BIG DATA MINING AND ANALYTICS
ISSN 2096-0654 01/10 pp391–403
Volume 6, Number 4, December 2023
DOI: 10.26599/BDMA.2022.9020041

VDCM: A Data Collection Mechanism for Crowd Sensing in Vehicular Ad Hoc Networks

Juli Yin, Linfeng Wei*, Zhiquan Liu, Xi Yang, Hongliang Sun, Yudan Cheng, and Jianbin Mai

Abstract: With the rapid development of mobile devices, aggregation security and efficiency topics are more important than past in crowd sensing. When collecting large-scale vehicle-provided data, the data transmitted via autonomous networks are publicly accessible to all attackers, which increases the risk of vehicle exposure. So we need to ensure data aggregation security. In addition, low aggregation efficiency will lead to insufficient sensing data, making the data unable to provide data mining services. Aiming at the problem of aggregation security and efficiency in large-scale data collection, this article proposes a data collection mechanism (VDCM) for crowd sensing in vehicular ad hoc networks (VANETs). The mechanism includes two mechanism assumptions and selects appropriate methods to reduce consumption. It selects sub mechanism 1 when there exist very few vehicles or the coalition cannot be formed, otherwise selects sub mechanism 2. Single aggregation is used to collect data in sub mechanism 1. In sub mechanism 2, cooperative vehicles are selected by using coalition formation strategy and auction cooperation agreement, and multi aggregation is used to collect data. Two sub mechanisms use Paillier homomorphic encryption technology to ensure the security of data aggregation. In addition, mechanism supplements the data update and scoring steps to increase the amount of available data. The performance analysis shows that the mechanism proposed in this paper can safely aggregate data and reduce consumption. The simulation results indicate that the proposed mechanism reduces time consumption and increases the amount of available data compared with existing mechanisms.

Key words: vehicular ad hoc networks (VANETs); crowd sensing; data collection; data aggregation security

1 Introduction

Big data are produced from multiple sources in different formats at very high speed[1]. Therefore, big data are difficult to obtain using traditional technologies. With the development of sensor technologies and the improvement of embedded computing devices, mobile crowd sensing (MCS) was first proposed as a new sensing mode by Ganti et al.[2] in 2011. This sensing mode combines the idea of crowd sourcing and uses the excellent sensing ability of mobile devices to obtain data. Despite its benefits, MCS experiences various security challenges against various security threats such as disinformation attack, data tampering attack, replay attack, eavesdropping attack, and denial of service (DoS) attack.

MCS is considered as an integrated part of the Internet of things (IoT) based services and applications[3]. With

- Juli Yin, Linfeng Wei, Zhiquan Liu, Hongliang Sun, Yudan Cheng, and Jianbin Mai are with the College of Cyber Security, Jinan University, Guangzhou 510632, China. E-mail: yjlyxx8@163.com; weilinuuu@163.com; zqliu@vip.qq.com; 776367476@qq.com; yxx775@163.com; yxx886@163.com.
- Zhiquan Liu is also with the Guangdong Provincial Key Laboratory of Cyber and Information Security Vulnerability Research, Guangzhou 510643, China.
- Xi Yang is with the College of Information Science Technology, Jinan University, Guangzhou 510632, China. E-mail: 1071402781@qq.com.

* To whom correspondence should be addressed.

Manuscript received: 2022-06-20; revised: 2022-09-18; accepted: 2022-10-19

the increasing demand for social events, crowd sensing is gradually applied to the vehicular ad hoc network (VANET)[4]. Mobile crowd sensing in VANET refers to that mobile vehicles use advanced sensing devices (e.g., wireless communication receiver, onboard sensor, tachograph, GPS, and camera) to collect and transmit sensing data in real time. If vehicles submit their sensing data to the processing center without any pre-processing, the sensing data may be abused or leaked by the processing center. Besides, some vehicles may also try to trick the processing center by providing false data. Data trustworthiness is of paramount importance when data mining techniques work with data in the context of secure and privacy preserving MCS[3]. Therefore, research on data collection mechanisms for mobile crowd sensing in VANET has become one of the most urgent issues to protect both the privacy of vehicle-provided data and reliability of data. Meanwhile, aggregation efficiency is also important. Data sources must be large enough to provide data mining services.

1.1 Related work

Due to the demand for huge amounts of data from crowd sensing in vehicular ad hoc networks, many vehicles need to participate in crowd sensing. During vehicle large-scale data collection, the security of data aggregation has also become the key issue for crowd sensing in VANET. Data are usually strongly related to vehicle information, thus if the data are leaked out, it may pose a threat to the safety of participating vehicles[5]. In addition, if the data are maliciously changed, it will hurt the reputation of the vehicle and reduce the enthusiasm of the vehicle to participate.

Many scholars at home and abroad have made a lot of research on the security of aggregated data in data collection. Basudan et al.[6] proposed a lightweight certificateless aggregation signcryption (CALASS) mechanism which is able to ensure the security of single data for fog computing. However, this mechanism fails to consider the security of data aggregation. Sookhak et al.[7] provided an identity based encryption (IBE) mechanism which is able to encrypt the data with a key and aggregates the data by the service provider. The IBE mechanism fails to consider reputation management and collection efficiency, which greatly limits its applications. Sun et al.[8] proposed a fog bus based vehicle crowd sensing (FBVCS) reporting privacy preservation mechanism which is able to realize data privacy and data aggregation. And compared with other mechanisms, FBVCS mechanism improves the amount of available data. However, this mechanism fails to guarantee the fog buses are credible. There is some additional consumption in the mechanism and the collection time is too long. Qian et al.[9] proposed the aggregation mechanism of homomorphic encryption method based on elliptic curves cryptography (ECC) which optimizes the collection time and protects the security of aggregated data. However, the theory of ECC homomorphic encryption method used in this mechanism has not been officially confirmed, and the security needs to be verified.

To ensure that service providers are provided a large amount of effective data, two questions need to be solved. The first question is the low collection efficiency which will lead to the fact that received data cannot provide effective data services and will result in insufficient sensing data samples. Another question is false data which will pollute the data results. Data aggregation is a common method to improve collection efficiency, and the excellent data aggregation mechanisms can help service providers collect available data safely and efficiently.

In recent years, scholars at home and abroad have done a lot of research on the aggregation efficiency of data collection. There are several kinds of aggregation mechanisms commonly used in wireless sensor networks (WSN) such as tree aggregation, cluster aggregation, chain aggregation, and unstructured method. Kuo et al.[10] proposed a tree aggregation mechanism which is able to use tree aggregation to construct a data aggregation tree to reduce the total energy consumption of data transmission. Naranjo et al.[11] proposed a method to find energy saving mechanism to select cluster head in WSN which is able to prolong the network life. Javaid et al.[12] proposed an application-oriented chain aggregation mechanism which is able to use mathematical modeling to find the local optimal path in the independent chain in each mechanism, so as to find the global optimal path for data transmission through its interconnection. Liu and Gao[13] proposed a hybrid structure aggregation mechanism which is able to minimize the monitoring overhead. However, these mechanisms cannot ensure aggregation data security. Rabieh et al.[14] proposed a homomorphic encryption self organization (HESO) mechanism based on homomorphic encryption which provides a data aggregation mechanism. It is able to reduce time consumption and ensure aggregation data security.

