Using UAV to Detect Truth for Clean Data Collection in Sensor-Cloud Systems

LI Xiuxian[1], LI Zhetao[1], OUYANG Yan[2], DUAN Haohua[3], XIANG Liyao[3]

(1. Xiangtan University, Xiangtan 411100, China;
2. Central South University, Changsha 410000, China;
3. Shanghai Jiao Tong University, Shanghai 200000, China)

Abstract: Mobile edge users (MEUs) collect data from sensor devices and report to cloud systems, which can facilitate numerous applications in sensor-cloud systems (SCS). However, because there is no effective way to access the ground truth to verify the quality of sensing devices' data or MEUs' reports, malicious sensing devices or MEUs may report false data and cause damage to the platform. It is critical for selecting sensing devices and MEUs to report truthful data. To tackle this challenge, a novel scheme that uses unmanned aerial vehicles (UAV) to detect the truth of sensing devices and MEUs (UAV-DT) is proposed to construct a clean data collection platform for SCS. In the UAV-DT scheme, the UAV delivers check codes to sensor devices and requires them to provide routes to the specified destination node. Then, the UAV flies along the path that enables maximal truth detection and collects the information of the sensing devices forwarding data packets to the cloud during this period. The information collected by the UAV will be checked in two aspects to verify the credibility of the sensor devices. The first is to check whether there is an abnormality in the received and sent data packets of the sensing devices and an evaluation of the degree of trust is given; the second is to compare the data packets submitted by the sensing devices to MEUs with the data packets submitted by the MEUs to the platform to verify the credibility of MEUs. Then, based on the verified trust value, an incentive mechanism is proposed to select credible MEUs for data collection, so as to create a clean data collection sensor-cloud network. The simulation results show that the proposed UAV-DT scheme can identify the trust of sensing devices and MEUs well. As a result, the proportion of clean data collected is greatly improved.

Keywords: sensor-cloud system; truth detection; trust reasoning and evolution; mobile edge user; unmanned aerial vehicle

DOI: 10.12142/ZTECOM.202103005

https://kns.cnki.net/kcms/detail/34.1294.TN.20210818.1538.005.html, published online August 18, 2021

Manuscript received: 2021-06-08

Citation (IEEE Format): X. X. Li, Z. T. Li, Y. Ouyang, et al., "Using UAV to detect truth for clean data collection in sensor-cloud systems," *ZTE Communications*, vol. 19, no. 3, pp. 30 – 45, Sept. 2021. doi: 10.12142/ZTECOM.202103005.

1 Introduction

With the development of techniques on microprocessor industry, sensing-based devices are becoming smaller while their computation and capacities are strengthened gradually[1 – 3]. Therefore, sensing technologies are widely deployed in areas with on-demand monitoring processes. According to a survey, there were more than 20 billion devices connected to the Internet of Thing (IoT) in 2020 and the number is growing at a faster rate[4 – 6]. These IoT devices are equipped with numerous sensing devices to realize the perception of the surroundings[9 – 10]. Thus, the sensor-cloud systems (SCS), within which the IoT devices and cloud services are well combined, can be more productive and effective on its functionality and solve such problems as the

This work was supported in part by National Natural Science Foundation of China under Grant No. 62032020, Hunan Science and Technology Planning Project under Grant No.2019RS3019, and the National Key Research and Development Program of China under Grant 2018YFB1003702.

sharing of sensor nodes and large amounts of data analysis due to memory and energy limitations[9]. In an SCS, a huge number of sensing devices are deployed at the edge of the network to sense the surrounding environment[11 - 13], and then upload the sensed data to the cloud. Due to the excellent computing power, cloud services can perform sophisticated computation and analytics, as well as orchestrate various applications. For example, the supervisory control and data acquisition (SCADA) system is one of the SCS and composes of smart sensing devices spreading over a wide area in order to remotely monitor physical phenomena[14]. These smart sensing devices can be deployed on demand in the areas that require temporary testing, and then collect data into the cloud in various ways to initiate and build up various applications[15 - 16]. The method of data collection has also changed a lot from the traditional methods in the past. In traditional wireless sensor networks (WSNs), many nodes are deployed in specific areas and self-organize into a network. The sensed data is routed to a specific node called sink through multi-hop routing[17 - 18] and the sink is connected to the Internet by a wired network; in this way, the data are reported to the cloud. However, the time and economic cost of deploying the network to establish the connection with the sink will be relatively high, so this system is hardly used on some scenarios such as urgent events and scenarios without complete infrastructure. Thus, many researchers have proposed more flexible and convenient data collection schemes. For example, BONOLA et al. [19] proposed a method of data collection using opportunistic routing through mobile vehicles (MVs), and in this way, the roadside is deployed with sensing devices to monitor the status of street lights, smart trash cans, and roads and bridges on demand. With this solution, the sensing hardware will be simple, only a short-distance wireless communication capability be required, and installing expensive 5G communication hardware be not necessary[19]. The reason is that, in a smart city, there are a large number of MVs moving on the roads of the city, and when the MVs pass through the communication range of sensing devices, they can collect data and transmit the data to the cloud through 5G communications. In the research of HUANG et al. [20], numerous deployed sensor nodes can also self-organize into a network; the nodes on both sides of a road act as gateways, which are responsible for converging the entire network, and pass data to the cloud through MVs[21]. Therefore, this method may be widely used in smart cities. More related studies have been conducted[22 - 24]. In fact, except MVs[25 - 27], smartphones, tablets and smart watches can also act as data collectors[28]. They are called mobile edge users (MEUs) in the research of WANG et al. [28]. Because these MEUs have 5G communication capabilities, they can communicate directly with the cloud. The MVs are only on the road, but there are multiple types of MEUs in the market[28] with a wider moving range. When these MEUs pass through sensing devices with weak communication capabilities, they can collect data from sensing devices within their communication range and relay the data to the cloud. The use of MVs for data collection[19] is also a form of data collection approaches using MEUs. Therefore, in this paper, MEU is the general term for the devices that have 5G communication capabilities to perform data collection in a relay mode, and sensing devices or sensing nodes refer to a type of simple hardware that can only communicate over a short distance and needs to rely on MEUs to relay data to the cloud.

In order to incentivize MEUs to collect data, the incentive mechanism[29] is widely used, which enables cloud to initiate data collection tasks, grant a reward for collecting data, and incentivize MEUs to collect data[29]. This mechanism simplifies the deployment requirements of sensing devices and many sensor devices can be deployed on demand without 5G communication capabilities. Therefore, it facilitates a dramatic cost reduction of numerous sensor devices[30]. Moreover, the data collected by these sensing devices will be reported to the cloud through a huge number of MEUs, rather than specifically deploying a network for device connection. Such a system based on the incentive mechanism has strong adaptability and has been widely studied and used.

However, in such applications, the pivotal point is how to ensure the security of data collection. The factors affecting the security of data collection mainly come from MEUs and sensing devices[28, 30]. The impact of MEUs on the security of data collection is mainly manifested in some MVs reporting false data in order to obtain rewards, and there are even some malicious MVs that deliberately report offensive data, making data-based applications unusable[20 - 22]. For sensing devices, due to their simple hardware design, they are vulnerable to face attacks. Once these sensing devices are attacked, various problems will occur. For example, a black hole drops the data packets that are passing through it, so that the cloud platform cannot receive data[30]. According to statistics, there are more than 30 types of attacks on the sensing network, and these attacks will generate false data, tamper with data, or block the collection of data to damage the network[30]. Therefore, how to create a safe and clean environment of data collection as well as collecting authentic and credible data is a challenge deserved to concern with.

Although the use of MEUs is a cost-effective method[20], it is more challenging to ensure the security of data collection in such a data collection mode. In addition to the inherent unsafe factors in sensing devices, the use of MEUs for data collection may bring more threatening factors[15]. In particular, MEUs participate in data collection voluntarily with no identification, so it is difficult to ensure that MEUs are trustworthy in such an open-ended network environment[24]. What is more serious is that it is incredibly hard to verify whether the data reported by MEUs are true, which is known as an information elicitation without verification (IEWV) problem[31]. Due to the IEWV problem, even if the MV report reports false data in order to obtain rewards, it is difficult to verify the data.

Using a credibility mechanism to choose trustable MVs for data collection is a feasible method. Because credible MVs will truthfully report the collected data, selecting credible MVs for data collection can improve the authenticity of the data[9, 15]. However, as mentioned earlier, it is difficult to verify the authenticity of data reported by MVs[32]; similarly, it is also difficult to identify the trustworthiness of MVs. In addition, for the sensing network, it is a major challenge to identify the credibility of these sensing devices[22, 24]. For a sensing network far from the edge of the network, it is very difficult to detect data attack[32]. Thus, we make the first attempt to deal with this challenge. In this paper, we propose a novel scheme that uses unmanned aerial vehicles (UAV) to detect the trust of sensing devices and MEUs (UAV-DT) to construct a clean data collection platform for SCS. The main contributions of this article are as follows:

1) We propose a framework using UAV to detect the trust of sensing devices and MEUs. In the proposed framework, the UAV is sent to the sensing network, deliver check codes to some selected sensing devices, and is required to route the code to the designated destination node. At the same time, the UAV collects information about data packets sent from sensing devices within a time span when passing through the sensing network. In this scheme, the checking code can act as a base truth indicator. If the UAV or cloud does not receive the verification code on the time that it should receive it, it can indicate that the verification code has been attacked during the data collection process. In this way, the IEWV problem that exists in this type of network can be effectively solved.

2) We propose an effective approach to sensing devices and MVs credibility computation. This method can construct a trusted data collection network environment. The information collected by the UAV will be checked by the cloud platform in two aspects to verify the credibility of sensor devices. On the one hand, the platform will check whether there is an anomaly in the data packet routing process for trust evaluation. It mainly checks whether the upstream and downstream nodes of the sensing devices receive and send data packets abnormally and therefore provide a performance evaluation about the trustworthiness. Besides, the data packets submitted by the sensing devices to the MEUs and those submitted by the MEUs to the platform will be checked and compared to verify the trustworthiness of the MEUs. On the other hand, according to the designed routing path of the verification code, it checks whether the verification code is successfully routed from the originated nodes to the MVs, and then submits the message to the cloud, which improves the trust of these sensing devices and MVs. The credibility computation method proposed in this paper enables accurate verification.

