Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 6244525, 20 pages
https://doi.org/10.1155/2021/6244525

Research Article

DC-LTM: A Data Collection Strategy Based on Layered Trust Mechanism for IoT

An He,[1] Guangwei Wu,[1] and Jinhuan Zhang[2]

[1]*School of Computer and Information Engineering, Central South University of Forestry and Technology, Changsha 410004, China*
[2]*School of Computer Science and Engineering, Central South University, Changsha 410083, China*

Correspondence should be addressed to Guangwei Wu; guangweiwu@csuft.edu.cn

Received 31 May 2021; Accepted 30 July 2021; Published 18 August 2021

Academic Editor: Zhaolong Ning

A large number of Internet of Things (IoT) devices such as sensor nodes are deployed in various urban infrastructures to monitor surrounding information. However, it is still a challenging issue to collect data in a low-cost, high-quality, and reliable manner through IoT technique. Although the recruitment of mobile vehicles (MVs) to collect urban data has proved to be an effective method, most existing data collection systems lack a trust detection mechanism for malicious terminal nodes and malicious vehicles, which should lead to security vulnerabilities in practice. This paper proposes a novel data collection strategy based on a layered trust mechanism (DC-LTM). The strategy recruits MVs as data collectors of the sensor nodes based on the data value in the city, evaluates the trustworthiness of the data reported by the nodes, and records the results to the cloud data center. Furthermore, in order to make the data collection system more efficient and trust mechanism more reliable, we introduce unmanned aerial vehicles (UAVs) dispatched by data centers to actively verify the core sensor node data and use the core sensor data as baseline data to evaluate the credibility of the vehicles and the trust value of the whole network sensor nodes. Different from the previous strategies, UAVs adopts the DC-LTM method to obtain the node data while actively obtaining the trust value of MVs and nodes, which effectively improves the quality of data acquisition. Simulation results show that the mechanism effectively distinguishes malicious vehicles that provide false data in exchange for payment and reduces the total cost of system recruitment payments. At the same time, the proposed incentive mechanism encourages vehicle to complete the evaluation task and improves the accuracy of node trust evaluation. The recognition rates of false data attacks and flooding attacks as well as the recognition error rate of normal nodes are 100%, 98.9%, and 3.9%, respectively, which improves the quality of system data collection as a whole.

1. Introduction

With the development of IoT technology, new network systems such as the Internet of Vehicles, Smart Medical, and Smart City have also developed rapidly [1–3]. A large number of sensing devices perceive their surroundings and generate a large amount of data in the intelligent network system [4, 5]. These data often involve sensitive data, which brings higher demand on the security. Edge computing introduces an emerging computing model that helps to protect the confidentiality of sensitive data. As edge computing is based on the acquisition of massive data, the quality of data collection plays a decisive role in the application. Therefore, collecting data in a low-cost, high-quality, and reliable manner is a challenge issue for edge computing.

There are many researches based on mobile sink data collection strategies in data collection, which mainly use mobile vehicles (MVs) as data collection tool [6, 7], aggregating data while the vehicle is in motion and transmitting the collected data to data centers through 4G or 5G networks. These strategies are based on the assumption that terminal node data and mobile data collectors are completely reliable [8, 9]. Actually, a malicious terminal node may be created due to network failure or intrusion, or the terminal node is normal but the data collector may report false data to cheat for payment [10]. Both cases mean that the

malicious terminal node data provided by normal vehicles or the normal terminal node data provided by malicious vehicles are unreliable. Systems that lack trust detection mechanisms for malicious terminal nodes and malicious vehicles are severely compromised.

The trust-based data collection mechanism is an effective strategy, which is widely used in the P2P network [11], WSN network [12], etc. The trust mechanism evaluates the trust values of data collectors by observing their interaction behaviors and belongs to passive trust mechanisms. If data reported by the data collector conforms to the predetermined behavior, the credibility of the data collector will be improved; otherwise, the credibility will be reduced. Trust value is an important criterion when selecting data collectors, thus discouraging low-trust data collectors from participating in data collection. In the edge network, the number of IoT devices is huge, and it is difficult to obtain data interaction behavior due to privacy and monitoring difficulties. Such passive trust acquisition method is not suitable for edge computing because of the difficulty of acquiring trust, inaccuracy of evaluation, and nonverifiability of evaluation.

This paper designs a layered trust mechanism, which recruits MVs to provide trust evaluation for terminal nodes, and uses UAVs to actively verify the credibility of data collectors and to inhibit malicious data collectors from participating in data collection. A novel data collection strategy based on layered trust mechanism (DC-LTM) is proposed to obtain secure and low-cost data for IoTs in this paper. Different from the previous strategy, UAVs are adapted to actively obtain the trust value of MVs and nodes while acquiring node data in the DC-LTM method, which are effective in improving the quality of data acquisition. The main innovations of this paper are as follows.

(i) A layered trust mechanism based on terminal nodes, MVs, and UAVs is proposed. The terminal nodes monitor the communication behavior of their neighbors and report it to the MVs. The MVs obtain the status of the terminal nodes directly and perform trust evaluation. UAVs perform trust evaluation of the core terminal nodes, the results of which can be used to verify the reliability of the vehicles that reporting these core terminal nodes. Three kinds of trust relationships are used in the trust evaluation scheme proposed in this paper, namely, recommendation trust, direct trust, and active trust, the weighting of which combines into the comprehensive trust. Among them, active trust is the basis of verifiability, which allows active verification of node data through UAV, and is considered the most reliable. Active trust is given the highest weighting. Both recommendation trust and direct trust unfold from the communication behavior and energy state between nodes, respectively, and are largely affected by the high weight of active trust. The usage of UAV is not only accurate for trust verification, but also expands the scope of trust assessment through recommendation trust and direct trust and makes trust verification faster

(ii) A trust evaluation mechanism of UAV participating in active evaluation is proposed. The mechanism uses UAVs as an authoritative verification tool to verify actively the authenticity of data reported by mobile vehicles and inhibit malicious vehicles from participating in data collection and getting rewards. At the same time, UAVs are adopted to acquire data while abtaining node trust in the DC-LTM method which are effective in improving the quality of data acquisition. The system assigns the task of data collection to the mobile vehicles, and the accuracy of the collected data depends on the behavior and performance of the mobile vehicles themselves. In practice, the behavior and performance of mobile vehicles are not always credible. It is difficult to distinguish the reported data from real data or false data. For example, some MVs may commit malicious fraud by providing false data in order to obtain more rewards, and some MVs may accidentally report incorrect data due to their own errors. All these situations can lead to the failure of system monitoring. Therefore, the UAV is dispatched at the right time to compare the data sensed by UAV with the data reported by vehicle and to judge the reliability of the vehicle, which changes the shortcomings of previous passive trust acquisition strategy

(iii) The proposed recruitment mechanism encourages MVs to upload node information accurately. In the DC-LTM method, the MVs are employed to collect data with an incentive mechanism. The mechanism considers the value of the collected sensor data and avoids the situation that no vehicle is willing to collect nodes far away from the data center. It is a reasonable incentive mechanism

(iv) Simulation results show that the proposed layer trust evaluation mechanism not only reduces the cost of recruitment payments but also improves the accuracy of trust evaluation. The recognition rate of false data attacks and flooding attacks and the recognition error rate of normal nodes are 100%, 98.9%, and 3.9%, respectively. The overall quality of data collection is improved

The rest of this paper is organized as follows. The related works are reviewed in Section 2. The system model and problem statement are described in Section 3. In Section 4, the DC-LTM scheme for IoT is proposed. The simulation results and performance are analyzed in Section 5. Finally, we conclude in Section 6.

2. Related Work

With the development of IoT technology and the Internet of Everything (IoE), more and more sensing devices have begun to be integrated into the Internet, leading to the development of new network systems such as Smart Cities and Internet of Vehicles. Numerous sensing devices are generating a large

amount of data, and intelligent network systems need to analyze and process the large amounts of data generated by sensing devices to improve network performance. Intelligent strategies are provided to improve network performance and provide high-quality services for users [13–15]. At the same time, these data often contains a large amount of sensitive data. It also puts higher requirements for the security of the Internet of things.

The edge computing has improved the security and real-time performance of data, but the security of the execution environment of the edge computing nodes is still a nonnegligible issue, which threatens the security of the entire edge computing model. The edge is often a high-value target for attackers, as it collects data from multiple terminals. Compared with terminal, edge is more vulnerable. Once the edge is broken, the attacker can not only access confidential data but also delete or forge data. Hardware vendors are now introducing Trusted Execution Environments (TEE) on various platforms and integrating TEE into edge computing nodes. It can effectively guarantee the computing security on these nodes. Common TEEs include Intel Software Guard Extensions (SGX) [16, 17], x86 system management mode (SMM) [18], and AMD platform security processor (PSP) [19]. On the other hand, edge computing is based on the acquisition of massive data, and the quality of data collection plays a decisive role in the application. Therefore, collecting data in a low-cost, high-quality, and reliable manner is a challenge issue. This paper focuses on data collection. As the number of vehicles in cities continues to increase, MV-based data sensing methods have gradually attracted the attention of researchers.

2.1. Low-Cost Data Collection Strategy. Bonola et al. proposed a low-cost data collection strategy for smart cities by opportunistic routing, in which sensing devices are connected to mobile vehicles via single or multihops, and data from sensor nodes is indirectly routed to the data center due to the 5G communication capability of mobile vehicles [20]. This strategy makes full use of the existing devices in the edge network and is a low-cost data collection strategy without the need to deploy dedicated data collection devices and other infrastructure. Abdelhamid et al. [9] combined sensing devices and mobile vehicles based on the literature [20], giving rewards based on the distance travelled by the vehicle and calculating the cost by considering only the distance travelled by the vehicle and the coverage of the city. He et al. propose a trajectory-based vehicle recruitment framework, which considers spatiotemporal availability and participant reputation, and recruit vehicles to achieve the desired spatial coverage within a limited budget [21].

2.2. High-Quality Data Collection Strategy. The quality of data collection is mainly reflected in the data collection rate and accuracy rate. Data collection rate mainly refers to the proportion of data collected or the proportion of collected data covering the whole collection area. Ren et al. proposed a vehicle recruitment strategy with the goal of optimizing data collection quality as an incentive for vehicles to perform data collection [8]. The literature [9] proposes a greedy approximation algorithm to solve the problem of vehicle participant recruitment, which achieves high quality mobile swarm intelligence perception on a limited budget. The aim of the scheme is to recruit participants to maximize coverage within a limited budget.

2.3. Trust-Based Data Collection Strategies. All of the above strategies are based on the assumption that terminal nodes and vehicles are completely reliable. In fact, there is a risk that the terminal nodes or vehicles may be attacked, or some vehicles may maliciously provide false data in order to get more rewards. If the data collected is not trustworthy, it will affect normal decision-making and have a serious impact on the system. Trust-based data collection strategy is an effective method to identify false data [22–24] and is commonly used in IoT.