However, this work fails to ensure data security and the amount of available data.

How to efficiently collect data on the premise of ensuring data security is one of the challenges faced by the current crowd sensing in VANET[15]. At present, there are many studies on the data collection of crowd sensing in vehicular ad hoc networks, but still exist unsolved problems such as low efficiency in data collection and aggregation security. And few studies consider the reduction of consumption and the increase of the amount of available data when considering the security of data aggregation.

1.2 Contribution

For crowd sensing in VANET, vehicles collect data through wireless networks. When making the data collection, vehicles are vulnerable to various attacks which result in privacy disclosure. In addition, there are some problems in the collection process such as data delay, data breach, and malicious data. This paper proposes a data collection mechanism named VDCM for the crowd sensing in vehicular ad hoc networks which is able to ensure the security of data aggregation, reduce consumption, and increase the amount of available data. The mechanism is based on Paillier homomorphic encryption technology, coalition formation strategy, and auction cooperation agreement. Two assumptions have been put forward based on specific situation to reduce the time consumption when ensuring the security of data aggregation. The VDCM mechanism improves existing mechanisms, supplements data update and scoring steps, and increases the amount of available data. Compared with the existing studies, the VDCM considers the reduction of consumption and the increase of the amount of available data when considering the security of data aggregation.

1.3 Road map

The remainder of the paper is organized as follows. In Section 2, we present some preliminaries such as coalition formation strategy and homomorphic encryption technology. In Section 3, we introduce the data collection mechanism which is proposed in this paper. Initially we elicit the system architecture and then present the workflow of the mechanism in detail. Before summarizing this paper in Section 5, we provide the performance analysis and simulation result analysis in Section 4.

2 Preliminary

The following subsections briefly introduce technologies used in this study, including coalition formation strategy and Paillier homomorphic encryption.

2.1 Coalition formation strategy

The user selects the coalition according to their preference[16] and sets the coalition as CO_1, CO_2, and CO_3. $\succ_n$ is defined as the binary relationship on all feasible coalitions that users n may form. If $\mathrm{CO}_1 \succ_n \mathrm{CO}_2$, users will get more reward from CO_1, and they are more inclined to join coalition CO_1. There is $\mathrm{CO}_1 \succ_n \mathrm{CO}_3$ according to the transferability. The preference order affects the final coalition structure. The coalition considers two different preference orders.

Pareto order[17]: For a user n and two coalitions CO_1 and CO_2, the Pareto order defines the following actions as

$$\begin{aligned}&\mathrm{CO}_1 \succ_n \mathrm{CO}_2 \Leftrightarrow\\&E_n(\mathrm{CO}_1) < E_n(\mathrm{CO}_2) \wedge\\&E_i(\mathrm{CO}_1) \leqslant E_i(\mathrm{CO}_1 \backslash \{n\}) \forall i \in \mathrm{CO}_1 \backslash \{n\} \wedge\\&E_i(\mathrm{CO}_2 \backslash \{n\}) \leqslant E_i(\mathrm{CO}_2) \forall i \in \mathrm{CO}_2 \backslash \{n\}\end{aligned} \tag{1}$$

Utilitarian order[17]: For a user n and two coalitions CO_1 and CO_2, the Utilitarian order defines the following actions as

$$\begin{aligned}&\mathrm{CO}_1 \succ_n \mathrm{CO}_2 \Leftrightarrow\\&\sum\nolimits_{i \in \mathrm{CO}_1} E_i(\mathrm{CO}_1) - \sum\nolimits_{i \in \mathrm{CO}_1 \backslash \{n\}} E_i(\mathrm{CO}_1 \backslash \{n\}) \leqslant\\&\sum\nolimits_{i \in \mathrm{CO}_2} E_i(\mathrm{CO}_2) - \sum\nolimits_{i \in \mathrm{CO}_2 \backslash \{n\}} E_i(\mathrm{CO}_2 \backslash \{n\})\end{aligned} \tag{2}$$

where $E(\mathrm{CO})$ is utility of coalition CO. In Pareto order or Utilitarian order, when users change the coalition, users will never damage utilities of other users. In Utilitarian order, users will never damage total utility of coalition when they change coalitions, but the utility of every user could not be ensured. For the coalition, it is important that user will not damage utilities of other users. This attribute ensures that the utility of the coalition will not be reduced, and ensures the existence of stable coalition partitions. Besides, multiple coalitions can form a new coalition. The formed behavior is defined as a consolidation rule. The consolidation rule is

$$\begin{aligned}&\{\mathrm{CO}_1 \cup \mathrm{CO}_2\} \Leftrightarrow\\&[\forall i \in \mathrm{CO}_1, (\mathrm{CO}_1 \cup \mathrm{CO}_2) \succ_i \mathrm{CO}_1] \wedge\\&[\forall i \in \mathrm{CO}_2, (\mathrm{CO}_1 \cup \mathrm{CO}_2) \succ_i \mathrm{CO}_2]\end{aligned} \tag{3}$$

2.2 Paillier homomorphic encryption

Homomorphic encryption technology can directly operate ciphertext without decryption. The Paillier homomorphic encryption technology[18, 19] is as follows.

Key generation: Randomly the user selects two large prime numbers p and q, and calculates $N = pq$ and the least common multiple $\lambda = \operatorname{lcm}(p-1, q-1)$. Then user randomly selects $g \in Z^*_{N^2}$, $Z^*_{N^2}$ indicating that the integer set has N^2 elements.

User defines function $L(x) = \dfrac{x-1}{N}$ and generates public key $K = (N, g)$ and the corresponding private key $k = (\lambda, \mu)$.

$$\mu = (L(g^{\lambda} \bmod N^2))^{-1} \bmod N \tag{4}$$

Encryption: User selects a random number r_i to encrypt the message m_i with K into ciphertext as follows:

$$C = E(m_i) = g^{m_i} {r_i}^N \bmod N^2 \tag{5}$$

Decryption: User uses k to decrypt the ciphertext C as follows:

$$m_i = D(C) = L(C^{\lambda} \bmod N^2) \times \mu \bmod N \tag{6}$$

Due to the additive homomorphism property of Paillier homomorphism encryption, for any message m_1 and m_2, user selects random number r_1 and r_2, and calculates the ciphertext product to get the aggregate message, as follows:

$$\begin{aligned} C_1 \cdot C_2 = E(m_1, r_1) \cdot E(m_2, r_2) = \\ g^{m_1+m_2}(r_1 \cdot r_2)^N \bmod N^2 \end{aligned} \tag{7}$$

Known $(r_1 \cdot r_2)^{\lambda N} \equiv 1(\bmod N^2)$ decrypts Eq. (7):

$$\begin{aligned} &D(C_1 \cdot C_2) = D(E(m_1, r_1) \cdot E(m_2, r_2) \bmod N^2) \equiv \\ &L(g^{\lambda(m_1+m_2)}(r_1 \cdot r_2)^{\lambda N} \bmod N^2) \cdot \mu \bmod N \equiv \\ &m_1 + m_2 \bmod N \end{aligned} \tag{8}$$

3 System Architecture and Various Stages

In this section, we first introduce the system architecture of our VDCM mechanism, then describe the various stages in our mechanism in two mechanism assumptions which include the initialization of the system, cooperative certification, data aggregation, data update, and information feedback.