3) Based on the proposed framework, we propose a data collection strategy based on credibility magnitude and incentivation mechanism. The simulation results show that the proposed UAV-DT scheme can identify the credibility magnitude of sensing devices and MEUs, and the amount of clean data collected has been proportionally incremented. The classification rate for trusted sensing devices is as high as 98.9%. Meanwhile, a data collection rate of 89.9% on average can be achieved.

2 Related Work

To protect the security of networks in smart cities, various safety mechanisms, e. g., cryptographic schemes, authentication mechanisms and secure storage, were proposed in the past. However, using a trust-based model, the trust evaluation mechanism has the advantages of efficiency, lightweight and low overheads. The trust models that have been proposed provide a better choice in terms of network security and safety.

In general, the trust-based evaluation mechanisms for network security can be classified into two categories: the centralized and distributed. For the former, the trust value of nodes can be calculated by themselves. KIM and SEO[33] have proposed a trust computation method using fuzzy logic (TCFL) for WSN. They suggest a trust model using fuzzy logic in sensor network, in which trust is an aggregation of consensus given a set of past interaction among sensors. They calculate the trust value of the path through the trust of the nodes, then the path with the highest trust value is selected to transmit data packets[33]. However, in the great majority of applications, smart network system is distributed with a large number of nodes and a node in the system only focuses on the trustworthiness of its neighbor nodes. Besides, centralized approaches always make high energy consumption.

A distributed mechanism, the beta-based trust and reputation evaluation system for wireless sensor networks (BTRES), is proposed in Ref. [34]. BTRES is based on monitoring nodes’ behavior and beta distribution is used to describe the distribution of nodes’ credibility. Another distributed trust computation scheme, the parameterized and localized trust management scheme (PLUS), is proposed by YAO et al.[35]. In PLUS, each sensor node maintains highly abstracted parameters, and rates the trustworthiness of its interested neighbors to adopt appropriate cryptographic methods, identifing the malicious nodes and sharing the opinion locally. Distributed mechanisms have obvious disadvantages as well, which include the excess energy node and time costs due to the cooperation and communication with neighbors and increasing memory costs with the increase of network density caused by the lack of centralized management.

To overcome the defects above, WANG et al.[28] propose a crowdsourcing mechanism for trust evaluation based on mobile edge computing. In this mechanism, through close access to end nodes, mobile edge users can obtain various types of information of the end nodes and determine whether the node is trustworthy. HUANG et al.[20] propose a novel baseline data based verifiable trust evaluation scheme, called BD-VTE, similar to the scheme in Ref. [28]. In BD-VTE, the trust of MVs is evaluated by sending UAVs to perceive IoT devices data as baseline data.

3 System Model and Problem Statement

3.1 System Model

Fig. 1 shows the SCS network model used in this paper. Our model includes sensing devices, MEUs and UAVs. The following is the description and symbol definition of each role.

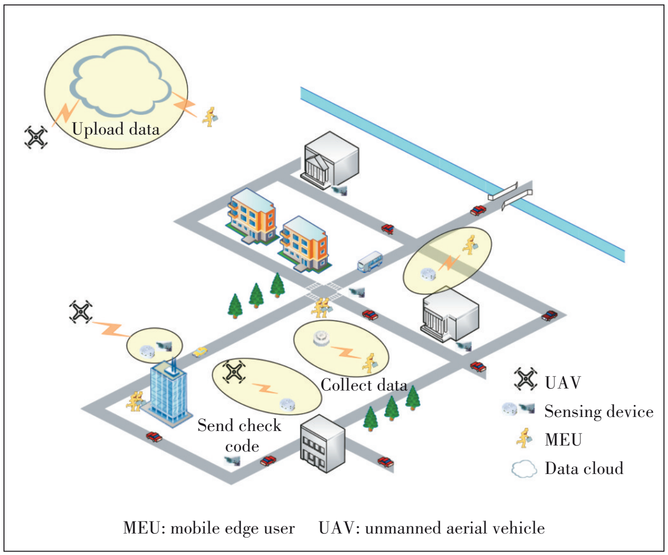
1) Sensing devices

As shown in Fig. 1, the SCS, the IoT devices are treated as sensor nodes and constitute the sensor network. There are N sensor nodes deployed in the network. The set of nodes is represented by $V = \{1,2,...,N\}$, and the sensor nodes perceive the environment in the city, output data and transmit them to the outside. A small number of nodes may be attacked and become malicious nodes, which is manifested as deliberate packet loss during data transmission. Suppose the number of malicious nodes is K and the set of malicious nodes is represented by $M = \{1,2,...,K\}$.

2) Unmanned aerial vehicles

The role of the UAVs is a bridge between the sensor network and the external network and they can communicate directly with the data center in the cloud. The UAVs can distribute the verification packets generated by the data center to the sensor nodes for transmission and directly check the transmission status of the nodes. In the scenario shown in Fig. 1, the UAVs pass the verification packet with a check code to a starting node, and then the node transmits it according to a certain routing rule.

3) Mobile edge users



▲Figure 1. Sensor-cloud system model

The number of MEUs with strong communication and storage capabilities distributed in the city far exceeds sensing devices. The MEU acts as a data collector in the system model and can directly communicate with the sensor node to obtain the data packet transmitted to the node. There is a total of L MEUs in the system, and their set is represented by $U = \{1,2,...,L\}$. Each MEU has its own active range, which is abstracted as a circle whose radius is r_i, and the abstract model of MEU is widely used by many researchers[28－29]. Within a certain period of time, the MEU can collect the data of all nodes covered in its active range, which means that the active range of the MEU indirectly refers to its ability to perform data collection tasks.

4) Transmission model

Considering communications between sender node n_1 and receiver node n_2, let p_{n_1} denote the transmitting power of n_1, and h_{n_1,n_2} denote the channel gain between n_1 and n_2. The channel gain follows the Rayleigh distribution. The distance between n_1 and n_2 is denoted by d_{n_1,n_2}, and the channel attenuation factor and Gaussian channel coefficient are donated by ϑ and h_0, respectively. Therefore, the channel gain holds as:

$$h_{n_1,n_2} = h_0 d^{-\vartheta}{}_{n_1,n_2} . \tag{1}$$

And the transmission rate between the sender node n_1 and receiver node n_2 can be denoted according to Shannon equation:

$$r_{n_1,n_2} = B\mathrm{logs}_2\left(1 + \frac{p_{n_1} \times h_{n_1,n_2}}{p_0 + N_0}\right) , \tag{2}$$

where B denotes the bandwidth, N_0 denotes the power spectral density of additive Gaussian white noise, and p_0 denotes the interference caused by reusing identical spectrum resources.

3.2 Problem Statement and Relevant Definition

The previous study has shown that using MEU can infer the trust value of the node according to its various states, e.g., the communication behavior, remaining battery, data content of target node, and so on[28]. In practice, it is difficult to obtain such information directly through the MEU, while obtaining data indirectly by monitoring neighbor nodes will add additional communication burden to each node, which will greatly reduce the life of the entire network. Due to the above limitations, it is unrealistic to directly or indirectly obtain the status of a node. Another problem that needs to be solved is the lack of an effective mechanism to ensure the authenticity of the data uploaded by MEUs. Therefore, we need to distinguish trusted nodes from malicious nodes in the network in an effective and realistic way. In general, when the nodes in the sensor network transmit data packets, trusted nodes can complete the data transmission task well. Occasionally, packet loss will occur when the network fluctuates greatly and the integrity of the data will not change significantly. However, malicious nodes will frequently drop packets or tamper with data, which will compromise the validity of the data. Meanwhile, as the third-party data collector, the credibility of MEU also needs investigating. It is also necessary to distinguish between trusted and malicious MEUs and hire trusted users to complete data collection tasks, thereby ensuring the quality of the collected data by MEUs. Thus, in this paper, the MEU that plays the role of data collector is regarded as a mobile node and it is referred to as a node with sensing devices when there is no special distinction. We also need to minimize the cost of evaluating and classifying nodes. Hiring MEUs for data collection is the main cost of the system. Therefore, the data center should adopt an efficient MEU incentive mechanism to hire a set of trusted MEUs to complete data collection tasks with high quality. This paper reflects the performance of system through the following trust indicators and overall costs:

1) Difference of trust values between normal and malicious nodes D, which is defined as

$$D = \overline{\sigma}_{\text{nor}} - \overline{\sigma}_{\text{mal}}\ , \tag{3}$$

where $\overline{\sigma}_{\text{nor}}$ is the average trust value of normal nodes and $\overline{\sigma}_{\text{mal}}$ is the average trust value of malicious nodes. The difference D between the two averages can show the difference of benefits to the network. When D is large, it means that the distinction between the two is obvious. Therefore, one of the goals of our strategy is $\text{Max}(D) = \max(\overline{\sigma}_{\text{nor}} - \overline{\sigma}_{\text{mal}})$.

2) The discrimination rate of trusted nodes $\mathcal{R}_t$ and discrimination rate of malicious nodes $\mathcal{R}_m$ are defined as

$$\mathcal{R}_t = \frac{num_{\bar{t}}}{num_t}\ , \tag{4}$$

$$\mathcal{R}_m = \frac{num_{\bar{m}}}{num_m}\ . \tag{5}$$

These two indicators refer to the ratio of the correct number of nodes judged to be trusted $num_{\bar{t}}$ and the total number of trusted nodes num_t, and the ratio of the correct number of nodes judged to be malicious $num_{\bar{m}}$ and the total number of malicious nodes num_m. Both $\mathcal{R}_t$ and $\mathcal{R}_m$ reflect the system's ability to classify nodes. Then, the goals of our strategy include $\text{Max}\left(\mathcal{R}_t\right) = \max\left(\frac{num_{\bar{t}}}{num_t}\right)$ and $\text{Max}\left(\mathcal{R}_m\right) = \max\left(\frac{num_{\bar{m}}}{num_m}\right)$ as well.

3) Total cost of system P is defined as

$$P = \sum_{i=1}^{R} \sum_{j=1}^{L} f_{i,j} \times p_{i,j} + \sum_{i=1}^{R} \mathcal{L}_i \times v\ , \tag{6}$$

where $f_{i,j}$ indicates whether the user labeled j in the data collector set U in the ith round participates in the data collection task; $f_{i,j} = 1$ indicates that the user participated in the data collection task, otherwise $f_{i,j} = 0$ means not; $p_{i,j}$ represents the remuneration received by the user labeled as j in the data collector set U in the ith round; $\mathcal{L}_i$ represents the number of nodes that need UAVs to inspect in the ith round and v represents the cost of UAV verification of one node. Therefore, one of the purposes of our strategy is $\text{Min}(P) = \min\left(\sum_{i=1}^{R} \sum_{j=1}^{L} f_{i,j} \times p_{i,j} + \sum_{i=1}^{R} \mathcal{L}_i \times v\right)$.