Wang et al. proposed an intelligent evaluation scheme based on mobile edge computing, and a probabilistic graphical model was designed to ensure the trustworthiness of nodes and reduce energy consumption [22]. Tanaka et al. proposed using the Beta function as a probability density function for trust evaluation of nodes [25]. Xu et al. proposed adding a unique hash value at the intermediate node when the packet is forwarded, verifying the hash value of the intermediate node uniformly at the destination node, and updating the counter of trust by the result of the verification [26]. These strategies are all passive trust mechanisms, and nodes that have not interacted for a long time may need to take more risks to establish communication. In recent years, there are many researches on active trust mechanism. Wang et al. proposed a trust assessment mechanism using crowdsourcing and intelligent mobile edge computing. Mobile edge users can obtain various information and determine whether a node is trustworthy through proximity access to the terminal node [23]. Sharma et al. proposed the use of a more trustworthy base station to periodically collect and check packets sent by neighbouring nodes, which has ensured the security of the routing path [27], while Wang et al. [28] argue that the active mechanism requires initiating routes to actively test other nodes, which also implies the consumption of additional energy. Hu et al. further proposed a data collection approach using a combination of UAV and mobile vehicles to collect data with less delay [24]. Regarding the data collection strategy of UAVs in WSNs, many researchers focus on the path planning and scheduling of UAVs. Jiang et al. proposed a UAV trajectory optimization algorithm based on ant colony [29]. Li et al. [30] proposed an evolutionary path planning algorithm to maximize the information collection of UAVs. Liu et al. [31] proposed that the city should be divided into several districts to maximize the capacity of each node according to the number of WSN nodes, the size of WSN cluster, and the number of UAVs.

There are few studies on recruiting MVs to conduct trust evaluation on terminal nodes and using UAVs as authorization verification tools to actively verify the authenticity of data reported by MVs. In this paper, the main responsibility of the UAV is to perform data validation and improve the quality of the system data collection in general.

3. System Model and Problem Statement

3.1. Smart City Network Model. The network model based on DC-LTM is shown in Figure 1, which includes sensing devices, data collection tools, and data center. In smart city, many city infrastructures, such as street lights and rubbish bins, are equipped with sensing devices to monitor surroundings. Typically, these sensors have simple hardware and a communication range of only a few tens of meters. They can use their limited computing power to store data temporarily, and data is transferred from the sensors to the mobile data collection tool when an Internet-connected data collection tool moves into these sensors. And then data is reported to a more advanced data center through data collection tools. Mobile vehicles (MVs) and unmanned aerial vehicles (UAVs) are data acquisition tools that have been used in many researches. The data generated by the sensors will be received and stored by the data acquisition tool. The data collected by the UAVs or MVs will be processed further in data center, including trust assessment, recruitment of MVs, rewards, and decision-making [23]. Some advanced data centers also have the ability to dispatch UAVs [32, 33].

3.2. Problem Statement. Figure 2 illustrates the structure of data collection based on layered trust mechanism (DC-LTM). Cloud data center recruits mobile vehicles for data collection, and the mobile vehicles obtain sensored data and calculate the trust value of the terminal nodes and then report it to the data center. At the same time, the UAV is sent to actively verify the trust value of the nodes and vehicles in the network at an appropriate time and then report it to the data center. The cloud data center aggregates the reported data and then evaluates the trust value of the whole network. Finally, the mobile vehicle is paid accordingly. A sensor node needs to be evaluated several times with different vehicles, and the cloud performs a comprehensive calculation to get the final evaluation result to ensure accuracy. The scheme proposed can collect data from the nodes. It obtains the trust assessment of each terminal node in the whole network, as well as determining the trustworthiness of the mobile vehicle.

The system uses the following performance indicators.

3.2.1. Evaluate Cost. The evaluation cost of the trust mechanism includes the cost of the recruited MVs and the cost of using the UAV, as shown in Equation (1):

$$\text{cost} = \sum_{j \in \text{MVs}} \text{ec}_j \bullet z_j + \sum_{k \in \text{UAV}} \text{uc}_k, \tag{1}$$

where ec_j denotes the evaluation cost of the jth MVs. And $z_j \in \{0, 1\}$ indicates whether MVs are recruited. uc_k represents the cost of a single UAV. Obviously, the evaluation cost of trust mechanism is as low as possible.

3.2.2. Evaluate Quality. The evaluation quality of the trust mechanism refers to the overall evaluation quality of the recruited MVs, as shown in Equation (2):

$$\text{quality} = \sum_{j \in \text{MVs}} \text{eq}_j \bullet z_j, \tag{2}$$

where eq_j denotes the evaluation quality of the jth MVs. Obviously, the recruitment of MVS should make the overall evaluation quality of the system higher.

3.2.3. Evaluation Accuracy. The evaluation accuracy of the trust mechanism is that the accuracy of the system can effectively distinguish between malicious and normal nodes, as shown in Equation (3):

$$\text{accuracy} = \frac{\sum_{j \in MVs} R\text{num}_j \bullet z_j}{\sum_{j \in MVs} \left(R\text{num}_j + W\text{num}_i\right) \bullet z_j}, \tag{3}$$

where $R\text{num}_j$represents the number of MVs_jcorrectly reported trust evaluation values of sensing nodes, and $W\text{num}_i$ represents the number of MVs_j incorrectly reported trust evaluation values of sensing nodes. Obviously, the trust evaluation mechanism is as accurate as possible.

Therefore, the goal of the data collection strategy of this mechanism is shown in Equation (4):

$$\begin{cases} \min\ (\text{Cost}) = \min\left(\sum_{j \in \text{MVs}} \text{ec}_j \bullet z_j + \sum_{k \in \text{UAV}} \text{uc}_k\right), \\ \max\ (\text{quality}) = \max\left(\sum_{j \in \text{MVs}} \text{eq}_j \bullet z_j\right), \\ \max\ (\text{accuracy}) = \max\left(\frac{\sum_{j \in \text{MVs}} R\text{num}_j \bullet z_j}{\sum_{j \in \text{MVs}} \left(R\text{num}_j + W\text{num}_i\right) \bullet z_j}\right). \end{cases} \tag{4}$$

4. Scheme

4.1. Overview. First, the cloud data center publishes the task to all MVs in the region. The publication includes the network topology of the sensor network and the location *positi* $on(x_i, y_i)$ of each sensor node n_i, as well as the collection value $\text{val}(n_i)$ of each sensor. Secondly, the interested MVs report its evaluation quality, evaluation area, evaluation value, and evaluation cost to the cloud data center after receiving the publication, and we name the four of them as bids. Third, the cloud data center decides the winner set based on minimizing the collection cost per sensor and maximizing the overall assessment quality according to each MVs' bid and recruits MVs to collect sensor node data. Fourth, the selected MVs use their mobility abilities to collect sensor node information and calculate the node's trust value. Fifth, UAVs are dispatched to verify the trust of sensor nodes and MVs. Sixth, the cloud data center implements the result summary and comparison mechanism to obtain the final trust evaluation of each sensor and MVs, and at the same

Figure 1: Smart city network model for DC-LTM scheme.

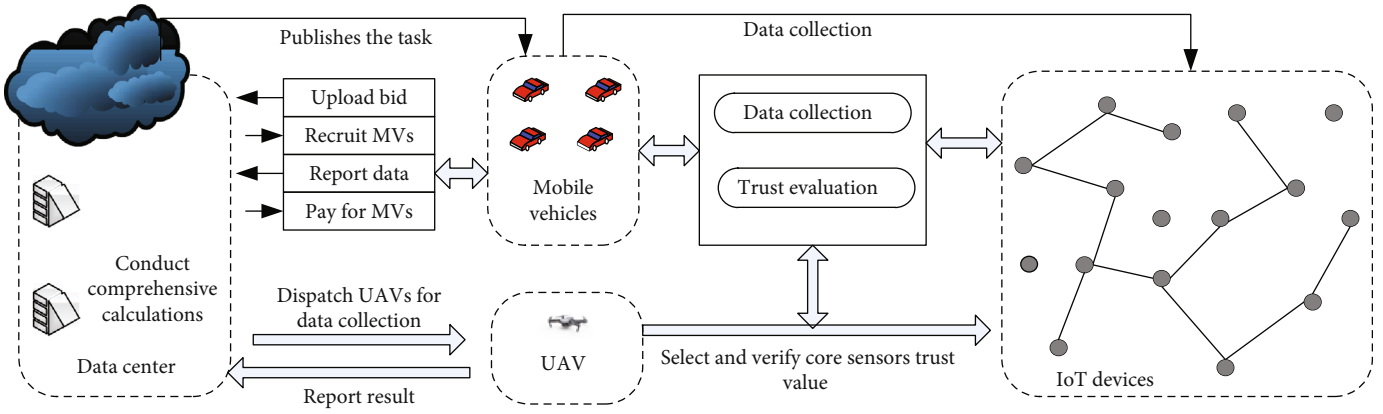

Figure 2: The structure of data collection based on layered trust mechanism (DC-LTM).

time, it also obtains the sensor collection information. Finally, the MVs are paid.

4.2. Programme

4.2.1. Task Publishing. In this section, the cloud data center publishes the task to all MVs in the region.

Definition 1. Task(topology, position(x_i, y_i), val(n_i)). Task is a ternary formula Task(topology, position(x_i, y_i), val(n_i)). It includes the network topology of the sensor network, the location position(x_i, y_i) of each sensor node, and the collection value val(n_i) of each sensor. n_i indicates the ith sensor.

Definition 2. The collection value val(n_i) *of the sensor node.* If every sensor collected received the same reward, there would be a reluctance for MVs far from the data center to collect it, as they would need to travel longer distances and spend more time and effort bringing it back to the data center for the same reward. MVs which are closer to the data center can get more rewards without driving long distances. In our recruitment mechanism, the collection value of a sensor node is related to the distance from data center. The longer the distance, the higher the value; the closer the distance, the lower the value. The collection value of each sensor is calculated by Equation (5), where distance$(|n_i,$ data center$|)$ refers to the distance of the sensor node from its own nearest data center.

$$\text{val}(n_i) = \log_a \text{distance}(|n_i, \text{data center}|). \tag{5}$$

4.2.2. MV Bids. In this section, the interested MVs report its evaluation quality, evaluation area, evaluation value, and evaluation cost to the cloud data center after receiving the publication, and we name the four of them as bids.

Definition 3. The bid for a MVs_j *is* $\text{bid}(\text{MVs}_j)$. This is a quadruple constraint containing the evaluation quality (eq_j), evaluation area (ea_j), evaluation value (ev_j), and evaluation cost (ec_j) of the MVs. The bids of MVs are obtained by Equation (6), where MVs_j denotes the jth MVs.

$$\text{bid}\left(\text{MVs}_j\right) = \left(\text{eq}_j, \text{ea}_j, \text{ev}_j, \text{ec}_j\right). \tag{6}$$

Definition 4. Evaluation quality $(\text{eq}_j), \text{eq}_j \in [0, 1]$. Assume that a sensor node has the truth of state value Tn_i, where $Tn_i \in \{-1, 1\}$, and $Tn_i = -1$ represents a malicious node, and $Tn_i = 1$ represents a normal node. Evaluation result $(\text{MVs}n_{ij})$ is a trust evaluation result obtained by MVs_j for node n_i. When Tn_i converges to $\text{MVs}n_{ij}$, the higher eq_j is, the higher the evaluation quality of MVs is. In this paper, we assume that eq_j is related to the intrinsic characteristics of the MVs. Thus, for any MVs, eq_j is a constant that is independent of the sensor being evaluated. The eq_j of each MVs is given randomly, $\text{eq}_j \in [0, 1]$.

Definition 5. Evaluation area (ea_j). Once a task is distributing, the vehicle is able to move freely within its action radius and collect sensor data. $r(\text{MVs}_j)$ is the action radius of the jth MVs, and ea_j is the set of sensor nodes within the action radius $r(\text{MVs}_j)$.