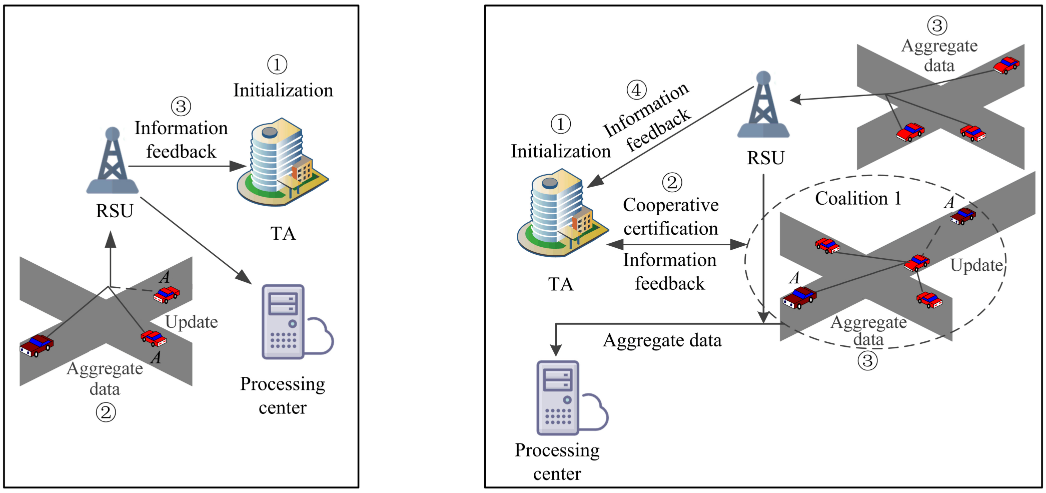
3.1 System structure

Aiming at the safety and efficiency of data collection for the crowd sensing in VANET, this paper proposes a data collection mechanism for crowd sensing in vehicular ad hoc networks. The system architecture of our VDCM mechanism is revealed in Fig. 1, where there exist four kinds of primary entities, i.e., trusted authority (TA), road side units (RSUs), processing center, and vehicles.

TA: The VDCM mechanism contains a trusted authority (TA). The TA contains a clock and divides the time into a series of equallength time intervals[20]. Furthermore, it stores and periodically updates vehicles reputation information based on the received reputation feedbacks and helps vehicles change coalition.

RSUs: In the VDCM mechanism, road side units (RSUs) help processing center collect vehicles-provided data. When RSU is authorized by TA, it can process data. Generally RSUs connect to vehicles and TA through the vehicle to roadside unit (V2R) wireless communication and wired communication, respectively[21].

Processing center: The VDCM mechanism which contains processing center is able to use various operations (e.g., K-means clustering, Lloyd's clustering, cloud computing, etc.) to process data. Generally,



(a) Sub mechanism 1

(b) Sub mechanism 2

Fig. 1 System architecture of the VDCM mechanism.

processing center refers to the processing equipment of the service provider layer in crowd sensing. The content of data processing is beyond the scope of this article, so it will not be introduced in detail.

Vehicles: In the VDCM mechanism, each vehicle is assumed to be equipped with sensing devices to collect and transmit sensing data. And each vehicle is assumed to generate random number which has a certain corresponding relationship with its sequence numbers. Besides, each vehicle periodically requests the TA for its new coalition and reputation certificates.

According to the specific situation, the VDCM can be divided into two mechanism assumptions, sub mechanism 1 and sub mechanism 2. When very few vehicles collect data or the coalition strategy cannot be adopted, for example, if the coalition cannot be formed or the coalition cooperation vehicle cannot be selected, etc., it should select sub mechanism 1. Sub mechanism 1 is revealed in Fig. 1a and includes three stages. In the first stage, TA initializes vehicle reputation score and coalition. In the second stage, the vehicle sends or updates the data to the RSUs, and the RSUs aggregate data which are sent to the processing center. In the third stage, the TA updates the vehicle reputation according to the score feedback of RSUs.

Except for the above special cases, sub mechanism 2 is selected. Sub mechanism 2 is revealed in Fig. 1b and includes four stages. In the first stage, TA initializes vehicle reputation score and coalition. In the second stage, the TA decides the cooperative vehicles in the coalition. In the third stage, the cooperative vehicle assists the RSUs to aggregate and update the data which are sent to the processing center. In the fourth stage, the TA updates the vehicle reputation and the coalition according to the score feedback of RSUs and cooperative vehicles.

3.2 Threat model

In this paper, it is assumed that the TA is completely trusted, RSUs are honest and curious, RSUs honestly perform data aggregation, but they are curious about the privacy information of the vehicle. The vehicles may be malicious and possibly pose threats as follows.

Disinformation attack: Malicious vehicles upload false data and pollute the dataset. The attack results in insufficient perceived data samples and makes the data unable to provide effective services.

Data tampering attack: Malicious vehicles tamper with the data uploaded by other vehicles. The attack affects the reputation score of the tampered vehicles and reduces the collection enthusiasm of the tampered vehicles.

Replay attack: Malicious vehicles upload duplicate data for many times. The attack affects the perception of data samples, and makes the data unable to provide effective services.

Eavesdropping attack: Malicious vehicles judge the information of the sending vehicles by receiving data, posing a threat to the participating vehicles.

DoS attack: Malicious vehicles refuse to provide services, which affects the efficiency of data collection.

3.3 Initialization

In order to avoid malicious transmission of data by vehicles and ensure the availability of vehicle-provided data, the TA generates time intervals $T_\alpha \in \{T_1, T_2, \dots\}$ and sets the initial reputation score $\mathrm{RM}_{V_i}^{0}$ for each vehicle V_i. TA stores the information in the database and updates the vehicle reputation score in time interval. The setting of the initial reputation score $\mathrm{RM}_{V_i}^{0}$ as[21]

$$\mathrm{RM}_{V_i}^{0}=\begin{cases}0.9, & \text{if } V_i \in \text{Law enforcenment vehicles;}\\ 0.5, & \text{if } V_i \in \text{Public service vehicles;}\\ 0.1, & \text{if } V_i \in \text{Other vehicles}\end{cases} \tag{9}$$

If the TA judges that the current situation should select sub mechanism 2, it will involve the allocation of the coalition. In order to facilitate the allocation of subsequent cooperative vehicles and preliminarily improve the efficiency of data collection, the TA will allocate the initial coalition. According to the time interval, the TA helps vehicles to update coalition, set coalition structure $P = \{\mathrm{CO}_1, \mathrm{CO}_2, \dots, \mathrm{CO}_k\}$, and participat vehicles $\upsilon = \{1, 2, \dots, N\}$. The allocation of the coalition shall comply with the coalition formation strategy[16] to ensure the transmission consumption is as little as possible and fits Eq. (10).

$$\begin{cases}\mathrm{CO}_a \cap \mathrm{CO}_b = \varphi,\\ (\mathrm{CO}_a \neq \mathrm{CO}_b \in \{\mathrm{CO}_1, \mathrm{CO}_2, \dots, \mathrm{CO}_k\});\\ \sum \mathrm{CO}_i = \upsilon\end{cases} \tag{10}$$

3.4 Cooperative certification

If the TA judges that sub mechanism 2 can be selected in the current situation, it will generate the initial coalition and start the cooperative certification. Otherwise, the RSUs will directly aggregate the data and skip the cooperative certification step. The execution steps of cooperative certification are as follows.