In summary, all objectives in this paper are shown in Eqs. (7) – (10).

$$\text{Max}(D) = \max\left(\overline{\sigma}_{\text{nor}} - \overline{\sigma}_{\text{mal}}\right), \tag{7}$$

$$\text{Max}\left(\mathcal{R}_t\right) = \max\left(\frac{num_{\bar{t}}}{num_t}\right), \tag{8}$$

$$\text{Max}\left(\mathcal{R}_m\right) = \max\left(\frac{num_{\bar{m}}}{num_m}\right), \tag{9}$$

$$\text{Min}(P) = \min\left(\sum_{i=1}^{R} \sum_{j=1}^{L} f_{i,j} {*} p_{i,j} + \sum_{i=1}^{R} \mathcal{L}_i {*} v\ .\right. \tag{10}$$

The notation of parameters in the model and for problem statement is shown in Table 1.

4 Proposed UAV-DT System

In this part, we present our UAV-DT scheme. The proposed scheme is divided into three parts: the UAV-assisted trust verification mechanism, trust reasoning mechanism based on communication behavior, and incentive mechanism based on cost performance and trust.

4.1 UAV-Assisted Trust Verification Mechanism

The most critical part of our scheme is to evaluate the trust value of nodes in the SCS, thereby distinguishing between trusted and malicious nodes. We propose a UAV-assisted trust verification mechanism to determine whether the communication behavior of the node is normal.

The mechanism is divided into four stages: generation and distribution of verification packets, result gathering of the tail node, review of data collection tasks, and verification of a suspect path. In each stage, UAVs play an important role.

At the beginning, the sensor network shown in Fig. 2(a) is untested. The following is an introduction to the four steps:

1) Generation and distribution of detection packets

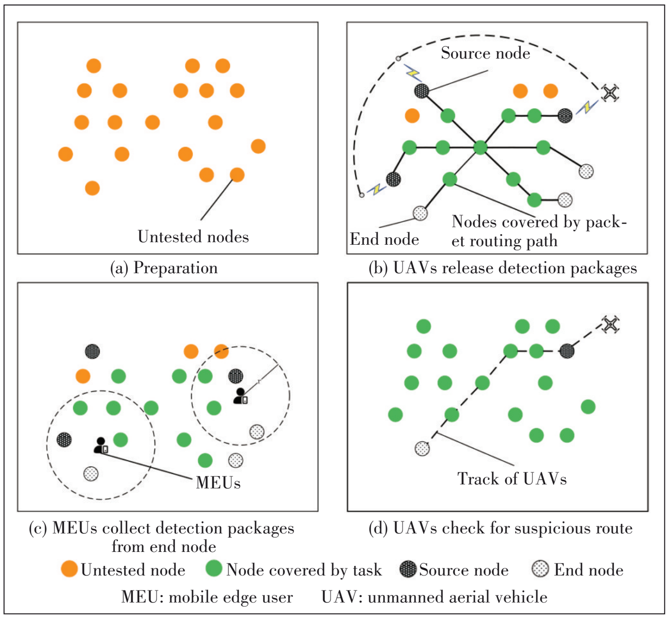
At this stage, the UAV selects a certain number of source nodes in the network as the start of data packet delivery. Our trust evaluation is conducted in multiple rounds, so the trust value of a node will change after each round of calculation. In this process, when the trust value of the node is higher than $\sigma_{\max}$, the node is judged to be trusted. On the contrary, when

▼**Table 1. Parameters in System Model and Problem Statement**

Parameter	Meaning
V	Collection of sensor nodes
U	Collection of mobile edge users
M	Collection of malicious nodes
T	Number of verification tasks per round
R	Number of rounds of a verification task
D	Difference of trust values between normal and malicious nodes
P	Total cost of system
$\overline{\sigma}_{\mathrm{nor}}$	Average trust of normal nodes
$\overline{\sigma}_{\mathrm{mal}}$	Average trust of malicious nodes
$\mathcal{R}_t$	Discrimination rate of trusted nodes
$\mathcal{R}_m$	Discrimination rate of malicious nodes
$num_{\bar{t}}$	Number of nodes judged to be trusted
num_t	Total number of trusted nodes
$num_{\bar{m}}$	Number of nodes judged to be malicious
num_m	Total number of malicious nodes
$f_{i,j}$	Participation flag for user labelled as j in the ith round
$B_{i,j}$	Result flag for node labelled as j in the ith task
$p_{i,j}$	Payment for user labelled as j in the ith round
$\mathcal{L}_i$	Number of nodes that need UAVs to inspect in the ith round
υ	Cost of UAV verification of one node
N	Number of sensor nodes
K	Number of malicious nodes
L	Number of mobile edge users
σ_0	Initial trust value
$\sigma_{\max},\sigma_{\min}$	Trust value threshold
σ_{com_i}	Communication trust value
$\sigma_{rec}(i,j)$	Cooperative recommendation coefficient between nodes labeled as i and j
σ_{rec_i}	Cooperative recommendation trust value
σ_{int_i}	Comprehensive trust value
bid_i	Bid of mobile edge user
b_i	Expected reward of mobile edge user
PoI_i	Number of task nodes in the active range
α_1,α_2	Weight coefficient of winning bids set selection algorithm
ω_1,ω_2	Weight coefficient of comprehensive trust

UAV: unmanned aerial vehicle

the trust value of the node is lower than σ_{min}, the node is considered untrusted. Therefore, our criterion for selecting a source node is the node whose trust value is between $\sigma_{\min}$ and $\sigma_{\max}$ in the suspicious state.



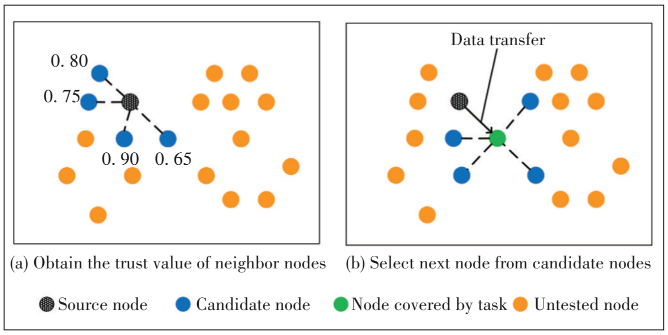
(a) Preparation

(b) UAVs release detection packages

(c) MEUs collect detection packages from end node

(d) UAVs check for suspicious route

▲**Figure 2. Using UAV to detect the trust of sensing devices and MEUs (UAV-DT)**



(a) Obtain the trust value of neighbor nodes

(b) Select next node from candidate nodes

▲**Figure 3. Route generation strategy**

As Fig. 2(b) shows, after a source node is selected, it is necessary to determine the route of the detection data packet transmission. We use the low-trust node diffusion strategy shown in Fig. 3 to generate the transmission path of the data packet. Starting from the source node, when determining the next hop node, the current trust value of each neighboring node is considered; the node that is closest to the initial trust value σ_0 is selected and the task is refused to repeat. It will be ensured that each node in the network receives a certain number of inspections.

The generation and distribution process of the verification data packet can be summarized in Algorithm 1.

Algorithm 1. Generation of detection packet algorithm

```
Input: $R$, $T$, $\sigma_{\max}$, $\sigma_{\min}$, $\sigma_0$, $V$
Output: $S$
Initialize $iter_{out} = 0$, $S = \varnothing$
While $iter_{out} < R$ Do
    $\rho = \varnothing$
    Randomly choose source node $n$ in $V$ and
    $\sigma_n < \sigma_{\max}$ and $\sigma_n > \sigma_{\min}$
    $\rho = \rho \cup n$
    $node_{cur} = n$
    $node_{source} = n$
    $iter_{in} = 0$
    While $iter_{in} < T$ Do
        Choose $nxt$ in $neb\left(node_{cur}\right)$
        $\min |\sigma_{nxt} - \sigma_0|$ and $nxt \notin \rho$
        $\rho = \rho \cup nxt$
```

```
12:            $node_{cur} = nxt$
13:            $iter_{in} = iter_{in} + 1$
14:        End While
15:        $node_{end} = node_{cur}$
16:        Generate check packet $p = (node_{source}, node_{end}, \rho)$
17:        $S = S \cup p$
18:        $iter_{out} = iter_{out} + 1$
19: End While
20: Return $S$
```

Algorithm 1 shows the process of multiple rounds of verification data packet generation and routing distribution. The input of the algorithm includes the number of verification task rounds R, the number of verification data packets in each round T, three constants related to trust value and node sets V. The outer loop (Lines 2 – 19) represents each round of verification tasks, and at the beginning of each mission, the system randomly chooses source node n in V. Then, the system generates a route for each verification task in the inner loop (Lines 9 – 14). Finally, the verification task set S is output.

2) Result gathering of the tail node

When the route of the verification packets is determined, the UAV distributes the data packets to the source node, and the packets are transmitted to the tail node in turn according to the routing path.

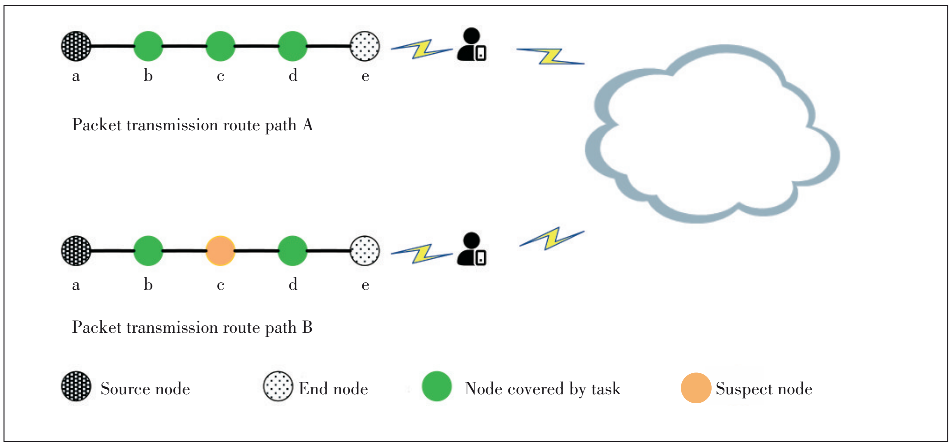
Then, we only need to collect the delivered data packets at the tail node to confirm whether there is any node loss behavior on the delivery path of the data packets. As shown in Fig. 2(c), MEUs are hired as data collectors to perform data collection tasks at the end nodes and hand data over to the data center in the cloud for further processing.