Definition 6. Evaluation value (ev_j). ev_j denotes the sum of the sensor values collected within the evaluation area of the jth MVs. The further the sensor is from the data center, the higher its value, as in Equation (7).

$$\text{ev}_j = \sum_{n_i \in \text{ea}_j} \text{val}(n_i) = \sum_{n_i \in \text{ea}_j} \log_a \text{distance}(|n_i, \text{data center}|). \tag{7}$$

Definition 7. Evaluation cost (ec_j). ec_j represents the collection cost per sensor in the evaluation area of the jth MVs, as in Equation (8). A higher sensor value means a higher collection cost. $n\text{ea}_j$ represents the number of sensor nodes in the evaluation area.

$$\text{ec}_j = \frac{\text{ev}_j}{n\text{ea}_j}, \tag{8}$$

$$\text{if } n\text{ea}_i = n\text{ea}_j, \text{ev}_i > \text{ev}_j, \text{then ec}_i > \text{ec}_j, \tag{9}$$

$$\text{if } n\text{ea}_i > n\text{ea}_j, \text{ev}_i = \text{ev}_j, \text{then ec}_i < \text{ec}_j. \tag{10}$$

4.2.3. Recruit MVs. In this section, the cloud data center determines the winner set according to the recruitment strategy and recruits MVs to collect sensor node data based on the bids of each MVs. The objective function of the recruitment strategy is shown in Equation (11).

$$\max_z \frac{\sum_{i \in \text{Sensors}} \text{benefit}_i}{\sum_{j \in \text{MVs}} \text{ec}_j \bullet z_j}, \tag{11}$$

$$s.t. \forall i \in \text{Sensors}, \text{benefit}_i = \max_{j \in \text{MVs}} \left(\text{eq}_j \bullet x_{ij} \bullet z_j\right). \tag{12}$$

The vector z is being optimized, where $z_j = 1$ refers to the jth MVs being recruited, and $z_j = 0$ indicates that the jth MVs is not recruited. The denominator refers to the recruitment cost of the system, and the numerator represents the overall evaluation quality of the recruited MVs.$\text{benefit}_i = \max_{j \in \text{MVs}} (\text{eq}_j \bullet x_{ij} \bullet z_j)$ refers to the adjusted evaluation quality of recruited MVs, $x_{ij} = 1$ indicating that the jth MVs is able to cover the sensors i.

For example, let the 1th, 2th, and 3th MVs cover the 1th sensor node, namely, $x_{11} = x_{12} = x_{13} = 1$, the eq_j of three vehicles are 0.9, 0.5, 0.1, respectively, and ec_j is 0.1. It makes more reasonable to have the 1th Mvs to cover this sensor, namely, $z_1 = 1$, and in order to minimize $\sum_{j \in \text{MVs}} \text{ec}_j z_j$, it should make $z_2 = z_3 = 0$.

The pseudocode of the recruitment strategy is presented in Algorithm 1.

Input: $MVs = \{MVs_1, MVs_2, \cdots, MVs_j\}$, $bid = \{bid(MVs_1), bid(MVs_2), \cdots, bid(MVs_j)\}$
Output: $SMVs$, $PMVs$
$SMVs = \varnothing$,//$SMVs$ represents the set of MVs recruited.
$PMVs = \varnothing$,//$SMVs$ represents the payment set of MVS recruited.
$S(SMVs) = \varnothing$,//$S(SMVs)$ represents sensor nodes set that $SMVs$ can collect data .
For each $j \in MVs$
caculate $bid(MVs_j)$ and form bid//Calculate the bid of each MVs and put the result into the set bid
End for
$ec = \{ec_1, ec_2, \cdots, ec_j\}$
While $S(SMVs)$ cannot cover the whole network do
$MVs_j = \max_z(\sum_{i \in Sensors} benefit_i / \sum_{j \in MVs} ec_j \bullet z_j)$
If$S(MVs_j) \subseteq S(SMVs)$
$bid = bid \setminus bid(MVs_j)$
else
$SMVs = SMVs \cup MVs_j$
$bid = bid \setminus bid(MVs_j)$
End if
End while//Find MVs with the maximum value-cost ratio
Call Algorithm 2(CTVN)
For each $i \in SMVs$
$PMVs = PMVs \bigcup bid(MVs_i)$
End for

Algorithm 1: Recruitment strategy (RS).

4.2.4. Trust Evaluation. In this section, the selected MVs use their mobility abilities to collect sensor node information and calculate the node's trust value.

(1) Trust Evaluation Mechanism. The trust evaluation mechanism consists of direct trust, recommended trust, active trust, and comprehensive trust. Direct trust is some monitoring evidence based on communication, energy, and data transmission. Recommended trust is the collection and calculation of recommended values from neighboring sensors. Active trust is trust evaluation of the UAV on the data reported by the MVs and the core sensor nodes. Comprehensive trust is a combination of recommended trust, direct trust, and active trust.

Identifying malicious nodes is crucial to the security of the IoT. We propose a scheme based on a layered trust evaluation mechanism in this paper. The scheme is as follows: MVs are sent out to collect data from nodes and calculate the trust value of nodes. However, some MVs may exaggerate their costs or report false data to get extra rewards. Therefore, UAVs are sent out to evaluate the data reported by MVs at the appropriate time, so that trusted vehicles can be obtained and paid accordingly.

We have designed a layered trust evaluation mechanism, as shown in Figure 3, which is divided into two layers, the WSN layer and the cloud layer. In the WSN layer, the sensors monitor the communication behavior of other sensors and perform the recommended trust calculation and upload the recommended trust list to the cloud layer. The cloud layer consists of two parts, MVs and UAV. The main function of MVs is to directly obtain the energy state of the sensor nodes and perform direct trust calculation and infer the primary comprehensive trust value of the nodes with the recommended trust uploaded from the WSN layer and send the results to the cloud. The main function of UAV is to verify the trust values of the core sensors and MVs and then calculate the active trust. Finally, the UAV obtains the final comprehensive trust value of the node based on the recommended trust, direct trust, and active trust and sends it to the cloud for trust status analysis, which is used to reason about the trust values of the whole network. The purpose of the scheme is to perform data analysis, calculation, and storage trust in the cloud. Therefore, it can defend against more attacks with less consumption of network resource within a tolerable delay.

(2) Recommended Trust. We use the communication behavior of neighbor sensor nodes as the recommended trust value. According to the literature [34], the communication behavior of sensor nodes is related to the number of packets sent and the success rate of communication.

We assume that the number of packets sent by a sensor node follows a normal distribution, which can be modeled by a probability density function $f(x) = (1/\sqrt{2\pi}\sigma) \exp(-(x-\mu)^2/2\sigma^2)$, where μ and σ^2 are the mean and variance of the packets sent, respectively. In general, the more interactions between two nodes, the higher the trust value. However, when the number of packets sent by a sensor node exceeds a threshold in a certain period of time, the trust value of the node should be reduced to prevent flooding attacks. The trust value for the number of packets $T_{d\text{num}}^{ij}$ from node i to node j

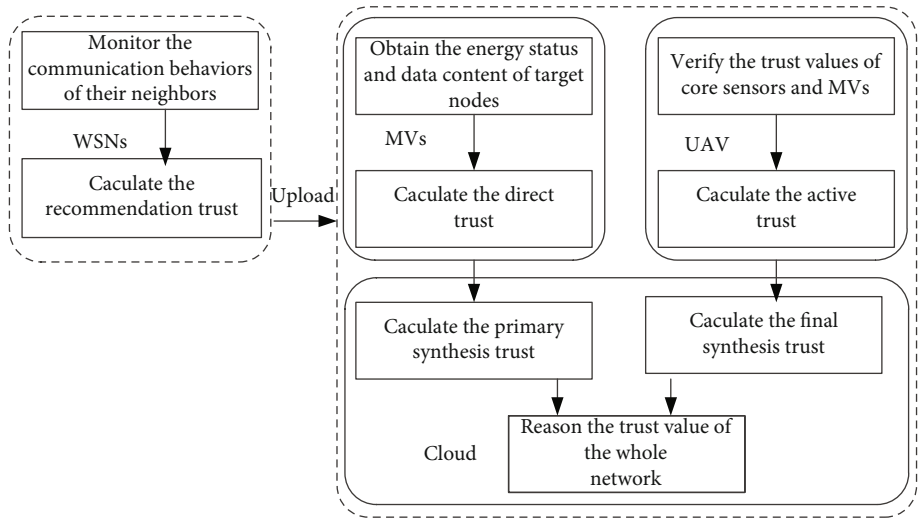

Figure 3: The cloud-based layered trust evaluation mechanism.

is defined as Equation (13):

$$T_{d\text{num}}^{ij} = \begin{cases} \dfrac{d\text{num}_{ij}}{\max\left(d\text{num}_j\right)} & d\text{num}_{ij} \le \lambda\mu_d \\ \exp\left(-\dfrac{\left(d\text{num}_{ij} - \mu_d\right)^2}{2\sigma^2}\right) & d\text{num}_{ij} > \lambda\mu_d \end{cases}, \tag{13}$$

where the variables $d\text{num}_{ij}$ are the number of packets sent by the node n_i to the node n_j, $\max\left(d\text{num}_j\right)$ is the maximum number of packets received by the node from its neighbors, μ_d is the average number of packets sent by the neighbor nodes of node j, and λ and σ are important factors used to adjust the threshold and reduce the trust value of the node.

The communication success rate is another important criterion to evaluate whether a sensor node is trustworthy or not. When a sensor node communicates with a target node, the higher the communication success rate is, the higher the trust value of the target node is, and vice versa. According to the literature [34], the trust value for communication success rate T_{com}^{ij} from node i to node j is defined as Equation (14):

$$T_{\text{com}}^{ij} = \frac{2\kappa_{ij} + \xi_{ij}}{2}, \tag{14}$$

where $\kappa_{ij} = s_{ij}/s_{ij} + f_{ij} + 1$, $\xi_{ij} = 1/s_{ij} + f_{ij} + 1$, s_{ij}, is the number of times that the node n_i and the node n_j have successfully communicated in a certain time, and f_{ij} is the number of times they have failed to communicate.

Therefore, the recommended trust value T_{rec}^{ij} that the node n_i reasoned the node n_j is calculated as shown in Equation (15):

$$T_{\text{rec}}^{ij} = \omega \bullet T_{d\text{num}}^{ij} + (1-\omega) \bullet T_{\text{com}}^{ij}, \tag{15}$$

where ω is the weight.

When MVs access the sensor node, they are able to obtain the overall recommended trust value T_{rec}^{j} about the node n_j, which is shown in Equation (16):

$$T_{\text{rec}}^{j} = \text{average}\left(\sum_{i=1}^{i} T_{\text{rec}}^{ij}\right), \tag{16}$$

where $\sum_{i=1}^{i} T_{\text{rec}}^{ij}$ is the sum of the recommended trust values of nodes 1, 2, ..., i to node n_j.

(3) Direct Trust. We assume that the energy state of the sensor nodes obtained by MVs as the direct trust value. We assume that the energy consumption of the sensor nodes follows a normal distribution. If malicious nodes launch an attack, they will consume additional energy. When the residual energy of a sensor node is less than a threshold, the energy trust is zero. When its residual energy is greater than the threshold, MVs consider the energy consumption rate to calculate the energy trust of sensor node. When MVs visits the jth node, it can obtain the energy trust value T_{en}^{j} about the node, as shown in Equation (17):

$$T_{en}^{j} = \begin{cases} 0 & E_{\text{left}}^{j} \le \lambda\mu_e \\ \exp\left(-\dfrac{\left(E_{\text{cost}}^{j} - \mu_e\right)^2}{2\sigma^2}\right) & E_{\text{left}}^{j} > \lambda\mu_e \end{cases}, \tag{17}$$

where μ_e is the average energy consumption rate of the jth node's neighboring nodes, and λ and σ are important factors used to adjust the threshold and reduce the trust value of the node.