Firstly, the TA retrieves the vehicle reputation scores in each coalition and broadcasts the number of vehicles that need to be aggregated N_{a} and the corresponding reward Pa to each vehicle. Then the vehicle V_i judges whether to participate in the cooperative certification bidding. If the vehicle decides to participate in the bidding, a price bidding Pb_{V_i} is calculated according to the bidding cost. In order to prevent message flooding and bidding attacks, each vehicle is allowed to quote only once. V_i generates a request response Ab_{V_i} and sent it to the TA.

$$\text{Ab}_{V_i} = (N_{\text{a}}, \text{Pb}_{V_i}, \text{RM}_{V_i}) \tag{11}$$

where RM_{V_i} is the vehicle reputation. TA calculates the pre-cooperation degree according to Eq. (12) after receiving the bidding information.

$$\text{CP}_{V_i} = \begin{cases} 0, & \text{if} \quad \text{RM}_{V_i} < \text{RL}; \\ \dfrac{\text{RM}_{V_i}}{\text{Pb}_{V_i} \sum d_{V_i}^{j}}, & \text{otherwise} \end{cases} \tag{12}$$

where $\sum d_{V_i}^{j}$ is the sum of the distance from vehicle V_i to vehicle V_j in the same coalition, and the TA selects a certain number of vehicles as partners according to the pre cooperation degree from CP_{V_i} high to low.

3.5 Data aggregation

According to the two mechanisms, the subsequent operations are divided into the single aggregation algorithm corresponding to sub mechanism 1 and the multi aggregation algorithm corresponding to sub mechanism 2.

3.5.1 Single aggregation algorithm

The TA randomly sorts the participating vehicles and generates a set of sequence numbers $k_1, k_2, \ldots, k_n$. Then the TA safely transmits the sequence numbers to each vehicle $V_1, V_2, \ldots, V_n$, and records the corresponding relationship between the sequence number and the vehicle. The TA will generate public key $K = (N, g)$ and private key $k = (\lambda, \mu)$, then the TA broadcasts the public key K to vehicles. V_i selects random number r_i, encrypts data message M_i to generate ciphertext C_i through Paillier homomorphic encryption technology, then sends ciphertext and sequence number to the RSU.

$$C_i = g^{M_i} r_i^N \bmod N^2 \tag{13}$$

RSU receives C_i from V_i and multiplies it according to Eq. (14) to obtain the aggregated encrypted data C_0.

$$C_0 = \prod_{i=1}^{n} C_i = \prod_{i=1}^{n} E(M_i, r_i) = g^{\sum\limits_{i=1}^{n} M_i} \prod_{i=1}^{n} r_i \bmod N^2 \tag{14}$$

After that, the processing center can retrieve the corresponding private key from the TA and obtain aggregate data after decryption, as

$$D(\prod_{i=1}^{n} C_i) \equiv L(\prod_{i=1}^{n} C_i^{\lambda} \bmod N^2) \times \mu \bmod N \equiv L(g^{\lambda \sum\limits_{i=1}^{n} M_i} \bmod N^2) \times \mu \bmod N \equiv \sum_{i=1}^{n} M_i \bmod N \tag{15}$$

In addition, if the TA chooses to authorize the RSU, then the RSU can decrypt, process, and analyze the aggregate ciphertext through the private key.

3.5.2 Multi aggregation algorithm

After passing the cooperative certification, the vehicle will become a cooperative vehicle to assist RSU in collecting aggregated data. The TA generates the public key $K = (N, g)$ and private key $k = (\lambda, \mu)$. Then the TA broadcasts the public key to vehicles.

The cooperative vehicle V_{c_1} collects data from vehicles $V_1, V_2, \ldots, V_q (q < n)$. Then the TA will randomly sort the participating vehicles and generate a set of sequence numbers $k_1, k_2, \ldots, k_q$ which will be safely transmitted to collected vehicles $V_1, V_2, \ldots, V_q$. The TA records the corresponding relationship between the sequence number and the vehicle.

V_i selects random numbers r_i in coalition A, then sends the ciphertext and the sequence number to V_{c_1}. The encryption process is similar to that in single aggregation, and finally ciphertext product $C_{V_{c_1}}$ was obtained by aggregation.

$$C_{V_{c_1}} = \prod_{i=1}^{q} C_i = \prod_{i=1}^{q} E(M_i, r_i) = g^{\sum\limits_{i=1}^{q} M_i} \prod_{i=1}^{q} r_i \bmod N^2 \tag{16}$$

The processing center can retrieve the corresponding private key from the TA. The final aggregate ciphertext C is obtained from the aggregated ciphertext product $\prod\limits_{i=1}^{w} C_{V_{c_i}}$ of the w cooperative vehicles and the RSUs aggregate ciphertext product C_0 according to Eq. (17). After processing center decrypts C by the private key to obtain final aggregate data, as shown in Eq. (18).

$$C = C_0 \prod_{i=1}^{w} C_{V_{c_i}} \tag{17}$$

$$D(C) \equiv L(\prod_{i=1}^{n} C^{\lambda} \bmod N^2) \times \mu \bmod N \equiv \sum_{i=1}^{n} M_i \bmod N \tag{18}$$

If the TA chooses authorization, the RSUs or cooperative vehicle can decrypt, process, and analyze the aggregated ciphertext through the private key.

3.6 Data update

V_i can choose to update the uploaded data within a certain period of time after the cooperative vehicle or RSU receives the data from the vehicle V_i. When V_i needs to change the uploaded data, it performs the following operations.

Data M_i' will be added, V_i uses the same random number r_i to obtain the updated ciphertext according to Eq. (19). Then V_i sends the updated ciphertext and its own sequence number to the RSU or cooperative vehicle.

$$C_i' = g^{M'_i} {r_i}^N \bmod N^2 \tag{19}$$

The RSU or cooperative vehicle will multiply the ciphertext to be added by the original ciphertext of the vehicle and obtain the updated ciphertext, as follows:

$$C_i'' = g^{M_i + M'_i} r_i^{2N} \bmod N^2 \tag{20}$$

$$\prod_{i=1}^{n} {r_i}^{\lambda N} \equiv 1 \bmod N^2 \tag{21}$$

When the updated ciphertext is multiplied by other ciphertexts, it can be aggregated into a new aggregated ciphertext. The data can be updated without decryption. The possibility of RSU or cooperative vehicle receiving vehicle data update request is affected by the willingness of vehicle data update. Our mechanism analyzes the two aspects of distance and benefit. The farther the vehicle distance is, the more time it needs to consume. Since the update is only limited to a certain period of time, with the increasing distance, the willingness of vehicles will decrease accordingly. In addition, the update intention is affected by the update effects and incomes, and the effects and incomes of our mechanism are reflected in the score. When the vehicle data are updated, the score will increase appropriately to affect the final reputation. However, the service continuity score is affected by the marginal utility[21], and the benefits will be reduced when the vehicle has updated multiple times. The b-th update willingness degree B_b^i is calculated as

$$B_b^i = \begin{cases} 0, \ \text{if } d_b^{i,o} + d_0 > t_b^0 v_i || A_b^i - A_{b-1}^i < \wp_i; \\ \delta A_b^i (d_b^{i,o} + d_0)^{-1}, \ \text{otherwise} \end{cases} \tag{22}$$

where A_b^i is b times that the vehicle is rewarded. $d_b^{i,o}$ is the distance from the vehicle to the update midpoint (RSU or cooperative vehicle) at the b times. d_0 is additional consumption. $\wp_i$ is vehicle benefit threshold and is determined by V_i. δ is the conversion scale factor.