3) Review of data collection tasks

Since the data collector is not necessarily credible, we still need to further verify the validity and completeness of the data collector's collection results.

Obviously, not all collected results need to be verified. When issuing data collection tasks, the system clarifies which data packets at the sensing devices need to be collected, and the data collector does not know the content of the data packet and its check code in advance. Therefore, in the case that the verification packet is normally delivered to the tail node, we can consider that the result of this collection must be credible and no further confirmation is required if the submitted by the data collector is consistent with the original packet and is accompanied by a true verification code.

If a data package uploaded by the data collector is inconsistent with the original package distributed by the UAV, it is necessary to rely on the UAV to recollect the data at the tail node to make a judgment on the communication behavior between the data collector and the sensing devices on the routing path. The UAV compares the original verification packet with the data collected by itself to determine whether the data collector has performed the data collection task honestly, or the verification packet has been modified by one or more malicious nodes during the transmission process.


a
b
c
d
e
Packet transmission route path A
a
b
c
d
e
Packet transmission route path B
Source node
End node
Node covered by task
Suspect node

▲Figure 4. Successful transmission and suspect paths

4) Verification of a suspect path

Through the previous stage, the system knows which verification packets have been modified during the transmission process, and defines such a routing path as a suspected path; that is, malicious nodes may appear on this path. In our system, the UAV checks the communication records of each node along the path, and judges the node(s) where the packet loss has occurred in the transmission process based on these records.

Then, we specifically describe the processing of nodes on the successful transmission path and the verification of suspected path (Fig. 4).

As shown in Fig. 4, in route path A, the verification packet is successfully transmitted to the tail node, and then collected by the MEUs faithfully, which means a successful transmission path. In this case, all the nodes on the transmission path honestly transmit the verification data packet to the next hop node. The system records the successful communication behavior once for Nodes a, b, c, and d. Since Node e is the tail node, it does not participate in the transmission process, so this verification task cannot perform trust evaluation on it.

On the contrary, after system verification in route path B, the verification data packet collected at the tail node has changed compared with the original data packet, which indicates that there is packet loss behavior by one or more nodes. In the figure, the node marked in orange is the node that has lost packets during transmission. Considering that there are network fluctuations, trusted nodes may also lose packets due to poor network communications, so the system cannot directly determine whether the node that has lost the packet during the data packet transmission is a malicious node, but can only define its communication behavior this time is malicious. According to the transmission result of the data packet, the system records the successful communication behavior for Nodes a, b, and d once, and correspondingly, it records the malicious communication behavior for Node c once. As mentioned above, Node e is the tail node and does not participate in the transmission of data packets.

When the MEU collects data packets at the tail node and uploads them to the data center in the cloud, it may also misrepresent the data. The system will also check the communication behavior of its uploaded data, and record the number of honest uploads and false uploads. As shown in Fig. 5, the real active range of an MEU is a light area with a radius of R_1, but it lies to the cloud data center that its active range is the dark area with a radius of R_2. Then the system assigns data collection tasks at four tail nodes, but the MEU only completes three collection tasks. The data at the node at the bottom right is not collected by the MEU because it exceeds its active range and at the same time it lies about a false result. Based on the above, the system will record the honest upload and false upload of the MEU.

After the above four steps finish, one round of a verification task is completed. With the obtained communication behaviors of the nodes and data collector, we can use various trust evaluation methods to calculate their trust values. The next section will focus on describing the trust reasoning mechanism used in our scheme.

4.2 Trust Reasoning Mechanism Based on Communication Behavior

The proposed trust reasoning mechanism in this paper is divided into three parts: the trust value initialization, trust evaluation, and trust state determination (Fig. 6).

4.2.1 Trust Value Initialization

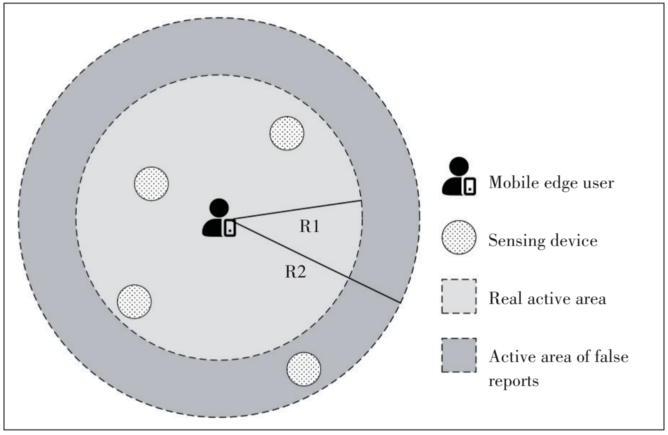
At the beginning, when all nodes and MEUs have not been fully checked, we cannot judge whether they are malicious or not, so we give each node the same initial trust value σ_0 and record it as a suspect state. After multiple rounds of trust inspection, the trust value of the nodes or MEUs participating in the transmission and collection of verification data packets will change through the trust reasoning mechanism and their status will increase or decrease accordingly.

4.2.2 Trust Evaluation

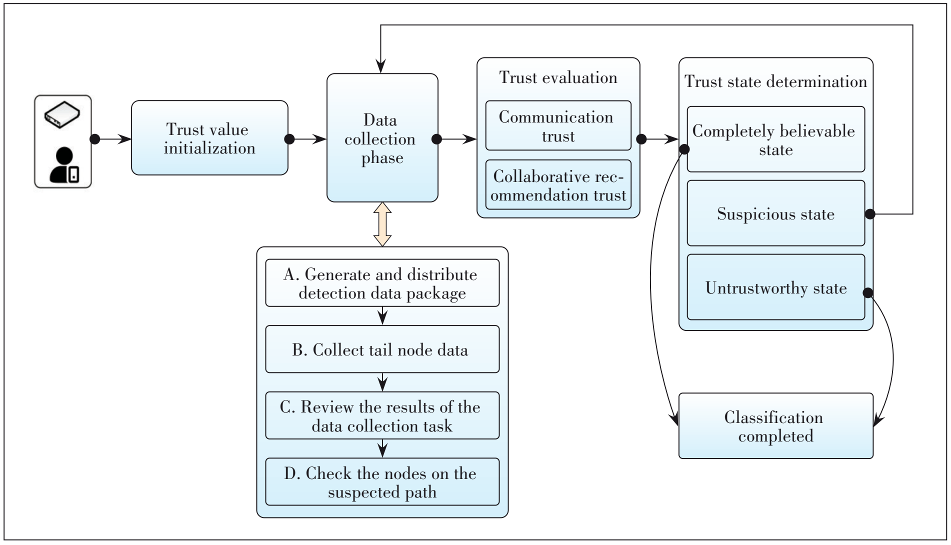
The trust value is the direct basis for judging whether the nodes and MEUs are malicious or not in our scheme. In our trust reasoning mechanism, two calculation methods are mainly used for trust value evaluation: communication behavior trust and collaborative recommendation trust.

1) Communication behavior trust

In each round of a trust verification task, the transmission of the verification data packet by the participating nodes is a communication behavior, and we can obtain the number of successful and malicious communication behaviors respectively. Similarly, we can also acquire two types of behaviors (honest upload behavior and false upload behavior) of MEUs participating in the data collection task. We can accordingly calculate the communication trust between the sensor nodes participating in the transmission task and the MEUs participating in the collection task, and mark them as $(\sigma_{com_i})^{wsn}$ and $(\sigma_{com_j})^{meu}$ respectively, in which i refers to the label of a sensor node in the node set V and and j refers to the label of an MEU in the user set U.



▲Figure 5. Mobile edge user (MEU) collects data package


Trust value initialization
Data collection phase
Trust evaluation
Communication trust
Collaborative recommendation trust
Trust state determination
Completely believable state
Suspicious state
Untrustworthy state
Classification completed
A. Generate and distribute detection data package
B. Collect tail node data
C. Review the results of the data collection task
D. Check the nodes on the suspected path

▲Figure 6. Trust reasoning mechanism based on communication behavior

We then record successful communication behaviors of the sensor node and honest upload behaviors of the MEU as positive communication behaviors, and the malicious communication behaviors of the sensor node and the false reports uploaded by the MEU as negative communication behaviors. We use a subjective logic framework (SLF)[36] to describe communication behavior trust and the formula for calculating the trust value is as follows:

$$\sigma_{com_i} = \frac{2A + B}{2}, \tag{11}$$

where $A = p/(p + n + 1)$, $B = 1/(p + n + 1)$, p is the number of positive communication behaviors of node i, and n is the number of negative communication behaviors of node i.

2) Collaborative recommendation trust

As shown in Fig. 7, the transmission and collection of a verification data packet involves multiple sensor nodes and multiple MEUs (MEU is regarded as mobile nodes), and these nodes coordinate to complete this verification task. With the packet as an intermediary, a virtual connection is created between the nodes, which is called collaborative recommendation in our scheme.

In our example, the nodes labeled a, b, e and f and the MEUs labeled B and C perform their task honestly, and then there is a positive virtual connection between them. However, the nodes labeled c, d and the MEU labeled A do not complete the task faithfully, so there is a negative virtual connection between them. Similar to the communication behavior trust, the subjective logic framework is used in the collaborative recommendation value calculation, and the formula is as follows:

$$\sigma_{rec}(i,j) = \frac{2A + B}{2}, \tag{12}$$

where $A = p/(p + n + 1)$, $B = 1/(p + n + 1)$, p is the number of positive virtual connections established by nodes i and j in multiple rounds of verification and connection tasks, and n is the number of negative virtual connections.