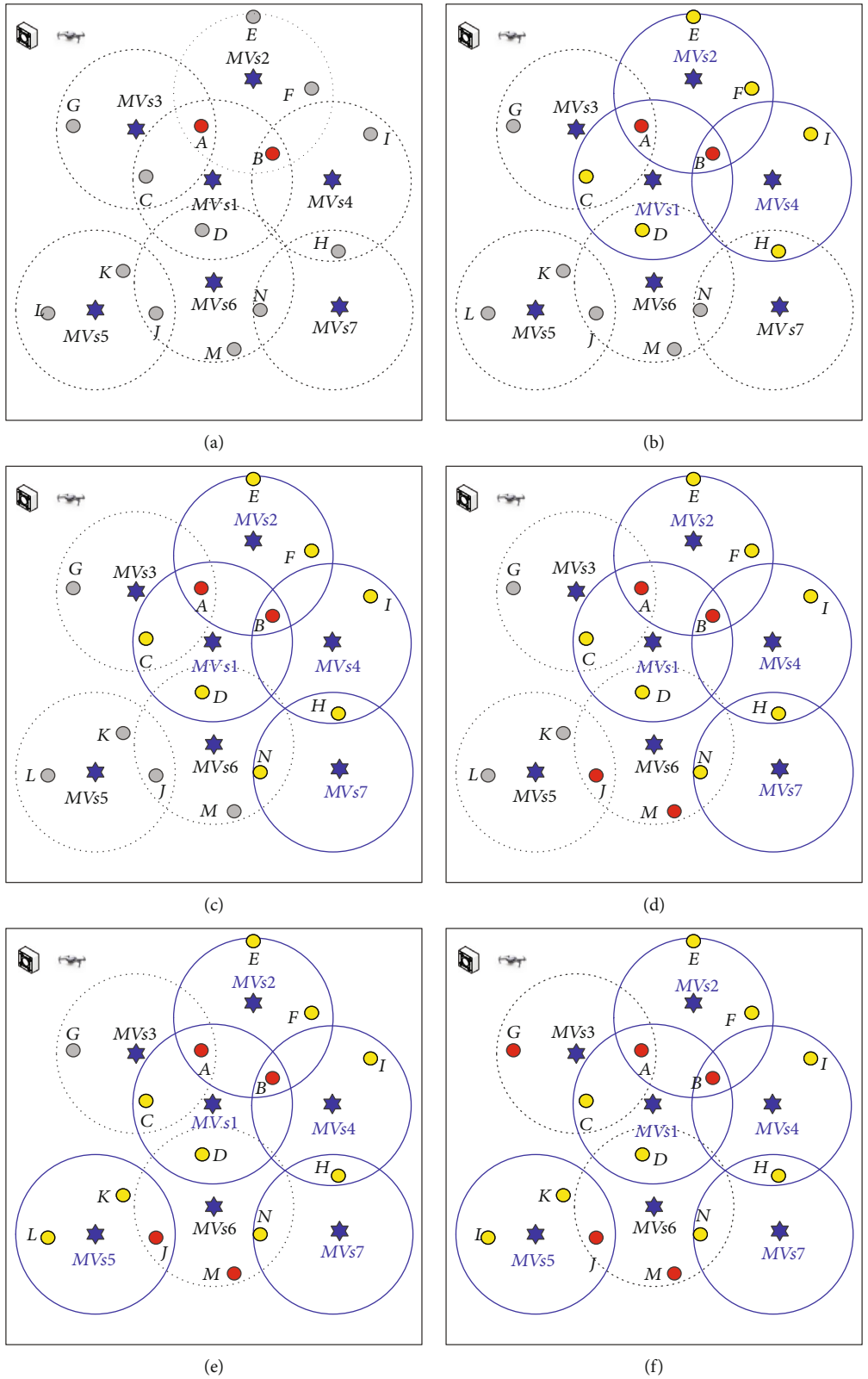

Figure 4: Selection and formation of core sensors, baseline sensors, and baseline MVs.

Input: $Sensor$, $SUAV$, $SMVs$
Output: T^j_{syn}, $BMVs$
$SUAV = \varnothing$//$SUAV$ represents the core sensor nodes selected by UAV for each round of flight
$BMVs = \varnothing$//$BMVs$ represents baseline MVs
$BSensor = \varnothing$//$BSensor$ represents baseline nodes
flight =1//flight represents the number of rounds that UAV dispatched
For each $node_j \in Sensor$//$Sensor$ represents the set of nodes
  Calculate T^j_{rec} by Equation (16)//Calculate the recommended trust value for each node
  Calculate T^j_{en} by Equation (17)//Calculate the direct trust value for each node
  T^j_{act}=0//The initial active trust value for each node is 0 when no UAV has been dispatched
End For
For each $node_j \in SUAV$ do//each node in the core sensor nodes selected by UAV
  IfT^j_{act}=0//The active trust value of node is 0
    Calculate T^j_{act} by Equation (18)//calculate the active trust value for this node
    Calculate T^j_{syn} by Equation (19)//calculate the comprehensive trust value for this node
    Verify the trust value of the MVs_k that report data from $node_j$
    While$BMVs$ can not increase **do**
      If the MVs_k is credibly
        $BMVs = BMVs \cup MVs_k$//set the MVs_k as baseline MVs
        Calculate the T^i_{act} of other sensors $node_i$ reported by $BMVs$ by equation (18)
        Calculate T^i_{syn} by Equation (19)//update the comprehensive trust value for this node
        $BSensor = BSensor \cup node_i$
      End If
        Verify the trust value of the $MVs_{k'}$ that report data from $BSensor$
        //data reported from BSensor can also be used as baseline data to verify the trust value of $MVs_{k'}$ which reporting them.
      If the $MVs_{k'}$ is credibly
        $BMVs = BMVs \cup MVs_{k'}$//Set the $MVs_{k'}$ as baseline MVs
      End if
    End while
  End If
  flight = flight+1//the core sensor node needs to be re selected in each round
End For

Algorithm 2: Calculating the trust value of nodes (CTVN).

(4) Active Trust. UAV can establish communication directly with the sensors, and the collected data can be considered as the most authoritative. We call the sensor collected by UAV as the core sensor and data reported by core sensor as baseline data. UAV can obtain the real state value Tn_i of the sensor, refer to definition 4 in Section 4.2.2. The baseline data can be used to validate the data reported by the MVS, thus verifying the trust value of the MVS reported that core sensor. For example, the core sensor data collected by the UAV is compared with the data collected by the MVs. The MVs are credible if data is equal and is not credible if data is unequal.

Adding active trust into the trust evaluation mechanism can not only verify the trust value of the MVs but also avoid recruiting malicious MVs to collect data. At the same time, it can evaluate the trust of each sensor in the network more accurately. The high cost of UAV makes it unable to work continuously. In order to reduce costs, some MVs have been verified as normal (called baseline MVs), and the sensor data reported by them can also be used as baseline data to verify the trust values of other MVs. Therefore, the UAV does not

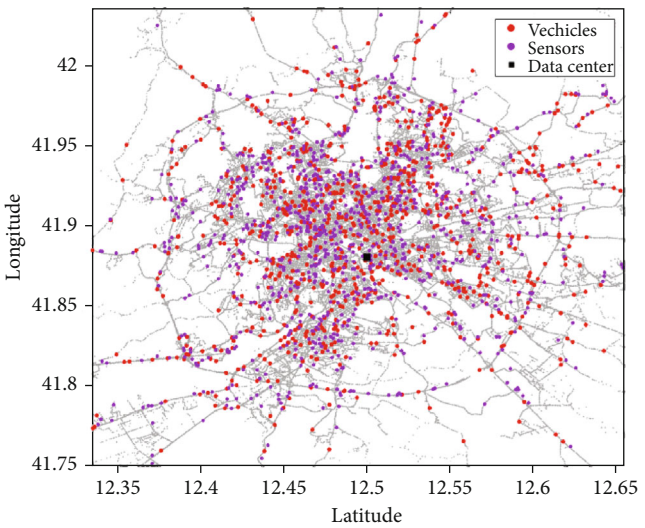

Figure 5: Network model.

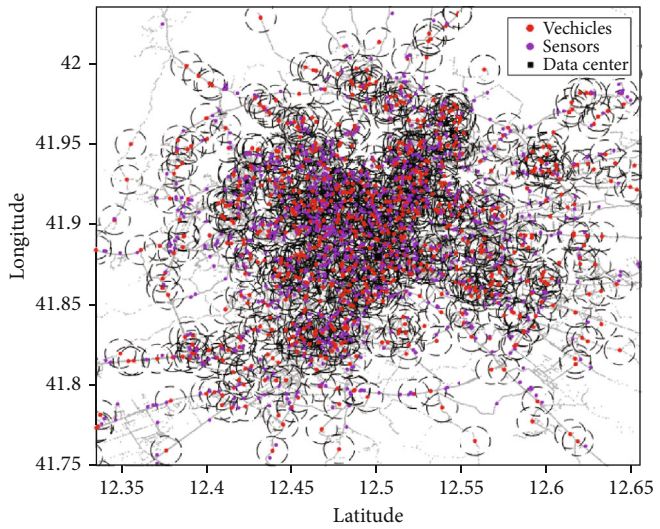

Figure 6: Coverage area of 500 MVs.

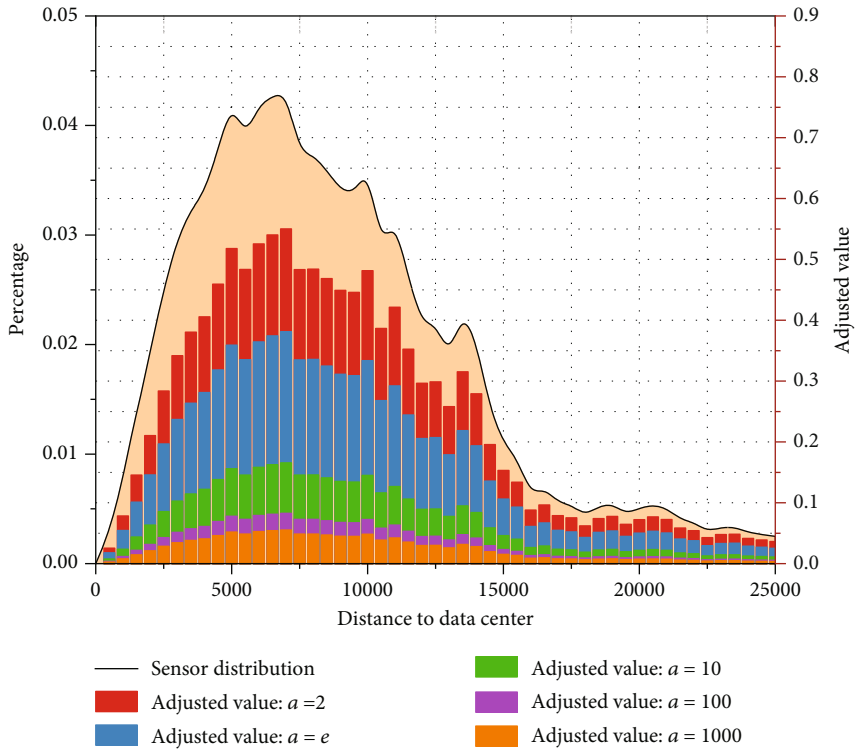

Figure 7: The effect of the base number in adjusting values.

need to collect sensors that can be verified by normal MVs, which means lower costs.