3.7 Information feedback

After the collection service, both RSU and cooperative vehicles score the collected vehicles which they come into contact with. The scoring directions include service continuity, data reliability, selfishness, and cooperation. After receiving the score, the TA should avoid malicious multiple updates and increase the workload of the intermediate aggregation point. It needs to calculate the service continuity score twice to get a new score.

$$\text{SC}' = \log_a(\text{SC} + 1) \tag{23}$$

where $(a > 1)$ is the marginal effect attenuation factor $\text{SC} \in [0, 100]$. Then, the TA calculates the final score formula as

$$\text{TS} = \text{SC}' \times \alpha + \text{DRL} \times \beta + \text{SF} \times \gamma + \text{CO} \times \theta \tag{24}$$

where weight of testimony is $\alpha + \beta + \gamma + \theta = 1$, scope of testimony is $\text{SC}', \text{DRL}, \text{SF}, \text{CO} \in [0, 100]$. TA will judge and generate vehicle feedback according to the final score. If $\text{TS} > \text{SL}$, the feedback will be recorded as $\alpha_{V_i}^1 = 1$. Otherwise, the feedback will be 0. The trusted data threshold SL is determined by the TA.

At the time interval of T_α, the TA updates the reputation scores RM_{V_i} of the current vehicle through Eq. (25) according to the recorded vehicle reputation feedback.

$$\text{RM}'_{V_i} = \begin{cases} \dfrac{\sum\limits_{j=n} (\alpha_{V_i}^j \times \omega_{V_i}^j)}{n} \text{RM}_{V_i}, \ \text{if } j = 0; \\ \gamma \times \text{RM}_{V_i}, \qquad \text{otherwise} \end{cases} \tag{25}$$

where the set of reputation feedback received by the vehicle in the current time interval is $\alpha_{V_i} = \{\alpha_{V_i}^1, \alpha_{V_i}^2, \ldots, \alpha_{V_i}^n\}$, the set of weights for each feedback time is $\omega_{V_i} = \{\omega_{V_i}^1, \omega_{V_i}^2, \ldots, \omega_{V_i}^n\}$ $(\omega_{V_i}^1 \leqslant \omega_{V_i}^n)$, and γ is the time attenuation factor.

At the time interval of T_α, TA assists vehicles to form a new coalition. The coalition formation follows the coalition formation strategy and uses the best response coalition formation algorithm[16].

Let $P = \{\text{CO}_1, \text{CO}_2, \ldots, \text{CO}_k\}$ be the coalition structure, and k is the maximum number of coalitions. Before the change, the utility function of all coalitions is $\bar{u}_n(P) = \{E_n(\text{CO}_1), E_n(\text{CO}_2), \ldots, E_n(\text{CO}_k)\}$. If the vehicle will join other coalitions, the TA analyzes the pre exchange status information and coalition information of the vehicle, finds the best utility, and assists the vehicle to join the corresponding coalition. The specific operations are as follows.

The two coalitions involved are recorded as CO_1 and CO_2, and V is the vehicle. TA calculates CO_1, CO_2, $\text{CO}_1 \cup \{V\}$, and $\text{CO}_2 \cup \{V\}$ utility and judges whether Eq. (26) is satisfied.

$$\begin{aligned}&\text{CO}_1 \succ_n \text{CO}_2 \Leftrightarrow E_n(\text{CO}_1) < E_n(\text{CO}_2) \wedge \\ &E_i(\text{CO}_1) \leqslant E_i(\text{CO}_1 \backslash \{V\}) \forall i \in \text{CO}_1 \backslash \{V\} \wedge \\ &E_i(\text{CO}_2 \backslash \{V\}) \leqslant E_i(\text{CO}_2) \forall i \in \text{CO}_2 \backslash \{V\}\end{aligned} \tag{26}$$

If Eq. (26) is satisfied, V can be switched from the coalition CO_2 to the coalition CO_1, otherwise it remains unchanged. The TA uses the above operations to allocate all vehicles to form a new coalition. In particular, if the cooperative vehicle needs to change the coalition, it is necessary to send the data collected in the current time period to the next cooperative vehicle or processing center before switching the coalition.

4 Performance Analysis and Simulation Evaluation

4.1 Performance analysis

This paper analyzes security of the VDCM mechanism. The collection mechanism can resist common attacks, such as disinformation attack, data tampering attack, replay attack, eavesdropping attack, and DoS attack. It is certified as follows:

Disinformation attack: When the vehicle collects data, the vehicle V_i may maliciously send false data and launch disinformation attacks to affect the collection efficiency. In this mechanism, RSUs and cooperative vehicles will score the vehicle behavior in many aspects. The probability that malicious data score meets threshold and is collected as useful data is almost negligible. In addition, the TA will update the vehicle reputation at time intervals, and the vehicle sending false data will affect its reputation, so this mechanism can resist disinformation attack.

Data tampering attack: When the vehicle collects data, the attacker may launch a data tampering attack to tamper with the vehicle data. In this mechanism, the vehicle selects a random number r_i and uses public key k to generate ciphertext C_i according to homomorphic secret technology. The attacker cannot get the private key k and the vehicle random number r_i decrypts the ciphertext, so the attacker cannot get the vehicle plaintext data and tamper with the plaintext data, so the mechanism can resist data tampering attack.

Replay attack: When the vehicle updates data, the attacker may launch replay attack, repeatedly upload data, and attempt to improve its reputation and increase the workload of intermediate aggregation points by improving the service continuity score. In this mechanism, data update is affected by marginal utility, and the attacker cannot improve its reputation through multiple updates. Therefore, this mechanism can resist replay attack.

Eavesdropping attack: When the vehicle collects data, the attacker may launch an eavesdropping attack, trying to judge the information of the vehicle sent through the eavesdropping data, and pose a threat to the vehicle. In this mechanism, the attacker cannot obtain the private key k and the vehicle random number r_i decrypts the ciphertext, so the attacker cannot eavesdrop on the data. In addition, the TA will update the alliance at time intervals to improve the unlinkability between vehicles and data to ensure the security of data and vehicles. Therefore, this mechanism can resist eavesdropping attack.

DoS attack: When the vehicle collects data, the vehicle V_i may refuse to provide data, affecting the efficiency of data collection. In this mechanism, RSUs and cooperative vehicles will score vehicle behavior in many aspects, and the TA will update reputation at time intervals. If the vehicle refuses to provide service, its reputation will be affected, so this mechanism can resist DoS attack.

