1–7, 2020.

[9] X. Cao, Y. Chen, and K. J. R. Liu, "Data trading with multiple owners, collectors, and users: An iterative auction mechanism," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 2, pp. 268–281, 2017.

[10] X. Huang, S. Garg, J. Nie, W. Y. B. Lim, Y. Qi, Y. Zhang, and M. S. Hossain, "Toward efficient data trading in ai enabled reconfigurable wireless sensor network using contract and game theories," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 1, pp. 98–108, 2022.

[11] Y. Zhan, Y. Xia, Y. Liu, F. Li, and Y. Wang, "Incentive-aware time-sensitive data collection in mobile opportunistic crowdsensing," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 7849–7861, 2017.

[12] C. Chen, J. Wu, H. Lin, W. Chen, and Z. Zheng, "A secure and efficient blockchain-based data trading approach for internet of vehicles," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 9110–9121, 2019.

[13] W. Mao, Z. Zheng, and F. Wu, "Pricing for revenue maximization in iot data markets: An information design perspective," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 2019, pp. 1837–1845.

[14] S. Su, Z. Tian, S. Liang, S. Li, S. Du, and N. Guizani, "A reputation management scheme for efficient malicious vehicle identification over 5g networks," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 46–52, 2020.

[15] Z. Ying, M. Ma, Z. Zhao, X. Liu, and J. Ma, "A reputation-based leader election scheme for opportunistic autonomous vehicle platoon," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 3519–3532, 2022.

[16] Z. Tian, X. Gao, S. Su, and J. Qiu, "Vcash: A novel reputation framework for identifying denial of traffic service in internet of connected vehicles," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 3901–3909, 2020.

[17] H. Chai, S. Leng, K. Zhang, and S. Mao, "Proof-of-reputation based-consortium blockchain for trust resource sharing in internet of vehicles," *IEEE Access*, vol. 7, pp. 175 744–175 757, 2019.

[18] Y. Liu, H. Yu, S. Xie, and Y. Zhang, "Deep reinforcement learning for offloading and resource allocation in vehicle edge computing and networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 11 158–11 168, 2019.

[19] X.-Q. Pham, T.-D. Nguyen, V. Nguyen, and E.-N. Huh, "Joint node selection and resource allocation for task offloading in scalable vehicle-assisted multi-access edge computing," *Symmetry*, vol. 11, no. 1, 2019. [Online]. Available: https://www.mdpi.com/2073-8994/11/1/58

[20] H. Xu, W. Huang, Y. Zhou, D. Yang, M. Li, and Z. Han, "Edge computing resource allocation for unmanned aerial vehicle assisted mobile network with blockchain applications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3107–3121, 2021.

[21] X. Huang, R. Yu, J. Kang, Z. Xia, and Y. Zhang, "Software defined networking for energy harvesting internet of things," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1389–1399, 2018.

[22] J. Kang, Z. Xiong, D. Niyato, D. Ye, D. I. Kim, and J. Zhao, "Toward secure blockchain-enabled internet of vehicles: Optimizing consensus management using reputation and contract theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2906–2920, 2019.

[23] S. O. Sharif and S. Mansoor, "Performance analysis of stream and block cipher algorithms," in *2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE)*, vol. 1, 2010, pp. V1–522–V1–525.

[24] S. Deb, S. K. Ghosh, and S. C. Ghosh, "A multi-arm-bandit based resource block allocation in ris assisted wireless network," in *2021 IEEE 20th International Symposium on Network Computing and Applications (NCA)*. IEEE, 2021, pp. 1–6.

[25] R. Kerkouche, R. Alami, R. Féraud, N. Varsier, and P. Maillé, "Node-based optimization of lora transmissions with multi-armed bandit algorithms," in *2018 25th International Conference on Telecommunications (ICT)*. IEEE, 2018, pp. 521–526.

[26] M. M. Drugan and A. Nowe, "Designing multi-objective multi-armed bandits algorithms: A study," in *The 2013 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2013, pp. 1–8.

[27] S. Agrawal and N. Goyal, "Analysis of thompson sampling for the multi-armed bandit problem," in *Conference on learning theory*. JMLR Workshop and Conference Proceedings, 2012, pp. 39–1.

[28] Y. Seldin, C. Szepesvári, P. Auer, and Y. Abbasi-Yadkori, "Evaluation and analysis of the performance of the exp3 algorithm in stochastic environments," in *European Workshop on Reinforcement Learning*. PMLR, 2013, pp. 103–116.

[29] J. Zhang, W. Xia, F. Yan, and L. Shen, "Joint computation offloading and resource allocation optimization in heterogeneous networks with mobile edge computing," *IEEE Access*, vol. 6, pp. 19 324–19 337, 2018.

[30] H. Ko, S. Oh, B. Kim, and C. Kim, "Simple bit allocation algorithms with ber-constraint for ofdm-based systems," in *2009 IEEE Wireless Communications and Networking Conference*. IEEE, 2009, pp. 1–5.