The active trust value of the jth node is calculated based on the validation results of the sensor nodes collected by the UAV or baseline MVs and is calculated in Equation (18), where s_j=1 and f_j=0 indicates successful validation, and s_j=0 and f_j=1 indicate failed validation.

$$T_{act}^{j} = \frac{s_j - f_j + 1}{2} \tag{18}$$

(5) Comprehensive Trust. Three kinds of trust relationships are used in the trust evaluation scheme proposed in this paper, namely, recommendation trust, direct trust, and active trust, the weighting of which combines into the comprehensive trust, as shown in Equation (19).

$$T_{syn}^{j} = \omega_1 \bullet T_{rec}^{j} + \omega_2 \bullet T_{en}^{j} + \omega_3 \bullet T_{act}^{j}. \tag{19}$$

4.2.5. UAV Active Verification. In this section, UAVs are sent to verify the trust of sensor nodes and MVs.

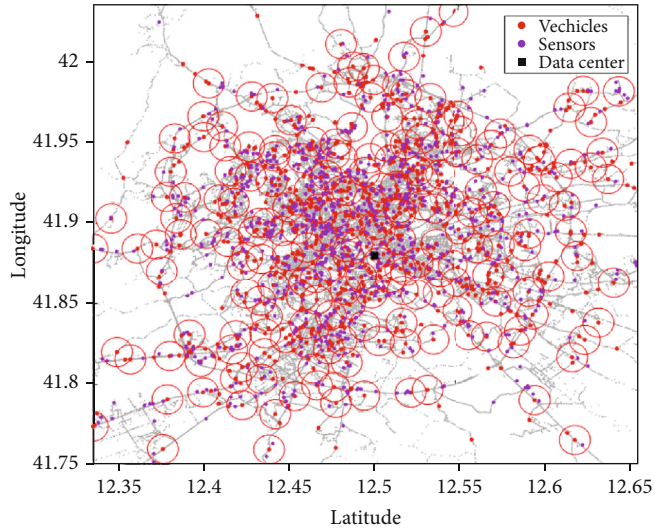

Figure 8: The distribution of vehicles for a recruitment strategy that considers cost only.

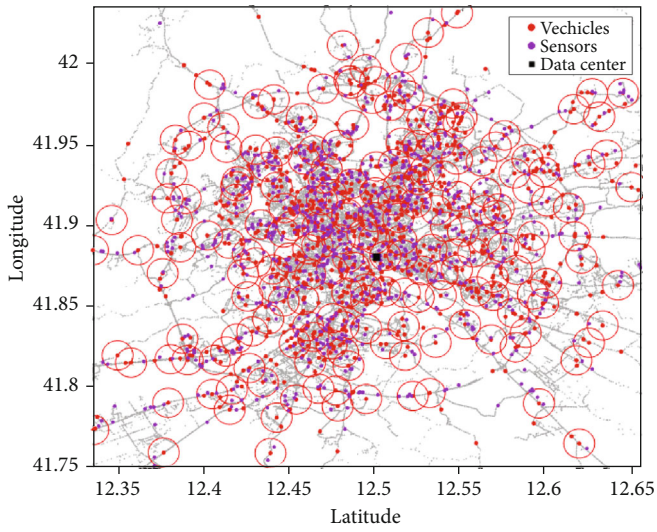

Figure 9: The distribution of vehicles for a recruitment strategy that considers evaluation quality only.

(1) Selection and Formation of Core Sensors, Baseline Sensors, and Baseline MVs. Due to the limited energy of the UAV, a single flight can only access a part of the sensors within its control range, which we call core sensors. In the initial stages of data collection, the UAV can only be used to calculate the active trust of sensor nodes without obtaining the baseline MVS. At this moment, sensor with more nonbaseline MVs reporting should be selected as the core sensor, so as to generate more baseline MVs and improve the efficiency of system trust verification. With the deepening of data collection, the trust verification of UAV is gradually replaced by more baseline MVS. Then, UAV should choose those sensors which have not yet calculated active trust and have the largest number of nonbaseline MVS reports as the core sensor.

The high cost of UAV makes it unable to work continuously. To reduce costs, data reported by some verified normal MVs (called baseline MVs) can also be used as baseline data to verify the trust values of other MVs. As a result, UAVs do not need to collect sensors that can be verified by normal MVs, which means lower costs.

The UAV active verification process is shown in Figure 4, assuming that the number of core sensors is 2.

In the initial stage of data collection in Figure 4(a), sensor node A is reported by MVs_1, MVs_2, and MVs_3, and sensor node B is reported by MVs_1, MVs_2, and MVs_4. Therefore, A and B are selected as the core sensors in the first round of UAV flight, namely SUAV $= \{A, B\}$, and the active trust value of the core sensors is calculated as T^A_{act} and T^B_{act}, and the comprehensive trust value is T^A_{syn} and T^B_{syn}.

In Figure 4(b), verify the trust value of MVs_1, MVs_2, MVs_3, and MVs_4 which reporting core sensor data. Assuming that the baseline MVs are MVs_1, MVs_2, and MVs_4, namely, $\text{BMVs} = \{\text{MVs}_1, \text{MVs}_2, \text{MVs}_4\}$, then calculate the active trust value T^C_{act}, T^D_{act}, T^E_{act}, T^F_{act}, T^H_{act}, and T^I_{act} of other

sensor nodes C, node D, node E, node F, node H, and node I which are collected by BMVs, calculating the comprehensive trust value T^C_{syn}, T^D_{syn}, T^E_{syn}, T^F_{syn}, T^H_{syn}, and T^I_{syn} and obtaining baseline sensor set BSensor = $\{C, D, E, F, H, I\}$.

In Figure 4(c), data reported from BSensor = $\{C, D, E, F, H, I\}$ can also be used as baseline data to verify the trust value of non baseline MVs which reporting them, such as MVs_6, MVs_7. When a new baseline MVs_7 is obtained, BMVs = $\{\text{MVs}_1, \text{MVs}_2, \text{MVs}_4, \text{MVs}_7\}$ is immediately updated. Meanwhile, the active trust value and the comprehensive trust value of nonbaseline node N are calculated. Where N is reported by MVs_7, and the baseline sensor set BSensor = $\{C, D, E, F, H, I, N\}$ is updated accordingly.

In this way, more baseline MVs can be generated.

When no more baseline MVs can be generated, send out UAVs for the second round of verification.

In Figure 4(d), there are still sensor nodes $\{G, J, K, L, M\}$ whose active trust values are 0. Select the sensor node with high reporting of nonbaseline MVs among the sensor nodes $\{G, J, K, L, M\}$ as the core sensor. Since node J is reported by MVs_5 and MVs_6, UAV selects J as the core sensor in the second round. G, K, L, and M all reported only once. Since two core sensors can be selected during each flight, and then select node M as the core sensor because the MVs reporting M can cover more sensors, namely, SUAV = $\{J, M\}$, calculate the active trust value T^J_{act} and T^M_{act} of the core sensor node J and node M and update the comprehensive trust value T^J_{syn} and T^M_{syn} of two nodes.

In Figure 4(e), verify the trust value of MVs_5 and MVs_6 which reporting core sensor node J and M. Assuming that the baseline MVs are MVs_5, namely, BMVs = $\{\text{MVs}_1, \text{MVs}_2, \text{MVs}_4, \text{MVs}_7, \text{MVs}_5\}$, then calculate the active trust value T^L_{act} and T^K_{act} of other sensor nodes L and node K which is collected by MVs_5, calculating the comprehensive trust value T^L_{syn} and T^K_{syn} and obtaining baseline sensor set BSensor = $\{C, D, E, F, H, I, N, L, K\}$.

In Figure 4(f), when no more baseline MVs can be generated, send out UAVs for the third round of verification. At this time, only the active trust value of sensor node G is 0; so, UAV selects node G as the core sensor, SUAV = $\{G\}$, calculating the active trust value T^G_{act} and the comprehensive trust value T^G_{syn} of node G. When the active trust value of all nodes is not 0, the active trust verification ends.

(2) Trust Evaluation Algorithm. The pseudocode of calculating the trust value of nodes is presented in Algorithm 2.

After multiple rounds of flight, the active trust values of all sensors can be calculated. This method obtains the comprehensive trust value and baseline MVs. The baseline MVs is the trusted MVs in this system, and the data center can pay for it according to the baseline MVs.

4.2.6. Payment. Based on the results reported by MVs and UAV, the cloud data center implements the result summary and comparison mechanism to obtain the final trust evaluation of each sensor and MVs, and at the same time, it also obtains the sensor collection information. Finally, the MVs are paid.

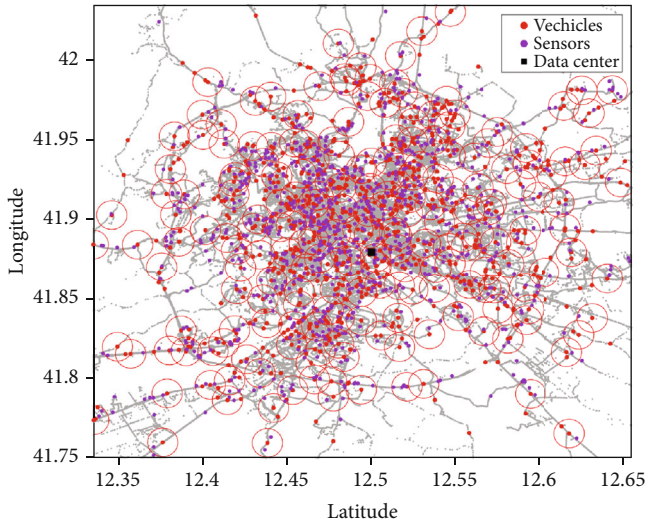

Figure 10: The distribution of vehicles for a recruitment strategy with comprehensive consideration of cost and evaluation quality.

5. Analysis of Experimental Results

5.1. Experimental Environment. The experiments are carried out in MATLAB R2020a. To evaluate the performance of the strategy, we simulated the experimental model using a real dataset. As shown in Figure 5, the red points are the vehicle location points at a certain time, and 1000 sensors and 1 fixed data center are deployed in a probabilistic manner based on the density of vehicle location points. The randomly generated locations of the sensor nodes are proportional to the density of vehicle initial points, making the sensor nodes heavily distributed within urban areas. The data center is deployed in the area with the highest density of GPS coordinates to enable the collection of reported data timely. Assuming that the activity radius of the vehicle is 1000 m, each MVs can collect sensor data within its activity radius, and the communication radius of sensor nodes is 400 m. Figure 6 shows the coverage area of 500 MVs, $\lambda = 1$, and the weight ω=0.5 of T_{rec}. The percentage of nodes sending false data is 20%, the percentage of nodes launching flooding attack is 5%, and the percentage of malicious vehicles is 20%.

An experiment for parameter adjustment was done for the base number a in Equation (5). Figure 7 shows the proportion of the total number of sensors in the network as the distance from the data center increases and the change trend of the adjusted value curve under different bases a.

If the data center is placed in the middle of the city, the majority of sensors are located between 2500 m and 12500 m from the data center, with a decreasing trend in the number of sensors whether the closer or the farther away the data center. Adjusted value means that for a given distance such as 10000 m, the value of the sensors can be obtained by using Equation (5) and then multiplying val in Equation (5) by the proportion of sensors at this distance to obtain the adjusted value. Since we will use val as the

Table 1: A comparison of the three recruitment strategies.