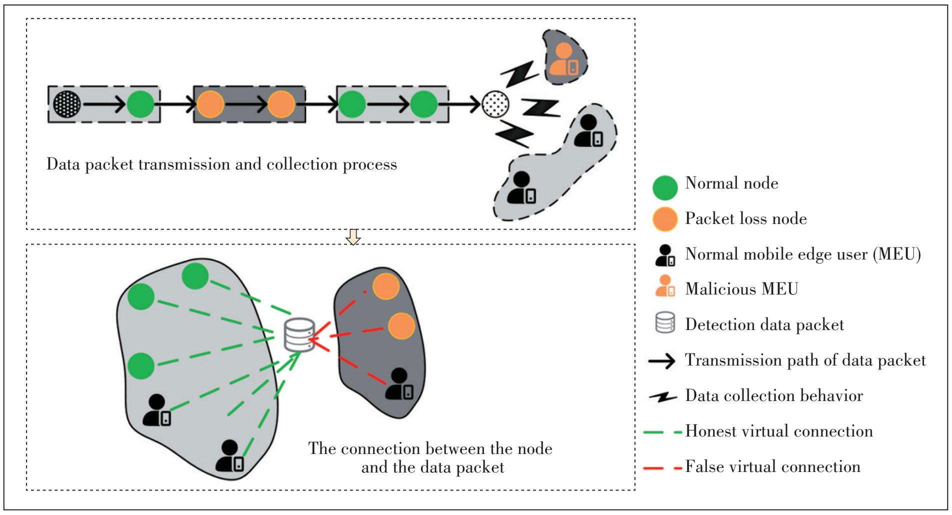
The collaborative recommendation trust coefficient between nodes i and j is $\sigma_{rec}(i,j)$, but if we want to calculate the recommendation trust value of node i, we need to synthesize the collaborative recommendation trust coefficients of all the nodes that have virtual connections with it. What's more, in order to ensure the reliability of recommendation, it is essential to consider the trust of the recommender's own communication behavior trust. In summary, the calculation formula of the node's collaborative recommendation trust is as follows:

$$\sigma_{rec_i} = \frac{\sum_j I_{i,j}(\sigma_{rec}(j,i))^2 \sigma_{com_j}}{\sum_j I_{i,j}\sigma_{rec}(j,i)}, \tag{13}$$

where $I_{i,j}$ is a status indicating whether there is a connection between nodes i and j. When $I_{i,j} = 1$, there is a connection between the two nodes, otherwise not; σ_{com_j} represents the communication behavior trust of node j and takes its own communication behavior trust as the weighting coefficient when node j recommends node i.

Algorithm 2. Algorithm of trust value evaluation (AoTVE) based on communication behavior

```
Input: $\rho_1$, $V$, $U$, $T$
Output: $\hat{V}$, $\hat{U}$
Initialize $\hat{V} = V$, $\hat{U} = U$
For each $\rho_{I,j} \in \rho_I$ Do
    Detection packets are transmitted on the router $\rho_{I,j}$
End For
For each $node_i \in (V^{mod} \cup U^{mod})$ Do
    Calculate $h$, $m$ of $node_i$ by using the data collector and UAV
    Calculate $\sigma_{com_i}$ using $h,m$ in Eq. (11)
    Let $x_i$ be a collection which node all in router path $\rho_I$
    $Q = \varnothing$
    For each $j$, i.e., $node_j \in x_i$ Do
        If $\sigma_{int_j} \neq \sigma_0$ and $\sum_{k=0}^{T} |B_{k,i} + B_{k,j}| \geqslant 2$ Do
            Calculate $p$, $n$ between $node_i$ and $node_j$
            Calculate $\sigma_{rec}(i, j)$ using $p$, $n$ in Eq. (12)
            $Q = Q \cup \{\sigma_{rec}(i,j)\}$
        End If
    End For
    Calculate $\sigma_{rec_i}$ using $Q$ in Eq. (13)
    Calculate $\sigma_{int_i}$ using $\sigma_{rec_i}$, $\sigma_{com_i}$ in Eq. (14)
    If $node_i$ in $V$ Do
```


Data packet transmission and collection process
The connection between the node and the data packet
Normal node
Packet loss node
Normal mobile edge user (MEU)
Malicious MEU
Detection data packet
Transmission path of data packet
Data collection behavior
Honest virtual connection
False virtual connection

▲Figure 7. Trust reasoning mechanism based on communication behavior

```
20:     $\hat{V}(i) = node_i$
21:     Else If $node_i$ in $U$ Do
22:         $\hat{U}(i) = node_i$
23:     End If
24: End For
25: Return $\hat{V}, \hat{U}$
```

Algorithm 2 is the algorithm of trust value evaluation (AoTV) based on communication behavior for two trust values in the trust reasoning mechanism. Line 11 (If $\sigma_{int_j} \neq \sigma_0$ and $\sum_{k=0}^{T} |B_{k,i} + B_{k,j}| \geqslant 2$ Do) restricts the conditions for node j to recommend node i. The restriction conditions require that node j has participated in the verification task before, that is, the comprehensive trust value is not the initial value ($\sigma_{int_j} \neq \sigma_0$), and node j has a virtual connection with node i. For example, suppose that node i and node j perform the task labeled 1 in this round and complete honestly, then $B_{1,i} = B_{2,j} = 1$. If they do not complete the task honestly, then $B_{1,i} = B_{2,j} = -1$. In both cases, $\sum_{k=0}^{T} |B_{k,i} + B_{k,j}| \geqslant 2$ is established, then node i and node j form a recommendation relationship with each other. After calculating the communication behavior trust and collaborative recommendation trust, the comprehensive trust of the node is obtained by the following formula:

$$\sigma_{int_i} = \omega_1 * \sigma_{com_i} + \omega_2 * \sigma_{rec_i}\ , \tag{14}$$

where ω_1 and ω_2 are aggregation constants and the best combination is found by subsequent experiments.

4.2.3 Trust State Determination

After each round of trust evaluation, the nodes participating in the task will update the trust value once. We take the comprehensive trust value of the node as the basis for its state judgment and use two constants $\sigma_{\max}$ and $\sigma_{\min}$ to divide the node into three states. When the comprehensive trust value of the node is greater than $\sigma_{\max}$, the node is judged to be trusted. When the comprehensive trust value of the node is less than $\sigma_{\min}$, the node is judged to be malicious state, and the node is classified as suspicious when it is in between $\sigma_{\max}$ and $\sigma_{\min}$.

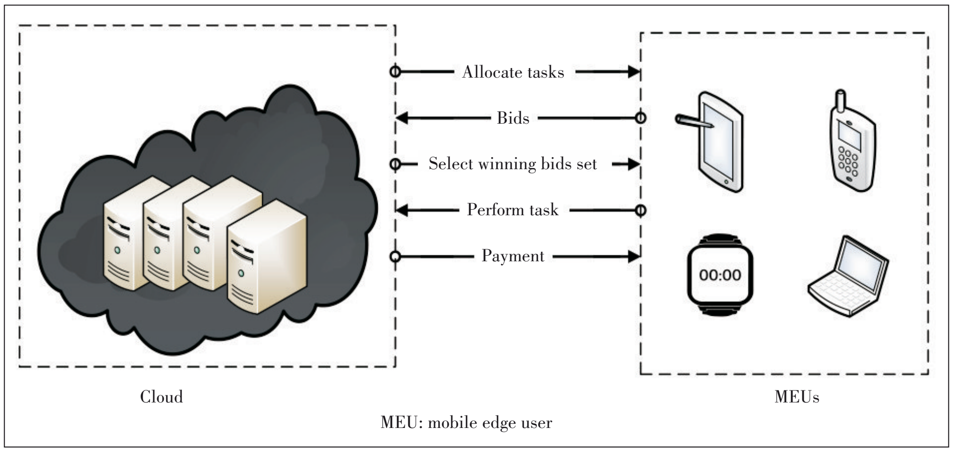
When a node is classified as a trusted state and a malicious state, for the sensor node, it no longer needs to be verified, but for the MEU, we should continuously verify its credibility because of its strong subjectivity. Besides, the higher the MEU' s comprehensive trust, the higher the probability and rewards that it will be selected for the task.

4.3 Incentive Mechanism Based on Cost Performance and Trust

In our scheme, the final result of the data packet transmitted via the sensor node is to be collected by third-party users to reduce the flying distance of the UAV and improve the efficiency of the system. We use mobile crowdsourcing (MCS)-based data collection scheme and complete the task of data collection by hiring a large number of MEUs distributed in the city. The MEUs have strong data storage capacity, computing power and high mobility. They can complete multiple data collection tasks in a short time and make full use of the idle computing power of the equipment.

However, hiring MEUs needs to consider two aspects. On the one hand, it is impossible for MEUs to unconditionally participate in data collection tasks. Participants hope to get actual rewards from providing data, rather than volunteering to provide data for free. Because the perception of data needs to consume resources such as battery power, computing resources and data flow of participants' mobile devices, the participants in this process also need to pay time and labor. Without proper return, participants are not interested in staying active in the MCS-based network for a long time. On the other hand, we need to select MEUs to participate in the task reasonably and give them appropriate remuneration to make full use of the remuneration budget. In other words, we need to focus on the selection criteria of participants.

Therefore, we propose an incentive mechanism based on cost performance and trust. Our incentive mechanism uses a reverse auction framework to describe the relationship between data centers and MEUs. Our mechanism is divided into five steps (Fig. 8). At first, the system selects all tail nodes, and sends data collection tasks to idle MEUs according to each round of verification data packets. Then, the MEUs whose active scopes cover the target node give their own quotes and, after the system receives the quotation from the MEUs, it uses Algorithm 3 to select a set of suitable bids, which is recorded as the winning bids set, and determines the



▲Figure 8. Reverse auction framework

reward based on its performance. Subsequently, the selected MEUs perform data collection tasks within the scope of their activities and upload the collected results to the data center. When we design the winning bids set selection algorithm, we consider two selection criteria:

1) The ratio of an MEU's expected revenue to its data collection capacity. The quotes of an MEU can be expressed as a two-tuple $bid_i = \left(b_i, PoI_i\right)$, in which b_i is the expected reward of the MEU labeled as i, while PoI_i represents the number of task nodes covered in the active range of the MEU labeled as i. This ratio can directly reflect the cost-effectiveness of the data benefits we can obtain by providing remuneration to users.

2) The comprehensive trust value of an MEU. Not all MEUs are authentic and the data they submit may be biased. We can divide untrusted MEUs into two categories: "Greedy Users" who may report falsehood by exaggerating their scope of activities for their own benefit and "Real Malicious Users" who deliberately misrepresent data, thus affecting the true collection of data packets, and have a certain strategy.