[31] Y. Mao, J. Zhang, and K. B. Letaief, "Joint task offloading scheduling and transmit power allocation for mobile-edge computing systems," in *2017 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2017, pp. 1–6.

[32] P. Di Lorenzo, S. Barbarossa, and S. Sardellitti, "Joint optimization of radio resources and code partitioning in mobile edge computing," *arXiv preprint arXiv:1307.3835*, 2013.

[33] Z. Chu, W. Hao, P. Xiao, D. Mi, Z. Liu, M. Khalily, J. R. Kelly, and A. P. Feresidis, "Secrecy rate optimization for intelligent reflecting surface assisted mimo system," *IEEE transactions on information forensics and security*, vol. 16, pp. 1655–1669, 2020.

[34] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms*. MIT press, 2022.

[35] L. Liu, M. Zhao, M. Yu, M. A. Jan, D. Lan, and A. Taherkordi, "Mobility-aware multi-hop task offloading for autonomous driving in vehicular edge computing and networks," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

[36] A. Feki and V. Capdevielle, "Autonomous resource allocation for dense lte networks: A multi armed bandit formulation," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*. IEEE, 2011, pp. 66–70.

[37] H. Seo and B. G. Lee, "A proportional-fair power allocation scheme for fair and efficient multiuser ofdm systems," in *IEEE Global Telecommunications Conference, 2004. GLOBECOM'04.*, vol. 6. IEEE, 2004, pp. 3737–3741.

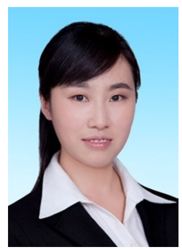
[38] A. A. Nasir, D. T. Ngo, X. Zhou, R. A. Kennedy, and S. Durrani, "Joint resource optimization for multicell networks with wireless energy harvesting relays," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6168–6183, 2016.

[39] J. Feng, W. Zhang, Q. Pei, J. Wu, and X. Lin, "Heterogeneous computation and resource allocation for wireless powered federated edge learning systems," *IEEE Transactions on Communications*, vol. 70, no. 5, pp. 3220–3233, 2022.

[40] H. Xiao, Q. Pei, X. Song, and W. Shi, "Authentication security level and resource optimization of computation offloading in edge computing systems," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

Huizi Xiao is currently pursuing a Ph.D. degree in Communication and Information Systems at Xidian University, Xi'an, China. From 2021 to 2023, she was with the University of Victoria, BC, Canada, as a visiting Ph.D. student. Her current research interests include edge computing, resource allocation, security and privacy, convex optimization, and stochastic network optimization.

Lin Cai (Fellow, IEEE) received her M.A.Sc. and Ph. D. degrees in electrical and computer engineering from the University of Waterloo, Waterloo, Canada, in 2002 and 2005, respectively. Since 2005, she has been with the Department of Electrical & Computer Engineering at the University of Victoria, and she is currently a Professor. She is an NSERC E.W.R. Steacie Memorial Fellow, an Engineering Institute of Canada (EIC) Fellow, and an IEEE Fellow. In 2020, she was elected as a Member of the Royal Society of Canadas College of New Scholars, Artists and Scientists, and a 2020 Star in Computer Networking and Communications by N2Women. Her research interests span several areas in communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia traffic and the Internet of Things. She is the Distinguished Lecturer of the IEEE Vehicular Technology Society and the IEEE Communications Society.



Jie Feng (Member, IEEE) received her Ph.D. degree in communication and information systems from Xidian University, China, in 2020. She is currently an Associate Professor at the Department of Electrical Engineering and Computer Science, Xidian University, Xian, China. From 2018 to 2019, she was with Carleton University, Ottawa, ON, Canada, as a visiting Ph.D. student. Her current research interests include mobile-edge computing, blockchain, deep reinforcement learning, the device to device communication, resource allocation and convex optimization, and stochastic network optimization.

Qingqi Pei (Member, IEEE) received his B.S., M.S., and Ph.D. degrees in Computer Science and Cryptography from Xidian University in 1998, 2005, and 2008, respectively. He is currently a Professor and a member of the State Key Laboratory of Integrated Services Networks, also a Professional Member of ACM and a Senior Member of IEEE, Senior Member of Chinese Institute of Electronics and China Computer Federation. His research interests focus on privacy preserving, blockchain and edge computing security.

Weisong Shi (Fellow, IEEE) is a Professor and Chair of the Department of Computer and Information Sciences at the University of Delaware (UD), where he leads the Connected and Autonomous Research (CAR) Laboratory. Dr. Shi is an internationally renowned expert in edge computing, autonomous driving and connected health. His pioneer paper entitled Edge Computing: Vision and Challenges has been cited more than 5000 times. Before he joined UD, he was a professor at Wayne State University (2002-2022). Dr. Shi has published more than 270 articles in peer-reviewed journals and conferences and served in editorial roles for more than 10 academic journals and publications, including EIC of Smart Health, AEIC of IEEE Internet Computing Magazine. He is a fellow of IEEE, and a distinguished member of ACM. More information can be found at http://weisongshi.org.