In conclusion, this mechanism can safely aggregate data.

To prove the advantages of this mechanism, this paper analyzes performance of the mechanism. The performance of this mechanism is compared with the homomorphic encryption self organization mechanism (HESO)[14], the privacy protection mechanism based on fog buses (FBVCS)[8], and the cloud oriented homomorphic encryption method of Qian et al.[9] According to Table 1, this paper compares the performance from the following: data security and integrity, data update ability, credibility of intermediate vehicles, and aggregation consumption.

Data security and integrity: The aggregation method used in this paper is based on Paillier homomorphic encryption technology which can aggregate the ciphertext directly to ensure data security. At the time intervals, the TA will update the coalition to improve the unlinkability of vehicles and data and further ensure the safety of data and vehicles. In addition,

Table 1 Performance comparison.

Method	Data security	Data integrity	Data update ability	Credibility of intermediate vehicles	Aggregation consumption (ms)
VDCM	✓	✓	✓	↑	29.82
HESO[14]	✓	✓	×	×	>59.64
FBVCS[8]	✓	✓	×	↓	33.51
Qian et al.[9]	×	×	✓	×	×

Note: ✓: The mechanism involves the performance comparison. ×: The mechanism does not involve the performance comparison. ↓↑: The mechanism is lower/higher than the other mechanisms.

the vehicle can update the encrypted data within a certain period of time, and RSUs or cooperative vehicles will score the vehicle behavior and feed back the reputation results in time. The TA updates the reputation according to the reputation feedback and filters malicious vehicles to improve the data integrity. Qian et al.[9] proposed homomorphic encryption technology based on ECC. However, the security of the reference basic mechanism has not been officially proved and recognized, thus the security needs to be verified. Therefore, the mechanism fails to guarantee the data security and integrity.

Data update ability: When devices update data, they are usually necessary to decrypt the ciphertext and then carry out corresponding processing operations. This paper directly processes the ciphertext to avoid unnecessary problems such as information disclosure. Qian et al. realize data update, but HESO and FBVCS fail to consider data update.

Credibility of intermediate vehicles: The intermediate vehicles of our mechanism are selected by the TA according to the auction cooperation agreement and pre cooperation degree. Credibility of intermediate vehicles is higher than FBVCS.

Aggregation consumption: In order to further reflect the performance advantages of our mechanism, the consumption of VDCM, FBVCS, and HESO in the process of data aggregation is compared. The measurement is from Ref. [8], running on the host configured with 1.8 GHz Intel processor, 8 GB memory, and windows 10 operating system, and obtains the following simulation time, homomorphic encryption time $T_{\text{ec}} = 29.82$ ms and mapping generation time $T_h = 3.69$ ms. In our mechanism, VDCM data aggregation only involves homomorphic encryption operation total consumption T_{ec}. In addition to homomorphic encryption, FBVCS mechanism also needs to consume additional mapping generation time T_h, with a total consumption of $T_{\text{ec}} + T_h$. HESO mechanism needs a times homomorphic encryption operation, and the total consumption is $(a+1)T_{\text{ec}} (a \geqslant 1)$. In conclusion, our mechanism has a lighter data aggregation process, reduces unnecessary consumption, and increases the willingness of vehicle collection. Compared with FBVCS and HESO mechanisms, our mechanism takes less time to aggregate the same amount of data.

The following is a comparative analysis of our mechanism. According to the specific situation, our mechanism VDCM is divided into two mechanism assumptions. Total consumption of the two mechanism assumptions is different under different vehicle numbers. Since sub mechanism 2 needs to increase cooperative vehicles additional consumption d_0 to reduce the consumption of other aspects (e.g., transmission and processing). It is assumed that both cases grow linearly. When the additional consumption is lower than the consumption of other aspects reduction in sub mechanism 2, sub mechanism 2 will be better than sub mechanism 1. Assuming that the additional consumption is consistent with the growth ratio and the optimization ratio is 50%, parameter $w \in \{1, 2, 3\}$. According to Fig. 2, under different growth ranges, when the number of vehicles is 200, the additional consumption is equal to the consumption of other aspects reduction in sub mechanism 2. As shown in Table 2, when the number of vehicles is 100, the consumption of sub mechanism 1 is lower than sub mechanism 2. When the number of vehicles is 600 and 800, the consumption of sub mechanism 2 is lower than sub mechanism 1. With the increase of the number of vehicles, the advantage of sub mechanism 2 is greater. The TA selects

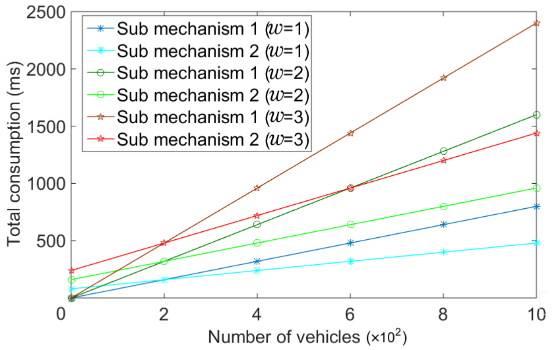


Fig. 2 Consumption of different numbers of vehicles under two sub mechanisms.

Table 2 Consumption of two sub mechanisms under the number of vehicles is 100, 600, and 800.

Number of vehicles	Consumption (ms)								
	$w=1$			$w=2$			$w=3$		
	Sub mechanism 1	Sub mechanism 2	D-value of two sub mechanisms	Sub mechanism 1	Sub mechanism 2	D-value of two sub mechanisms	Sub mechanism 1	Sub mechanism 2	D-value of two sub mechanisms
100	80	120	40	160	240	80	240	360	120
600	480	320	–160	960	640	–320	1440	960	–480
800	640	400	–240	1280	800	–480	1920	1200	–720

the corresponding assumptions according to the actual situation and adjusts the mechanism steps to reduce time consumption.