Based on the above two criteria, we designed Algorithm 3. The input of the algorithm includes the MEU bid set *BID* and node set *V*. The output of the algorithm is the winner set *S* and their payment set *P*. Then the algorithm uses the greedy method to find the MEU with the maximum sensing performance-price ratio and their trust value in first loop (Lines 3 – 11). In second loop (Lines 13 – 26), for each MEU in the winner set *S*, the algorithm removes the MEU from *S* and continues to select other MEUs in $\overline{BID}$ to join $\bar{S}$ until all the nodes can be accessed. Finally, according to the element in $\bar{S}$, the algorithm gets the payment of each MEU.

Algorithm 3. Winning bids set selection and payment determination

```
Input: $BID, V$
Output: $S, P$
1: $S, P = \varnothing$
2: $\overline{BID} = BID$
3: While $PoI(S)$ is not contain all node of $V$ Do
4:     Select participant $u$ from $\overline{BID}$ by using Eq. (15)
5:     If $PoI(\{u\}) \subseteq PoI(S)$ Then
6:         $\overline{BID} = \overline{BID} \setminus u$
7:     Else
8:         $S = S \cup \{u\}$
9:         $BID = BID \setminus u$
10:    End If
11: End While
12: $\bar{S} = \varnothing$
13: For each $u \in S$ do
14:     $\bar{S} = S \setminus u$
15:     While $PoI(\bar{S})$ is not contain all node of $V$ Do
16:         Select participant $\bar{u}$ from $\overline{BID}$ by using Eq. (15)
17:         If $PoI(\{\bar{u}\}) \subseteq PoI(\bar{S})$ Then
18:             $\overline{BID} = \overline{BID} \setminus \bar{u}$
19:         Else
20:             $S' = S' \cup \{\bar{u}\}$
21:             $BID = BID \setminus \bar{u}$
22:         End If
23:     End While
24:     Calculate $p_b$ by using Eq. (16)
25:     $P = P \cup \{p_b\}$
26: End for
27: Return $S, P$
```

Our incentive mechanism uses the following formula to select the current best participant:

$$u = \max_{i \in BID}\left(\alpha_1 * \frac{b_i}{PoI_i} + \alpha_2 * \sigma_{com_i}\right), \tag{15}$$

where b_i / PoI_i is the ratio of MEU's expected revenue to its data collection capacity and σ_{com_i} is comprehensive trust value of the participant labeled as i. We use proportional coefficients α_1 and α_2 to aggregate the participants' bid scores.

Algorithm 3 uses the ratio of the best data benefit in the alternative set S' to calculate the participant's payment (Lines 12 – 24). The calculation formula is:

$$p_b = \max_{j \in S'}\left(\frac{r_b}{r_j} * PoI_j\right), \tag{16}$$

where r_b is expected revenue of participant u_b, r_j is revenue of participant of u_j, and PoI_j is the number of task nodes in the active area of u_j.

5 Performance Analysis

5.1 Experiment Setup

We realized the UAV-DT scheme in Python 3.7 and ran a simulation experiment on IdeaPad Air 14 with 16 GB 2 133 MHz LPDDR4 RAM, whose CPU parameters is 2.10 GHz AMD Ryzen 5 4600U with Radeon Graphics.

The important parameters used in our experiments are listed in Table 2. Each experiment was carried out in a network area of 100×100 m^2, where 1 000 smart devices and 500 ME-

▼Table 2. Experimental parameters

Parameter	Value
Size of area/m^2	100 ×100
Number of sensor nodes	1 000
Number of MEUs	500
Active radius of MEU/m	[5, 10]
Payment of hiring MEU	[10, 25]

MEU: mobile edge user

Us were randomly deployed. We randomly created 20 different network scenarios in total and ran them once in each experiment. The results were averaged to ensure the robustness of our strategy in different network scenarios.

For a normal sensor device, there was a small probability of packet loss in the process of transmitting data packets due to network fluctuations. However, a malicious sensor device would deliberately discard a part of the data packet with a greater probability. We gave 5% and 20% probabilities for two different packet loss situations, which were reflected in the form of random functions in the simulation experiments. In the simulation, the data MEU reported had a 10% – 40% probability of being false.

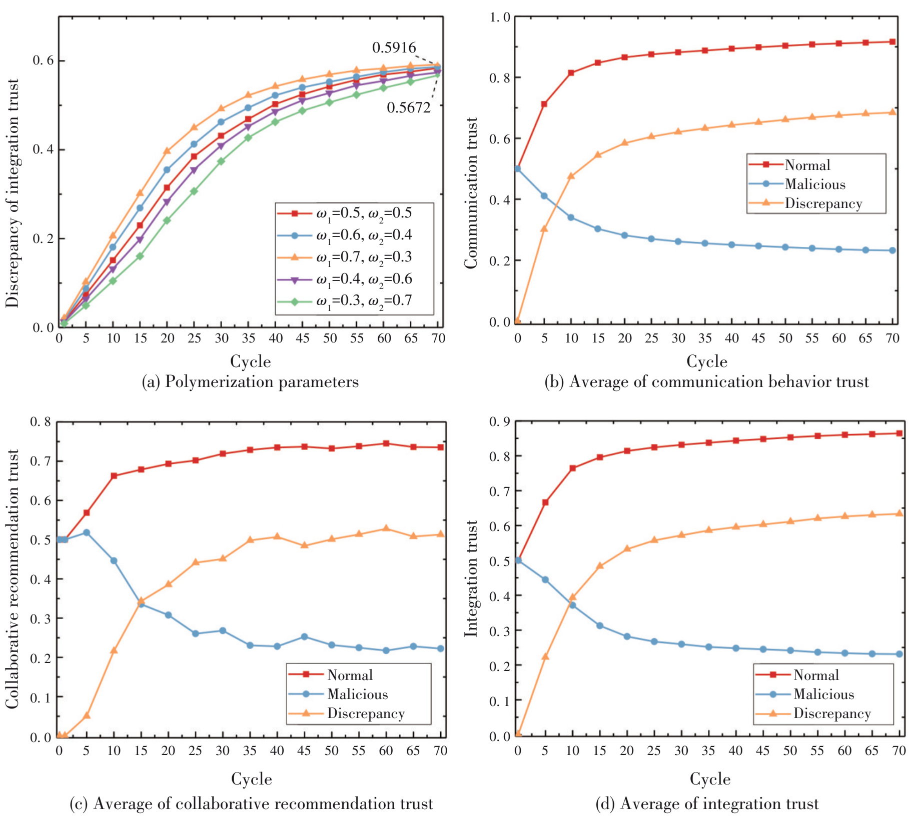
In the experiments, 70 rounds of verification tasks were carried out in each scenario, and in each round, we used the drone to release 30 data packets starting with random sensor device. The length of the routing path of each packet was fixed to 10 nodes.

5.2 Discrepancy of Trust Values

In our scheme, the communication behavior trust and collaborative recommendation trust are aggregated into integration trust, then whether a behavior is malicious or not is determined by setting two thresholds (σ_{max} and σ_{min}), which involves two aggregation coefficients ω_1 and ω_2.

In the discrepancy experiments, we set five sets of coefficients to test the effect of different coefficients on the discrepancy of trust values. We set $\hat{\omega}_1 = (0.5, 0.5)$, $\hat{\omega}_2 = (0.6, 0.4)$, $\hat{\omega}_3 = \left(0.7, 0.3\right)$, $\hat{\omega}_4 = (0.4, 0.6)$, and $\hat{\omega}_5 = (0.3, 0.7)$, and guaranteed $\omega_1 + \omega_2 = 1$. Besides, since the optimal classification threshold has not been determined, so our experiments did not classify nodes.

Fig. 9(a) shows that after 70 rounds of verification tasks are completed, the values of discrepancy between normal and malicious nodes in each set of experiments are in the interval between 0.5672 and 0.5916. Obviously, when we set $\hat{\omega} = \hat{\omega}_3 = \left(0.7, 0.3\right)$, the discrepancy reaches a peak, equal to 0.5916. The result shows that our scheme has a high degree of discrimination between normal and malicious nodes and the discrepancy between the two reaches a high value. Thus, we still use this set of aggregation coefficients in subsequent experiments.



(a) Polymerization parameters

(b) Average of communication behavior trust

(c) Average of collaborative recommendation trust

(d) Average of integration trust

▲**Figure 9. Influence of different parameters on the average trust between normal and malicious nodes**

We continue to advance the discrepancy experiment subsequently and Figs. 9(b), 9(c) and 9(d) show the discrepancy of communication behavior trust, collaborative recommendation trust, and integration trust between normal and malicious nodes. The three trust values of normal nodes are 0.916, 0.735 and 0.864 after 70 rounds of verification tasks, and the three trust values of malicious nodes are 0.684, 0.513 and 0.633. Compared with collaborative recommendation trust, the curve of communication behavior trust is smoother and the convergence speed is faster. After 20 rounds, the average of discrepancy in communication behavior trust has reached a high level. In terms of collaborative recommendation trust, the numerical curve has fluctuations in 70 rounds. As shown in Fig. 9(c), the average trust of normal nodes is not high enough so that the distinction between the two is not obvious.

The following is the conclusions of this group of experiments:

1) It is reasonable to trust a higher aggregation coefficient for communication behavior trust, and when the classification thresholds (σ_{max} and σ_{min}) are not set, we use the communication behavior trust to distinguish normal nodes from malicious nodes clearly.

2) The verification effect of communication behavior trust and collaborative recommendation trust can still be further optimized. We set classification thresholds σ_{max} and σ_{min}, and the system excludes the nodes that can be clearly identified as

the trustworthy or malicious from subsequent tasks, so that the left nodes with doubtful status (that is $\sigma_{\min} < \sigma_{int_i} < \sigma_{\max}$) can get a chance to be verified.

From Fig. 9(d), we can clearly see the average integration trust of normal nodes is 0.864 and has a slight upward trend. The average integration trust of malicious nodes is 0.231 and has a downward trend stably. Then we use 0.85 and 0.25 as the central values to find the best classification thresholds.

5.3 Discriminant Rate of Normal and Malicious Nodes

In this section, we set two groups of thresholds to conduct classification discrimination rate experiments. We use 0.85 and 0.25 as the central values of two groups ($\overline{\sigma_{\max}} = 0.85$, $\overline{\sigma_{\min}} = 0.25$), and find the best value in the interval between the upper and lower domains is 0.1. Then the interval of $\sigma_{\max}$ is ($\overline{\sigma_{\max}} - 0.05, \overline{\sigma_{\max}} + 0.05$) and the interval of $\sigma_{\min}$ is $\left(\overline{\sigma_{\min}} - 0.05, \overline{\sigma_{\min}} + 0.05\right)$.