	CCEQ strategy	CEQ strategy	CC strategy
Total cost	2353.63826846331	2425.73520998621	2254.82195098336
Average EQ of recruited MVs	0.793907288837423	0.794030076960136	0.726972632004356

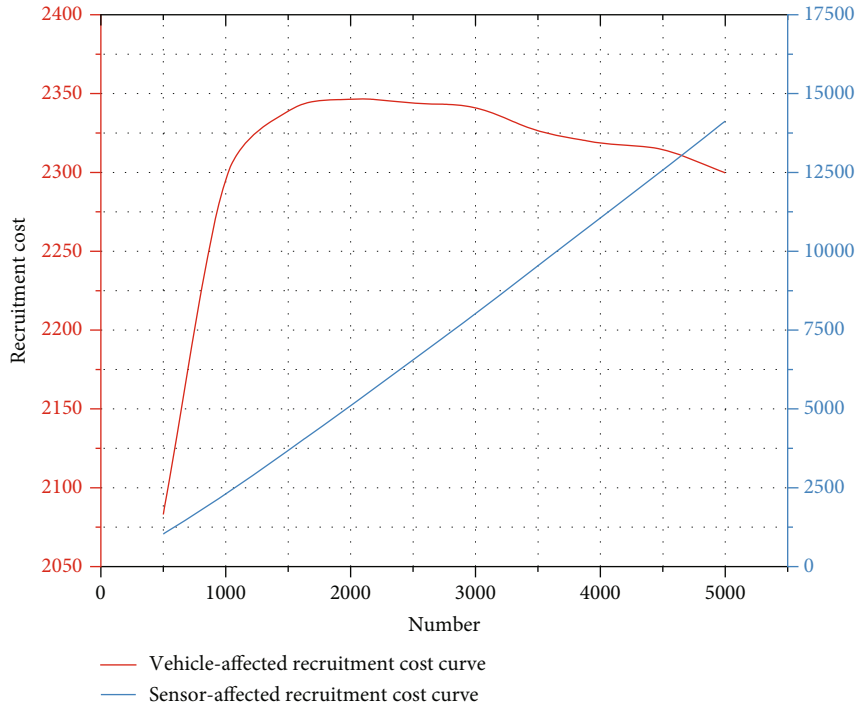

Figure 11: Impact of different numbers of MVs and number of sensor nodes on the total cost.

standard to measure the cost, then adjusted value means the distribution of the cost when recruiting. The smoother the curve, the less difference is got in the total cost when recruiting vehicles from different regions such as 2000 m and 15000 m. The steeper the curve is, the greater the cost difference is got, and the more unfair the recruitment is. Therefore, the base number $a = 100$ has been chosen in order to ensure that the curve is smooth enough, while ensuring that vehicles collecting data from different areas will be rewarded with different levels. For example, 15000 m is definitely greater than 2000 m.

5.2. Comparison of Recruitment Strategies. After the cloud data center publishes the task to all MVs in the region, each interested MVs generate its own bid including evaluation quality, evaluation area, evaluation value, and evaluation cost and reports it to the cloud. The cloud data center determines the winner set according to different objectives and recruits MVs to collect sensor node based on the bids of each MVs.

The distribution of vehicles for a recruitment strategy that considers cost only (CC strategy) is shown in Figure 8, which covers the full network with the smallest sum of costs per sensor collected by the recruited MVs.

Figure 9 shows the distribution of vehicles for a recruitment strategy that considers evaluation quality only (CEQ strategy). That is, the recruited MVSs maximize the overall evaluation quality of the system while satisfying the coverage of the entire network.

Figure 10 shows the distribution of vehicles for a recruitment strategy with comprehensive consideration of cost and evaluation quality (CCEQ strategy). That is, the vehicle distribution map for a recruitment strategy maximizes the overall evaluation quality of the selected MVs while considering cost minimization when the full network is covered.

A comparison of the total cost of the three recruitment strategies and the average evaluation quality of the recruitment vehicles is shown (see Table 1).

5.3. The Impact of Parameters on the Total Cost. The effect of different numbers of vehicles or sensors on the total cost of recruiting MVs is shown in Figure 11. The red line and blue line indicate the cost of recruitment at different numbers of MVs (fixed number of sensors 1000) and sensors (fixed number of vehicles 2000), respectively.

As can be seen from Figure 11, at the beginning, there are not enough vehicles to cover all the sensors. As the number of MVs increases, the vehicles are able to cover more sensors, which also led to increased costs. The total cost increases as the number of MVs increases. However, as the MVs increase further, the cloud data center can choose more cost effective MVs, and the total cost decreases slightly with the number of MVs.

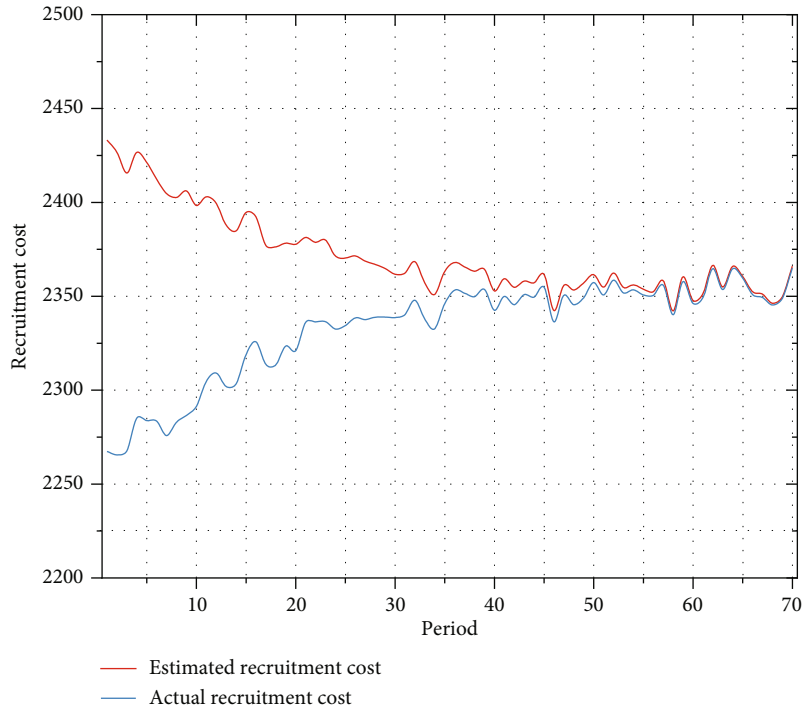

Figure 12: Payment differences when UAV are added to the trust evaluation mechanism.

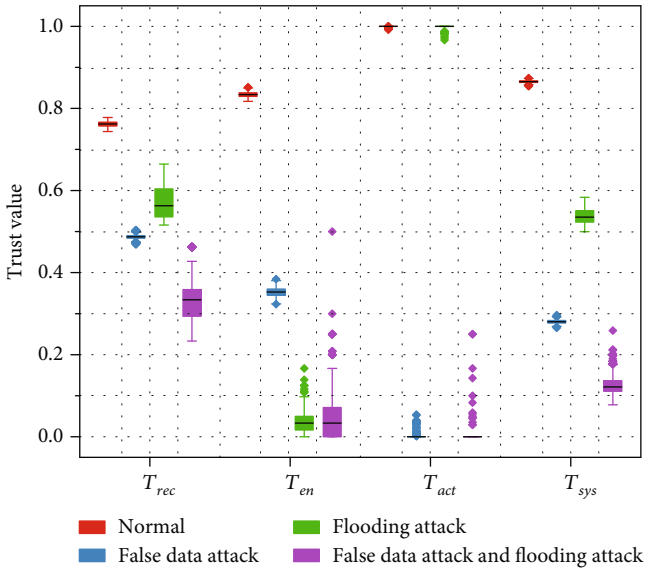

Figure 13: The variability in trust for each type of node after the introduction of UAV.

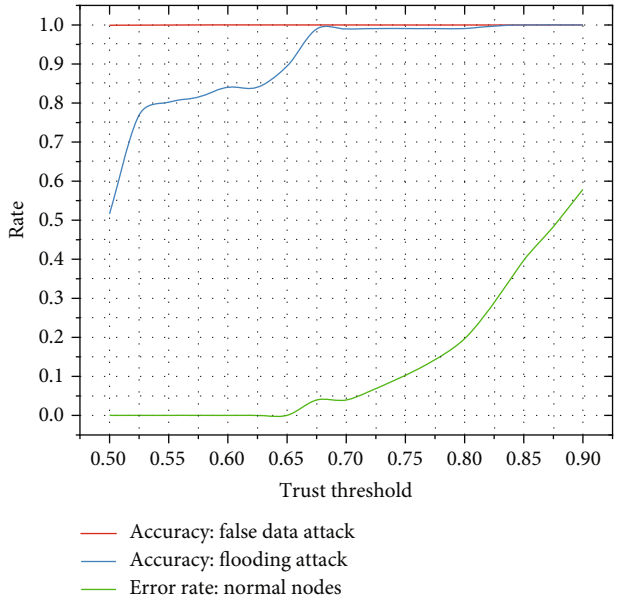

Figure 14: The influence of different trust thresholds.

The total costs increase with the number of sensor nodes. As the number of sensor nodes increases, the cloud data center needs to recruit more MVs to perform the evaluation tasks and therefore the total cost increases.

5.4. The Impact of UAV Participating in Active Verification on System Performance

5.4.1. Total Cost. Based on the results reported by MVs and UAV, the cloud data center performs a result aggregation comparison mechanism to obtain a final trust evaluation for each sensor and MVs. For some malicious vehicles, no payment is made thus reducing payment costs.

Figure 12 illustrates the total payment difference between verifying whether the MVs are reporting true data or not after the UAV joins the trust mechanism. The estimated

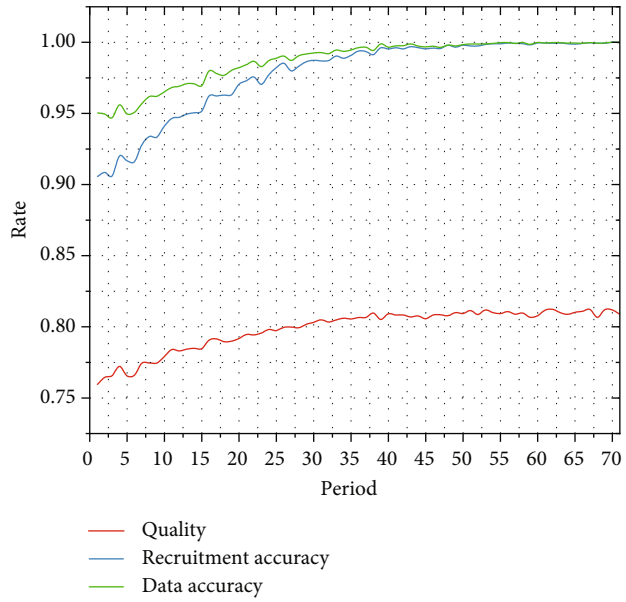

Figure 15: The impact of UAV on evaluation quality.

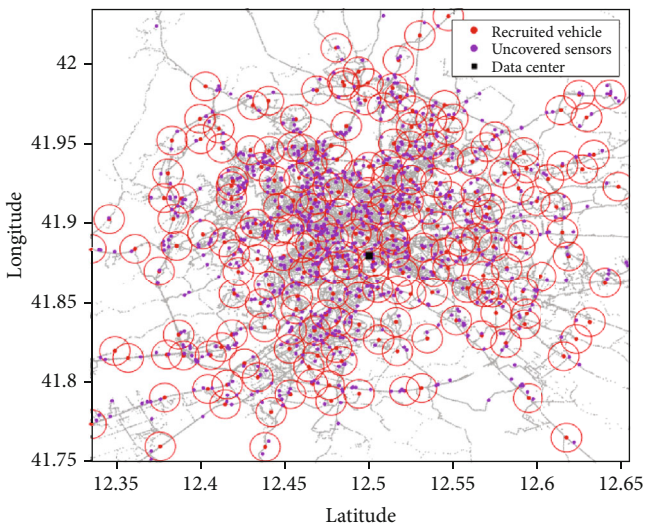

Figure 16: The initial stage.