4.2 Simulation evaluation

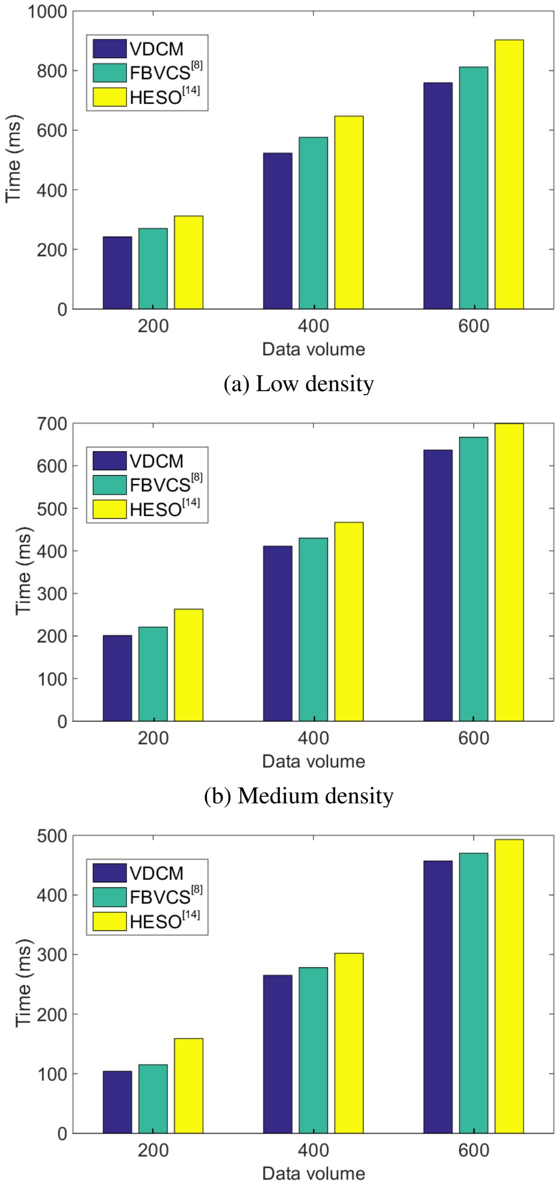
Based on the mobile traffic model simulation software, the simulation data collection process is carried out under specific circumstances. When the vehicle enters the communication range, vehicle participation is determined. The map adopts double intersections, involving a length of about 2.1 km. According to the number of data taken, this paper does the simulation update and scoring mechanism. The simulation sets the data provided by vehicles whose reputation is higher than the reputation threshold as available data. In order to facilitate comparison with various mechanisms, the specific simulation parameter values are set in Table 3. The set in Table 3 is from Refs. [8, 22].

FBVCS and HESO introduce fog bus and route planning strategies, respectively, and additional consumption is increased. In order to optimize the comparison, in the following simulation comparison, our mechanism introduces cooperative vehicle assisted aggregation. The strategies of our VDCM, FBVCS, and HESO are adjusted to the best. The following simulation compares the time when VDCM, FBVCS, and HESO mechanisms collect the same amount of data. The simulation records the time required for VDCM, FBVCS, and HESO to collect the same amount of data at different densities when the amount of data is 200, 400, and 600. As shown in Fig. 3, when the vehicle density is medium and high, the time required by our mechanism is lower than the above mechanism. When the amount of data collected is 200, the gap is most obvious. In addition, according to Table 4, at low density, the time gap becomes more obvious with the increase of the amount of data collected. Through the above analysis, we can easily find that the VDCM mechanism takes less time to collect the same amount of data when compared with FBVCS and HESO mechanisms. So the VDCM mechanism reduces time consumption.

Table 3 Simulation parameter value setting.

Parameter	Value
Total data volume	340−700
Vehicle communication range	200 m
Maximum vehicle speed	22 m/s
Bus communication range	400 m
Maximum bus speed	15 m/s
Map width	400 m
Map length	1300 m
Map size	2.1 km^2
Number of low density vehicles	60
Number of medium density vehicles	130
Number of high density vehicles	210
Update data volume	200, 400, 600
Reputation threshold	0.5
Distance classification	3
Distance threshold	2
Random update probability	0.5
Update threshold	0.6
Distance classification probability	1/3
Attenuation degree	50%

Not only is the reputation feedback obtained by the subsequent processing of this mechanism not limited to the data, but it involves the scores of all aspects. According to the above parameters, this paper simulates the reputation update after the data collection operation in time interval, and this paper generates the reputation scores according to Ref. [22]. The reputation scores are generated randomly and conform to the normal distribution. According to Fig. 4, the amount of data available for this mechanism is higher than FBVCS and HESO mechanisms. Table 5 indicates the proportion of available data of each mechanism under different data volumes. The VDCM has a higher proportion of available data than FBVCS and HESO. The advantage is more obvious when the total number of data is 200. When the data volume is 200, compared with the other two mechanisms, the proportion of available data of the VDCM mechanism is higher than that of the FBVCS and HESO mechanisms by 19.5% and 22.5%, respectively. So the VDCM mechanism increases the amount of available data compared with existing mechanisms.



(a) Low density

(b) Medium density

(c) High density

Fig. 3 Time consumption of collecting different amounts of data at each density.

Table 4 Time consumption of collecting different amounts of data at low density.

Amount of data	Time consumption (ms)		
	VDCM	FBVCS[8]	HESO[14]
200	242(0)	270(+32)	312(+70)
400	523(0)	576(+53)	647(+124)
600	759(0)	815(+56)	903(+144)

5 Conclusion

In this article, we have proposed a data collection mechanism named VDCM for the crowd sensing in vehicular ad hoc networks. Firstly, the VDCM adopts the Paillier homomorphic encryption technology to ensure the security of data aggregation. In addition, the VDCM leverages the coalition formation strategy and auction cooperation protocol to aggregate vehicle-provided data. Moreover, it supplements the data update aspect not mentioned in most data collection mechanisms, and the update probability is analyzed from two aspects of distance and benefit. Finally, the VDCM updates reputation scores according to the reputation feedback and updates the coalition. Performance analysis shows that part of the performance of the proposed mechanism is better than existing data collection mechanisms and the VDCM can safely aggregate data. We have not only evaluated the performance of the VDCM mechanism but also compared it with FBVCS and HESO mechanisms. The simulation results show that compared with the FBVCS and HESO, our VDCM can reduce time consumption and get a higher amount of available data for vehicles in the process of data collection.

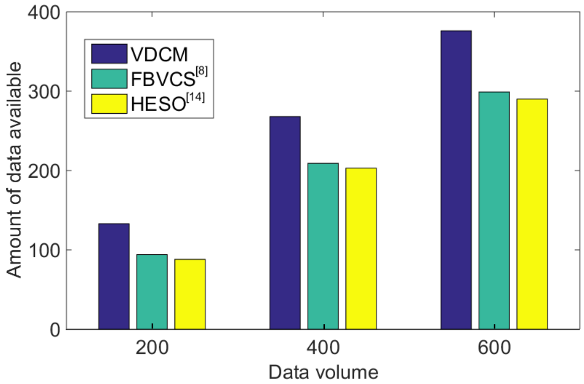


Fig. 4 Amount of data available for each mechanism under different data volumes.

Table 5 Proportion of available data of each mechanism under different data volumes.

Data volume	Proportion of available data (%)		
	VDCM	FBVCS[8]	HESO[14]
200	66.50	47.00	44.00
400	67.00	52.25	50.75
600	62.67	49.83	48.33

In future work, we will further evaluate and improve the performance of our VDCM mechanism. Furthermore, we will introduce the famous ECC into our mechanism so as to reduce encryption consumption.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Nos. 62272195 and 61802146); the Guangdong Province Science and Technology Planning Project (No. KTP20200022); the

Guangdong Basic and Applied Basic Research Foundation (No. 2019A1515011017); the Science and Technology Program of Guangzhou of China (No. 202201010421); the Fundamental Research Funds for the Central Universities (Nos. 21621417 and 21622402); and the Guangdong Provincial Key Laboratory of Cyber and Information Security Vulnerability Research (No. 2020B1212060081).

References

[1] A. K. Sandhu, Big data with cloud computing: Discussions and challenges, *Big Data Mining and Analytics*, vol. 5, no. 1, pp. 32–40, 2022.