As shown in Fig. 10, in the first three sets of experiments, the discrimination rates of trusted nodes converge quickly and the final result is around 0.988, which means the correct discrimination rate of trustworthy results reaches 98.8%. Even if the result of the last set in more stringent conditions reaches 0.930, the correct discrimination rate can reach up to 93%. The results of the experiments on the discrimination rate of malicious nodes are shown in Fig. 11. The discrimination effect of groups 2 - 5 is better and the final discrimination rate is between 0.821 and 0.933. Under the most stringent threshold conditions, when $\sigma_{\min} = 0.20$ in the experiment, the discrimination rat is still higher than 65%.

From the experimental results, it can be seen that our scheme has excellent effect and robustness on the recognition ability of trusted and malicious nodes, and can obtain a high recognition rate even after 20 rounds. Then we choose $\sigma_{max} = 0.85$ and $\sigma_{\min} = 0.30$ as the best thresholds for our system to classify nodes. The node labeled as i will be treated as a trusted node when $\sigma_{int_i} > \left(\sigma_{\max} = 0.85\right)$ and a malicious node when $\sigma_{int_i} < (\sigma_{min} = 0.30)$.

Under the best classification threshold, the classification results are shown in Fig. 12. Trusted and malicious nodes are thoroughly classified after 30 rounds. Finally, the classification rate of trusted nodes is as high as 98.9%, while the classification rate of malicious nodes also reaches 94.2%. In other words, only 15 nodes are still in doubt status after 70 rounds in our simulation environment.

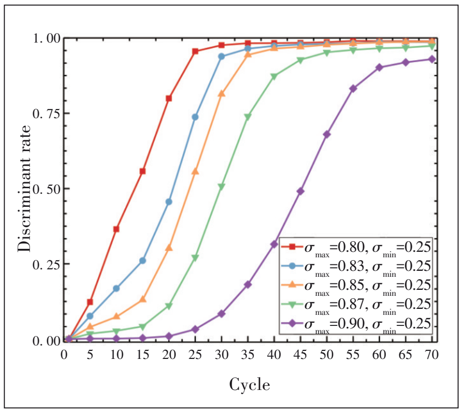
After the best classification thresholds ($\sigma_{max} = 0.85$ and $\sigma_{min} = 0.30$) are set, we repeat the experiment in Section 5.2 with $\omega_1 = 0.7$ and $\omega_2 = 0.3$. As shown in Figs. 13(a), 13(b) and 13(c), our scheme has a good improvement on the trust evaluation ability of nodes after setting the classification thresholds.

In terms of communication behavior trust, the average trust of normal nodes, the average trust of malicious nodes and the discrepancy between normal and malicious nodes are 0.929, 0.241, 0.688 respectively, which are slightly improved. In terms of collaborative recommendation trust, the average trust of normal nodes, the average trust of malicious nodes and the discrepancy between normal and malicious nodes have changed from 0.735, 0.222, and 0.513 to 0.890, 0.282 and 0.608, respectively. The discrepancy has increased by 18.5%, which means that the evaluation effect of collaborative recommendation trust has been improved. Combining the above two types of trust, the result of discrepancy in integration trust has increased by 4.8%.

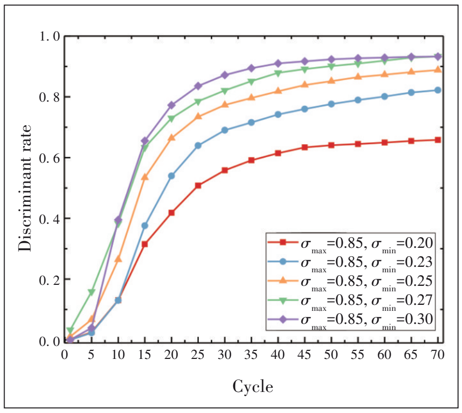
Based on the experiments, we determine the classification thresholds for node classification in our scheme, and the results prove that the node classification ability of the system is very significant.

5.4 Collection Rate

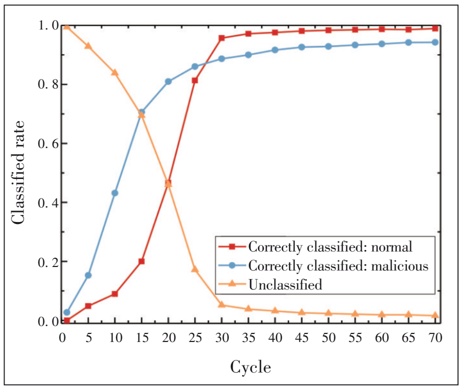
In our network scenario, there are some malicious nodes, which randomly discard some data packets passing through it with a certain probability. In this section, we generate the same number of regular data packets as the verification data packets, transmit them on the network, and hand them over to the MEU



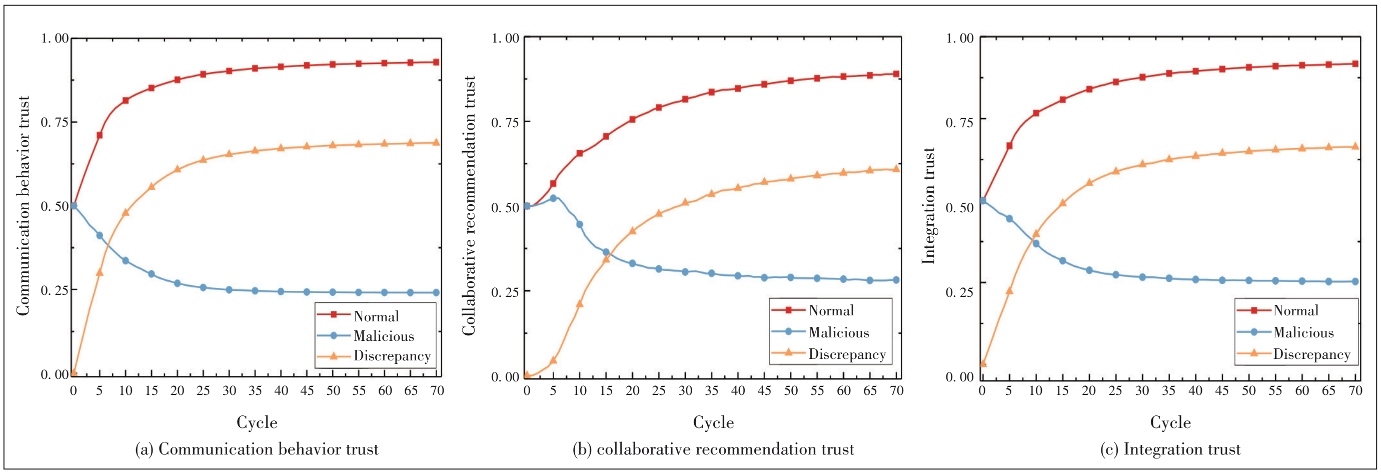
▲Figure 10. Discrimination rate of trusted nodes



▲Figure 11. Discrimination rate of malicious nodes



▲Figure 12. Discriminative ability under the optimal classification threshold



(a) Communication behavior trust (b) collaborative recommendation trust (c) Integration trust

▲Figure 13. Trust evaluation capability under the optimal threshold

for collection. In contrast, we also simulated the collection of regular data packets in the original network scenario without our scheme, which is usually called unverified network.

The results are shown in Figs. 14 and 15. In our scheme, the collection rate curve first rises quickly and stabilizes in a very high value range, while the collection rate curve fluctuates at a relatively low position in the unverified network. After 14 rounds (when the most nodes in the network are classified), the collection rate of our scheme stays within the range 0.88 to 0.92, and the collection rate of unverified network maintains between 0.78 and 0.82. The average of the former is 0.899, while that of the latter is 0.808, that is, our scheme improves 11.2% compared with the unverified network.

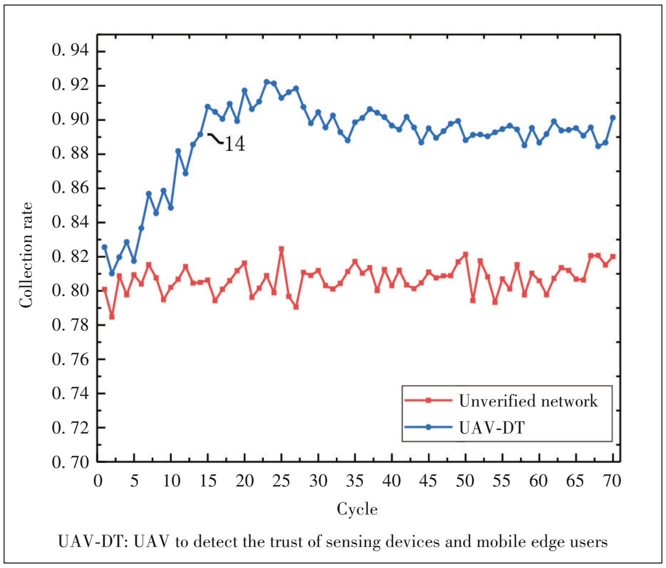
From another perspective, in our network scenario, there are two main reasons for packet loss: network fluctuations and the malicious node that deliberately loses packets. We also conducted two comparative experiments on the causes of packet loss.

As shown in Fig. 15, the proportions of the packet loss ratio caused by network fluctuations and malicious nodes are relatively stable in multiple rounds of experiments in the unverified network. The former is 26.6% in average and the latter is 73.4% correspondingly. In our scheme, the proportions of the two keep changing with the increase of rounds and the proportion of malicious nodes intentionally losing packets is slowly decreasing from 0.638 to 0.468.

From the above experimental results, it can be seen that our scheme effectively detects a large number of trusted nodes and malicious nodes in the network, thereby avoiding malicious nodes during data packet transmission and improving the collection rate of data packets.