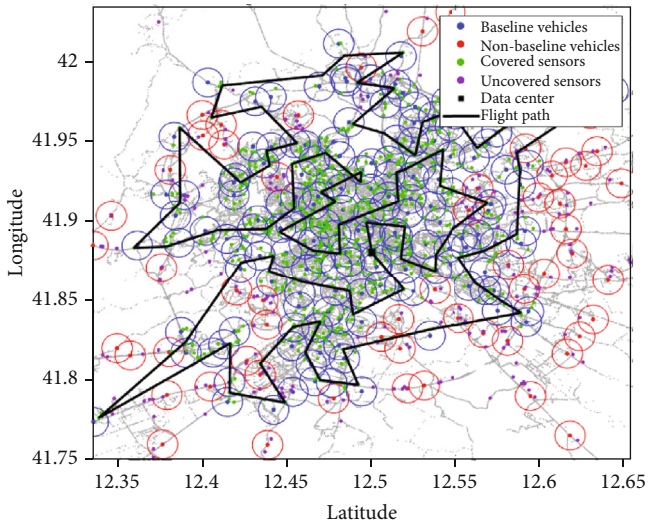

Figure 17: The first flight.

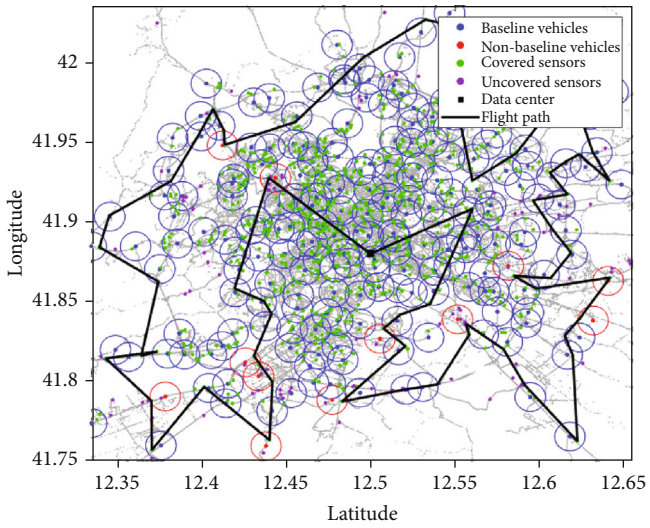

Figure 18: The second flight.

recruitment cost in Figure 12 is the sum bids of all the vehicles recruited, while the actual recruitment cost refers to the payment cost after excluding the malicious vehicles. It can be seen from Figure 12 that the proposed layered trust evaluation mechanism can effectively distinguish between normal vehicles and malicious vehicles. As time goes on, more malicious vehicles are tested and excluded from the recruitment system, and the proposed layered trust evaluation mechanism with UAV participation can effectively distinguish malicious vehicles providing false data from normal vehicles.

5.5. Evaluation Accuracy. To verify the accuracy of the system in distinguishing between malicious nodes and normal nodes, we consider four types of node, namely, normal nodes, nodes that generate false data, nodes that launch flooding attacks, and nodes that launch both false data attacks and flooding attacks. Figure 13 illustrates the variability in trust for each type of node after the introduction of UAV. For example, active trust T_{act} can accurately determine if a node has provided false data, but it cannot identify a flooding attack because the data provided is correct despite the flooding attack being launched. Energy trust T_{en}, however, can accurately determine whether a flooding attack has been launched because the node launching the flooding attack will consume more energy, and it can also determine whether a false data attack has been launched because additional energy is required. T_{rec} is valid for both flooding attack and false data attack. In addition, we assume that all three trusts have the same weight, and then the distribution of trust is as shown in Figure 13. The normal node has a mean trust value of 0.8651, while a node launching a flooding attack has the next highest trust value at 0.5373, a node launching a false

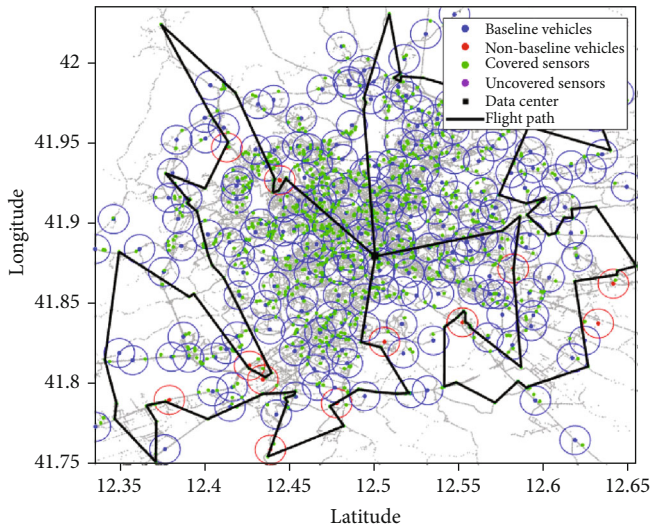

Figure 19: The third flight.

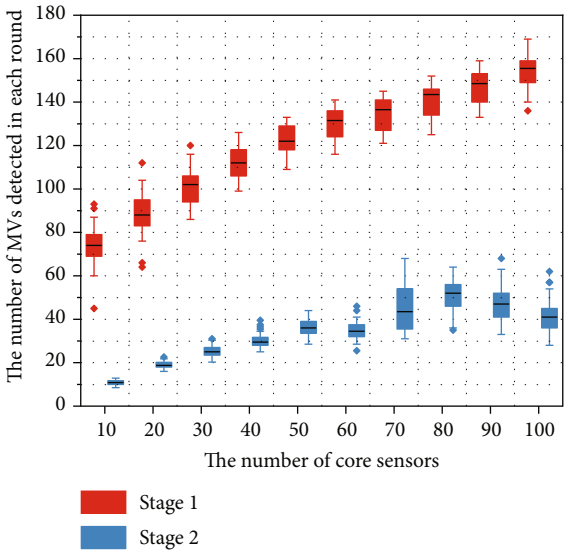

Figure 20: The impact of the number of core sensors on the identification of MVs.

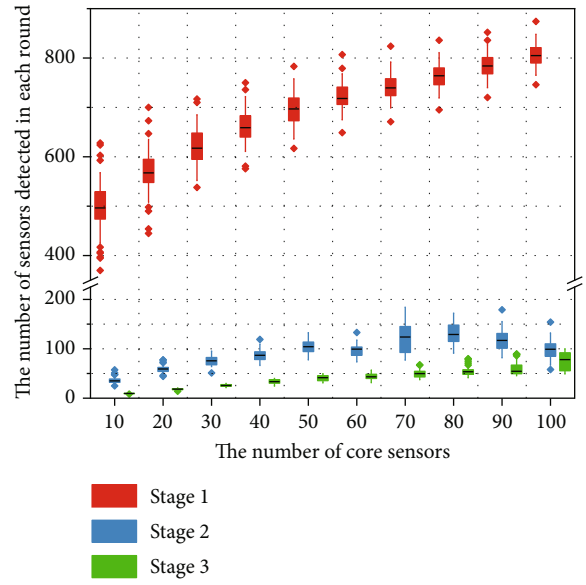

Figure 21: The impact of the number of core sensors on the identification of sensors.

data attack has a lower trust value at 0.2803, and the worst would be a node launching both attacks, with the trust closest to zero at about 0.1252.

In order to distinguish normal nodes from malicious nodes, a trust threshold is obviously needed. We can see in Figure 13 that it should be reasonable to set the threshold between the trust value T_{sys} of the normal node and the highest trust value $T_{sys}$of the malicious node. To further reduce misjudgement behavior including mistaking the node launching the attack as a normal node, or mistaking the normal node generating data frequently as malicious node, we have simulated the threshold interval 0.5 to 0.9.

As shown in Figure 14, red and blue refer to the identification accuracy rate for false data attacks and launching flooding attacks, respectively, while green is the error rate for normal nodes. The results show that when the trust threshold reaches 0.7, it is able to exclude 100% of the false data attacks and 98.9% of the flooding attacks, but may introduce a small identification error rate of 3.9% for normal nodes. However, when the trust threshold is 0.65, 100% of false data attacks is still ruled out, but the identification rate for flooding attacks is reduced to 89.4%, and its identification error rate for normal nodes is 0%. After a compromise, we chose a trust threshold of 0.675, which resulted in a recognition rate of false data attack and flooding attack, and the recognition error rate of normal nodes is 100%, 98.9% and 3.9%, respectively.

5.5.1. Evaluation Quality. Quality in Figure 15 refers to the average eq of recruited vehicles, recruitment accuracy refers to the proportion of normal vehicles to total recruited

vehicles, and data accuracy refers to the proportion of data reported by recruited normal vehicles to total reported data. All three show a slow upward trend because of the gradual exclusion of malicious vehicles.

5.6. Flight Trajectory of UAV. The red dots in Figure 16 show the recruited vehicles, and the red circles are the coverage of the recruited vehicles. The purple dots refer to the sensors that are not covered by the baseline vehicles, and it is need to wait for the UAV to verify and gain coverage of the baseline vehicles. The goal is to evaluate the trust of all sensors and vehicles in the network.

Figure 17 shows the trajectory of the UAV in the first flight and its coverage status for nodes and vehicles. The UAV needs to visit nodes covered by more nonbaseline vehicles in the first stage, so that more vehicles and sensors can be evaluated and thus more baseline devices can be acquired as possible. The blue dots indicate the baseline vehicles that could be evaluated in the first flight, and the blue circles represent its coverage area. Red dots indicate nonbaseline vehicles that have not yet been evaluated or verified as malicious. Green nodes indicate sensors that can be evaluated by the above baseline vehicles. The number of sensors covered by UAV and baseline vehicles increases the most in the first flight.

Figure 18 illustrates the trajectory of UAV's second flight. In this stage, the UAV is dispatched to collect sensors which have the largest number of nonbaseline MVS reports. The purpose is to evaluate nonbaseline vehicles that have unverified, and it is capable of providing effective sensor data if verified as credible. The second flight increases the number of baseline vehicles furtherly, which are generally located in low-density areas of the city. However, there are still some nonbaseline vehicles after multiple flights, which provide false data and have unverified, as the red dots shown in Figure 18. Sensors covered by these malicious vehicles are not credible and need to be recollected. Moreover, there are some sensors that are still not being covered by any MVs. Therefore, a third flight of the UAV is required to achieve coverage of all sensors in the city.

Figure 19 shows the trajectory of the UAV's third flight. In this stage, UAV will go to visit the nodes covered by the malicious device, as well as the nodes that are not within the coverage of the baseline vehicle, thus achieving full coverage of the whole city.