[2] R. K. Ganti, F. Ye, and H. Lei, Mobile crowdsensing: Current state and future challenges, *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 32–39, 2011.

[3] N. Banerjee, T. Giannetsos, E. Panaousis, and C. C. Took, Unsupervised learning for trustworthy IoT, in *Proc. 2018 IEEE International Conference on Fuzzy Systems* (*FUZZ-IEEE*), Rio de Janeiro, Brazil, 2018, pp. 1–8.

[4] G. Xu, H. Li, S. Liu, M. Wen, and R. Lu, Efficient and privacy-preserving truth discovery in mobile crowd sensing systems, *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 3854–3865, 2019.

[5] X. Liu, K. Ota, A. Liu, and Z. Chen, An incentive game based evolutionary model for crowd sensing networks, *Peer Peer Netw. Appl.*, vol. 9, no. 4, pp. 692–711, 2016.

[6] S. Basudan, X. Lin, and K. Sankaranarayanan, A privacy-preserving vehicular crowdsensing-based road surface condition monitoring system using fog computing, *IEEE Internet Things J.*, vol. 4, no. 3, pp. 772–782, 2017.

[7] M. Sookhak, A. Gani, M. K. Khan, and R. Buyya, Dynamic remote data auditing for securing big data storage in cloud computing, *Inf. Sci.*, vol. 380, pp. 101–116, 2017.

[8] G. Sun, S. Sun, J. Sun, H. Yu, X. Du, and M. Guizani, Security and privacy preservation in fog-based crowd sensing on the internet of vehicles, *J. Netw. Comput. Appl.*, vol. 134, no. 5, pp. 89–99, 2019.

[9] P. Qian, M. Wu, and Z. Liu, A method on homomorphic encryption privacy-preserving for cloud computing, (in Chinese), *Journal of Chinese Computer Systems*, vol. 36, no. 4, pp. 840–844, 2015.

[10] T. Kuo, K. C. Lin, and M. Tsai, On the construction of data aggregation tree with minimum energy cost in wireless sensor networks: NP-completeness and approximation algorithms, *IEEE T. Comput.*, vol. 65, no. 10, pp. 3109–3121, 2016.

[11] P. G. Naranjo, M. Shojafar, H. Mostafaei, Z. Pooranian, and E. Baccarelli, P-SEP: A prolong stable election routing algorithm for energy-limited heterogeneous fog-supported wireless sensor networks, *J. Supercomput.*, vol. 73, no. 2, pp. 733–755, 2017.

[12] N. Javaid, M. Jafri, Z. Khan, N. Alrajeh, M. Imran, and A. Vasilakos, Chain-based communication in cylindrical underwater wireless sensor networks, *Sensors*, vol. 15, no. 2, pp. 3625–3649, 2015.

[13] C. Liu and G. Cao, Distributed monitoring and aggregation in wireless sensor networks, in *Proc. 29th Conference on Information Communications*, San Diego, CA, USA, 2010, pp. 2097–2105.

[14] K. Rabieh, M. Mahmoud, and M. Younis, Privacy-preserving route reporting schemes for traffic management systems, *IEEE Trans. Veh. Technol.*, vol. 66, no. 3, pp. 2703–2713, 2017.

[15] P. Zhao, Y. Fu, and C. Li, Privacy and security-oriented crowdsensing system framework for vehicular networks, (in Chinese), *Mobile Communications*, vol. 45, no. 6, pp. 37–42, 2021.

[16] Y. Zhang, Y. Xu, Q. Wu, Y. Luo, Y. Xu, X. Chen, A. Anpalagan, and D. Zhang, Context awareness group buying in D2D networks: A coalition formation game-theoretic approach, *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12259–12272, 2018

[17] F. Xiong, H. Zheng, L. Ruan, H. Wang, L. Tang, X. Dong, and A. Li, Energy-saving data aggregation for multi-UAV system, *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 9002–9016, 2020.

[18] M. Huang, Research on improved homomorphic encryption scheme based on integer, (in Chinese), *China Science & Technology Overview*, no. 16, pp. 22–23, 2021.

[19] Y. Zhang and J. Ling, Improved algorithm for privacy-preserving association rules mining on horizontally distributed databases, (in Chinese), *Computer Science*, vol. 44, no. 8, pp. 157–161, 2017.

[20] Z. Liu, F. Huang, J. Weng, K. Cao, Y. Miao, J. Guo, and Y. Wu, BTMPP: Balancing trust management and privacy preservation for emergency message dissemination in vehicular networks, *IEEE Internet Things*, vol. 8, no. 7, pp. 5386–5407, 2021.

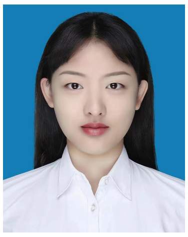
[21] Z. Liu, J. Weng, J. Ma, J. Guo, B. Feng, Z. Jiang, and K. Wei, TCEMD: A trust cascading-based emergency message dissemination model in VANETs, *IEEE Internet Things*, vol. 7, no. 5, pp. 4028–4048, 2020.

[22] Z. Liu, J. Guo, F. Huang, D. Cai, Y. Wu, X. Chen, and K. K. Igorevich, Lightweight trustworthy message exchange in unmanned aerial vehicle networks, *IEEE T. Intell. Transp.*, vol. 24, no. 2, pp. 2144–2157, 2023.

Linfeng Wei received the MEng degree from Xidian University, Xi'an, China in 2006, and the PhD degree from Jinan University, Guangzhou, China in 2020. He is currently a senior engineer with the College of Cyber Security, Jinan University, China. His research focuses on cloud computing security, block chain security, and mobile security.

Juli Yin received the BEng degree from South China Agricultural University, Guangzhou, China in 2019. She is currently pursing the MS degree with the College of Cyber Security, Jinan University, Guangzhou, China. Her current research interests include privacy preservation for crowd sensing in vehicular networks.

Zhiquan Liu received the BS and PhD degrees from Xidian University, Xi'an, China in 2012 and 2017, respectively. He is currently an associate professor with the College of Cyber Security, Jinan University, Guangzhou, China. His current research focuses on trust modeling and privacy protection in internet of vehicles.



Xi Yang received the BEng degree from Yunnan University, Yunnan, China in 2019. She is currently pursuing the MS degree with the College of Information Science Technology, Jinan University, Guangzhou, China. Her current research interests are the application of machine learning to forecasting problems.

Hongliang Sun received the BEng degree from East China Jiaotong University, Jiangxi, China in 2019. He is currently pursuing the MS degree with the College of Cyber Security, Jinan University, Guangzhou, China. His current research interest includes privacy preservation and trust management in vehicular networks.

Yudan Cheng received the BS and MS degrees from Northwest Normal University, Lanzhou, China in 2016 and 2019, respectively. She is currently pursuing the PhD degree with the College of Cyber Security, Jinan University, Guangzhou, China. Her current research interests include trust management and privacy preservation in vehicular networks.

Jianbin Mai received the BS degree from Northeast Agricultural University, Harbin, China in 2018 and MS degree from Hainan University, Haikou, China in 2021. He is currently pursuing the PhD degree with the College of Cyber Security, Jinan University, Guangzhou, China. His current research interests include security detection and privacy preservation in vehicular networks.