5.7. The Impact of the Number of Core Sensors on the Performance. Following simulation explains how the number of core sensors impacts on the efficiency of the trust verifications of MVs and sensors and on the number of flight rounds and the cost of UAV. We only consider the first two flight rounds of UAV when analyzing the impact of the number of core sensors on evaluating MVs, because UAV evaluates MVs at just the first two rounds in the simulation. Figure 20 shows that with the increasing number of core sensors, the identification rate of UAV on MVs gets decreasing at the first round and gets increasing first then becomes decreasing at the second round. In the simulation, the number of the recruited MVs is about 200. When the number of

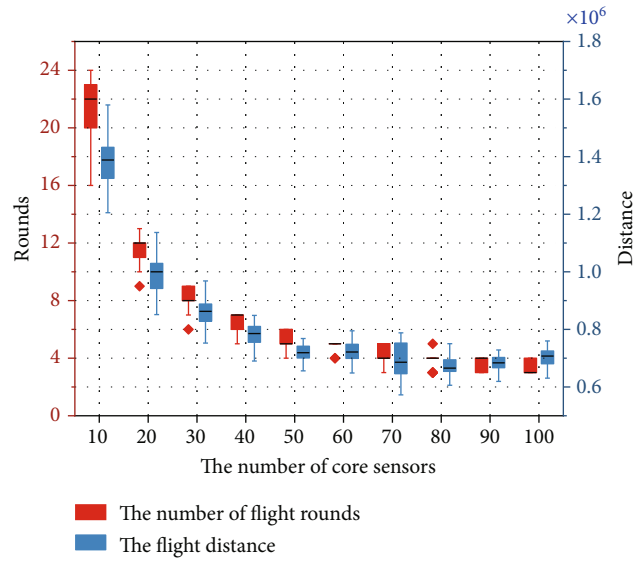

Figure 22: The impact of the number of core sensors on the flight cost of UAVs.

core sensors is set to a small number (smaller than 80), e.g., 60, the UAV covers about 130 MVs at the first round, thus leaving 70 MVs to be covered by UAV at the second round, which makes the identification rate of UAV high at that round. On the other hand, when the number of core sensors is set to a large number, e.g., 100, the UAV can cover about 160 MVs at the first round with the help of core sensors, and the identification rate of UAV at the second round is definitely low because there is only 40 MVs left after the first round.

Figure 21 indicates that the impact of the number of core sensors on the identification of sensors is similar to that of MVs shown in Figure 20. That is, the increasing number of core sensors also helps UAV identify sensors more efficiently. The only difference between them is that the UAV visits only sensors without covering by UAVs after the first two rounds.

From Figure 22, we see a similar result that the decreasing number of core sensors leads to the increasing number of the flight rounds of UAV. UAV has to flight more rounds to cover all the sensors without enough help from core sensors. It is natural to suppose that the cost of UAV is proportionate to its flight distance, and the flight distance of UAV is related to the number of its flight rounds. Therefore, the cost of UAV will decrease with the increasing number of core sensors. For example, in the simulation, UAV has to flight 10 rounds to cover all sensors when setting the number of core sensors to 10 and to flight only 2 rounds when setting the number to 50. Though a single round flight distance of the UAV on 50 core sensors may be larger than that on 10 core sensors, the total flight distance of UAV on 10 core sensors is much longer than the distance on 50 core sensors. That is, the flight cost when there are 10 core sensors is larger than that when there are 50 core sensors.

Since active trust depends on UAV and baseline MVs, it is particularly important in the early stages of trust verification. UAV verifies active trust by accessing the core sensors

perceived during flight, which means that the more core sensors there are, the more active verification times there are, and therefore, the faster active trust update. The above experimental results show that with the increase of the number of core sensors, the verification speed of active trust improves slightly. However, it is worth noting that more core sensors will also lead to higher flight costs. To sum up, the number of core sensors with better performance is 80 in this experiment.

6. Conclusions

This paper proposes an efficient data collection system, including a layered trust evaluation mechanism consisting of sensor nodes, MVs and UAVs, and a reasonable recruitment mechanism for MVs. The data collection system reaches a balance between cost and performance of data collection, which makes a reasonable cost while improving the accuracy and efficiency of data collection. Moreover, we show that UAVs can play an important role in designing a trust evaluation mechanism for IoTs. There are some other properties that need to be further considered, for example, whether the layered trust assessment mechanism aggravates the energy consumption to sensor node. It is also interesting to investigate the privacy-preserving methods in the mechanism during data gathering, which we plan to deal with by introducing blockchain for future work.

Data Availability

I have not included a data availability statement in my manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

This work was supported in part by the National Natural Science Foundation under Grant 61902432, Natural Science Foundation of Hunan Province under Grant 2020JJ4949, Excellent Youth Project of Scientific Research of Hunan Provincial Education Department under Grant 19B604, Talent Introduction Project of Central South University of Forestry and Technology under Grant 2020YJ016, and Course Construction Project of Central South University of Forestry and Technology under Grant A11030010.

References

[1] Z. Ning, P. Dong, X. Wang et al., "Mobile edge computing enabled 5G health monitoring for Internet of Medical Things: a decentralized game theoretic approach," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 463–478, 2021.

[2] F. Li, G. Huang, Q. Yang, and M. Xie, "Adaptive contention window MAC protocol in a global view for emerging trends networks," *IEEE Access*, vol. 9, no. 1, pp. 18402–18423, 2021.

[3] C. Huang, G. Huang, W. Liu, R. Wang, and M. Xie, "A parallel joint optimized relay selection protocol for wake-up radio enabled WSNs," *Physical Communication*, vol. 47, article 101320, 2021.

[4] M. Yu, A. Liu, N. Xiong, and T. Wang, "An intelligent game based offloading scheme for maximizing benefits of IoT-edge-cloud ecosystems," *IEEE Internet of Things Journal*, no. 99, pp. 1–1, 2020.

[5] Z. Ning, E. Ngai, R. Y. Kwok, and M. S. Obaidat, "Special Section on Pervasive Edge Computing for Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 5010-5011, 2021.

[6] T. Li, A. Liu, N. N. Xiong, S. Zhang, and T. Wang, "A trustworthiness-based vehicular recruitment scheme for information collections in distributed networked systems," *Information Science*, vol. 545, pp. 65–81, 2021.

[7] H. Teng, K. Ota, A. Liu, T. Wang, and S. Zhang, "Vehicles joint UAVs to acquire and analyze data for topology discovery in large-scale IoT systems," *Peer-to-Peer Networking and Applications*, vol. 13, no. 5, pp. 1720–1743, 2020.

[8] Y. Ren, A. Liu, M. Zhao, C. Huang, and T. Wang, "Quality utilization aware based data gathering for vehicular communication networks," *Wireless Communications and Mobile Computing*, vol. 2018, Article ID 6353714, 25 pages, 2018.

[9] S. Abdelhamid, H. S. Hassanein, and G. Takahara, "Reputation-aware, trajectory-based recruitment of smart vehicles for public sensing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1387–1400, 2017.

[10] L. Sánchez-Casado, G. Maciá-Fernández, P. García-Teodoro, and R. Magán-Carrión, "A model of data forwarding in MANETs for lightweight detection of malicious packet dropping," *Computer Networks*, vol. 87, pp. 44–58.

[11] Y. Ren, Z. Zeng, T. Wang, S. Zhang, and G. Zhi, "A trust-based minimum cost and quality aware data collection scheme in P2P network," *Peer-to-Peer Networking and Applications*, vol. 13, no. 6, pp. 2300–2323, 2020.

[12] J. Jiang, G. Han, F. Wang, L. Shu, and M. Guizani, "An efficient distributed trust model for wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 5, pp. 1228–1237, 2015.

[13] Z. L. Ning, K. Y. Zhang, X. J. Wang et al., "Intelligent edge computing in internet of vehicles: a joint computation offloading and caching solution," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2212–2225, 2021.

[14] Z. L. Ning, X. P. Hu, Z. K. Chen et al., "Corrections to "A cooperative quality-aware service access system for social Internet of Vehicles"," *IEEE Internet Things*, vol. 7, no. 7, article 6663, 2020.

[15] Z. L. Ning, Y. Li, P. Dong et al., "When deep reinforcement learning meets 5G-enabled vehicular networks: a distributed offloading framework for traffic big data," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1352–1361, 2020.

[16] I. Anati, S. Gueron, S. Johnson, and V. R. Scarlata, "Innovative technology for CPU based attestation and sealing," May 2017, https://software.intel.com/en-us/articles/innovative-technology-for-cpu-based-attestation-and-sealing.

[17] M. Hoekstra, R. Lal, P. Pappachan, C. Rozas, V. Phegade, and J. D. Cuvillo, "Using innovative instructions to create trustworthy software solutions," May 2017, https://software.intel.com/en-us/articles/using-innovative-instructions-to-create-trustworthy-software-solutions.

[18] Intel, "Intel 64 and IA-32 architectures software developer manuals," May 2018, https://software.intel.com/en-us/articles/intel-sdm.

[19] AMD, "AMD secure technology," https://www.amd.com/en-us/innovations/software-technologies/security.

[20] M. Bonola, L. Bracciale, P. Loreti, R. Amici, A. Rabuffi, and G. Bianchi, "Opportunistic communication in smart city: Experimental insight with small- scale taxi fleets as data carriers," *Ad Hoc Networks*, vol. 43, pp. 43–55, 2016.

[21] Z. He, J. Cao, and X. Liu, "High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, pp. 2542–2550, Kowloon, Hong Kong, 2015.

[22] T. Wang, H. Luo, W. Jia, A. Liu, and M. Xie, "MTES: an intelligent trust evaluation scheme in sensor-cloud-enabled industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 2054–2062, 2020.

[23] T. Wang and H. Luo, "Crowdsourcing mechanism for trust evaluation in CPCS based on intelligent mobile edge computing," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, pp. 1–19, 2019.

[24] L. Hu, A. Liu, M. Xie, and T. Wang, "UAVs joint vehicles as data mules for fast codes dissemination for edge networking in Smart City," *Peer-to-Peer Networking and Applications*, vol. 12, no. 6, pp. 1550–1574, 2019.

[25] T. Tanaka, T. Suzuki, and K. Kurihara, "Energy harvesting technology for maintenance-free sensors," *Fujitsu Scientific & Technical Journal*, vol. 50, pp. 93–100, 2014.

[26] L. Xu and Y. Zhang, "A New reputation-based trust management strategy for clustered ad hoc networks," in *2009 International Conference on Networks Security, Wireless Communications and Trusted Computing*, pp. 116–119, Wuhan, Hubei, 2009.

[27] S. Sharma, R. Mishra, and I. Kaur, "New trust based security approach for ad-hoc networks," in *2010 3rd International Conference on Computer Science and Information Technology*, pp. 428–431, Chengdu, 2010.

[28] J. Wang, X. Li, and Y. Zhang, "Research of P2P network trust model," in *2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics*, pp. 70–73, Hangzhou, 2013.

[29] B. Jiang, G. Huang, T. Wang, J. Gui, and X. Zhu, "Trust based energy efficient data collection with unmanned aerial vehicle in edge network," *Transactions on Emerging Telecommunications Technologies*, pp. 1–32, 2020.

[30] T. Li, W. Liu, T. Wang, Z. Ming, X. Li, and M. Ma, "Trust data collections via vehicles joint with unmanned aerial vehicles in the smart Internet of Things," *Transactions on Emerging Telecommunications Technologies*, vol. 7, pp. 1–24, 2020.

[31] Y. Liu, M. Dong, K. Ota, and A. Liu, "Activetrust: secure and trustable routing in wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 2013–2027, 2016.

[32] M. Shen, A. Liu, G. Huang, N. N. Xiong, and H. Lu, "ATTDC: an active and traceable trust data collection scheme for industrial security in smart cities," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6437–6453, 2021.

[33] W. Mo, T. Wang, S. Zhang, and J. Zhang, "An active and verifiable trust evaluation approach for edge computing," *Journal of Cloud Computing*, vol. 9, no. 1, 2020.

[34] G. Zhang, T. Wang, G. Wang, A. Liu, and W. Jia, "Detection of hidden data attacks combined fog computing and trust evaluation method in sensor-cloud system," *Concurrency and Computation-Practice & Experience*, Article ID cpe.5109, 2018.