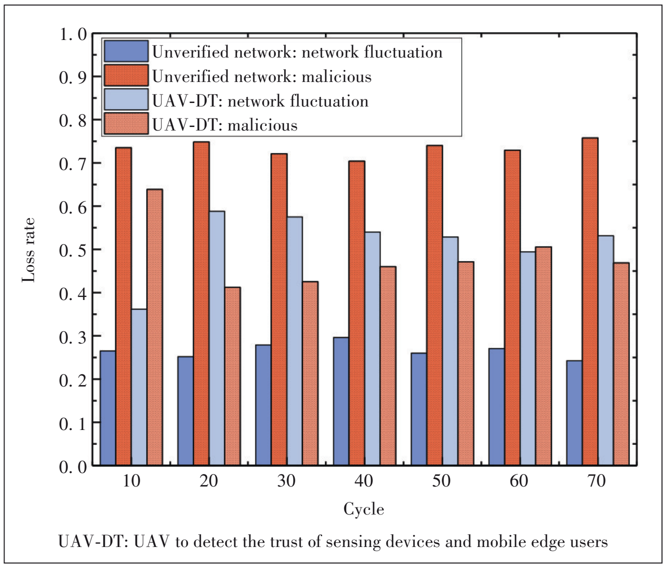
5.5 Cost

In this section, our experiments compare the winning bids



UAV-DT: UAV to detect the trust of sensing devices and mobile edge users

▲Figure 14. Collection rate



UAV-DT: UAV to detect the trust of sensing devices and mobile edge users

▲Figure 15. Rate of packet loss

set selection algorithm in our scheme with the conventional greedy algorithm (Fig. 16). The total cost of our system is divided into two parts: the cost of hiring MEUs and additional costs (the cost of sending drones for additional verification). We assume that the cost of each additional verification by the drone is as five times much as the cost of hiring an MEU.

According to the experimental results, the cost of hiring MEU in each round is much higher than the cost of additional verification of drones. The reason is: in our network scenario, there are more normal IoT devices than malicious ones and the frequency of transmission errors is relatively low compared to the total number of transmissions. Therefore, the cost of the system is mainly focused on hiring MEUs. In addition, it is obvious that the cost of our scheme for hiring MEUs is lower than that of the greedy strategy, with an average reduction of 23.4%. Although the additional costs are slightly higher, our scheme is still the best in terms of total cost, with an average reduction of 10.7%. Our UAV-DT scheme spends significantly less on employment than the greedy strategy. The greedy strategy does not consider the trust value in the selection range when selecting MEUs to participate in the data collection, which causes the suspected path shown in Fig. 4. This will inevitably lead to an increase in the cost of using UAV for review. On the contrary, UAV-DT uses the trust value of MEUs as the selection criterion shown in Algorithm 2.

6 Conclusions

In this paper, we propose a low-cost and efficient UAV-DT security scheme, including the UAV-assisted trust verification mechanism, incentive mechanism based on cost performance and trust, and trust reasoning mechanism based on communication behavior. By continuously verifying the trust of the nodes in the network, the trusted and malicious nodes can be quickly distinguished by comparing their trust values.

Our experimental results show that our security scheme has high discrimination for malicious nodes, and provides an effective solution to efficient and safe data collection in the city.

However, we did not use a good path planning scheme in the UAV broadcast verification packet stage and the UAV secondary inspection stage. Further studies are needed in future and we will focus on how to develop an efficient UAV flight path in the future research.

References

[1] HILLS G, LAU C, WRIGHT A, et al. Modern microprocessor built from complementary carbon nanotube transistors [J]. Nature, 2019, 572(7771): 595 – 602. DOI: 10.1109/tcad.2015.2415492

[2] REN Y Y, WANG T, ZHANG S B, et al. An intelligent big data collection technology based on micro mobile data centers for crowdsensing vehicular sensor network [J]. Personal and ubiquitous computing, 2020: 1 – 17. DOI: 10.1007/s00779-020-01440-0

[3] YU M Y, LIU A F, XIONG N N, et al. An intelligent game based offloading scheme for maximizing benefits of IoT-edge-cloud ecosystems [J]. IEEE Internet of Things journal, 2020, early access. DOI: 10.1109/JIOT.2020.3039828

[4] Gartner. 20.4 billion connected things by 2020 [EB/OL]. (2017-2-09) [2021-05-01]. https://www.itp.net/611397-204-billion-connected-things-by-2020-gartner

[5] LI T, LIU W, ZENG Z W, et al. DRLR: a deep reinforcement learning based recruitment scheme for massive data collections in 6G-based IoT networks [J]. IEEE Internet of Things journal, 2021, early access. DOI: 10.1109/JIOT.2021.3067904

[6] HUANG M F, ZHANG K, ZENG Z W, et al. An AUV-assisted data gathering scheme based on clustering and matrix completion for smart ocean [J]. IEEE Internet of Things journal, 2020, 7(10): 9904 – 9918. DOI: 10.1109/JIOT.2020.2988035

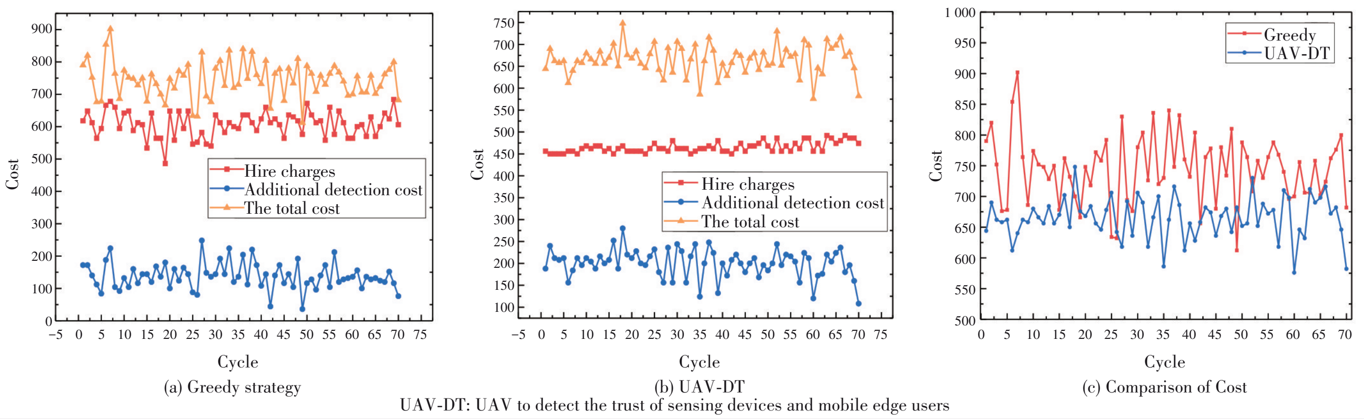
[7] OUYANG Y, LIU A F, XIONG N X, et al. An effective early message ahead join adaptive data aggregation scheme for sustainable IoT [J]. IEEE transactions on network science and engineering, 2021, 8(1): 201 – 219. DOI: 10.1109/TNSE.2020.3033938

[8] LI A, LIU W, ZENG L J, et al. An efficient data aggregation scheme based on differentiated threshold configuring joint optimal relay selection in WSNs [J]. IEEE access, 2021, 9: 19254 – 19269. DOI: 10.1109/ACCESS.2021.3054630

[9] WANG T, ZHANG G X, BHUIYAN M Z A, et al. A novel trust mechanism based on fog computing in sensor-cloud system [J]. Future generation computer systems, 2020, 109: 573 – 582. DOI: 10.1016/j.future.2018.05.049

[10] LIU S, HUANG G S, GUI J S, et al. Energy-aware MAC protocol for data differentiated services in sensor-cloud computing [J]. Journal of cloud computing, 2020, 9(1): 1 – 33. DOI: 10.1186/s13677-020-00196-5

[11] LI F F, HUANG G S, YANG Q, et al. Adaptive contention window MAC protocol in a global view for emerging trends networks [J]. IEEE access, 2021, 9: 18402 – 18423. DOI: 10.1109/ACCESS.2021.3054015



(a) Greedy strategy (b) UAV-DT (c) Comparison of Cost

UAV-DT: UAV to detect the trust of sensing devices and mobile edge users

▲**Figure 16. System Costs**

[12] HUANG C Q, HUANG G S, LIU W, et al. A parallel joint optimized relay selection protocol for wake-up radio enabled WSNs [J]. Physical communication, 2021, 47: 101320. DOI: 10.1016/j.phycom.2021.101320

[13] GUO J L, LI F F, WANG T, et al. Parameter analysis and optimization of polling-based medium access control protocol for multi-sensor communication [J]. International journal of distributed sensor networks, 2021, 17(4): 155014772110074. DOI: 10.1177/15501477211007412

[14] PALADINO, FISSORE, NEVIANI. A low-cost monitoring system and operating database for quality control in small food processing industry [J]. Journal of sensor and actuator networks, 2019, 8(4): 52. DOI: 10.3390/jsan8040052

[15] HUANG S B, ZENG Z W, OTA K, et al. An intelligent collaboration trust interconnections system for mobile information control in ubiquitous 5G networks [J]. IEEE transactions on network science and engineering, 2021, 8(1): 347 – 365. DOI: 10.1109/TNSE.2020.3038454

[16] ZHU X Y, LUO Y Y, LIU A F, et al. Multiagent deep reinforcement learning for vehicular computation offloading in IoT [J]. IEEE Internet of Things journal, 2021, 8(12): 9763 – 9773. DOI: 10.1109/JIOT.2020.3040768

[17] TENG H J, DONG M X, LIU Y X, et al. A low-cost physical location discovery scheme for large-scale Internet of Things in smart city through joint use of vehicles and UAVs [J]. Future generation computer systems, 2021, 118: 310 – 326. DOI: 10.1016/j.future.2021.01.032

[18] DENG Q Y, OUYANG Y, TIAN S J, et al. Early wake-up ahead node for fast code dissemination in wireless sensor networks [J]. IEEE transactions on vehicular technology, 2021, 70(4): 3877 – 3890. DOI: 10.1109/TVT.2021.3066216

[19] BONOLA M, BRACCIALE L, LORETI P, et al. Opportunistic communication in smart city: Experimental insight with small-scale taxi fleets as data carriers [J]. Ad hoc networks, 2016, 43: 43 – 55. DOI: 10.1016/j.adhoc.2016.02.002

[20] HUANG S B, LIU A F, ZHANG S B, et al. BD-VTE: A novel baseline data based verifiable trust evaluation scheme for smart network systems [J]. IEEE transactions on network science and engineering, 2021, 8(3): 2087 – 2105. DOI: 10.1109/TNSE.2020.3014455

[21] GUO J L, LIU A F, OTA K, et al. ITCN: an intelligent trust collaboration network system in IoT [J]. IEEE transactions on network science and engineering, 2021, early access. DOI: 10.1109/TNSE.2021.3057881

[22] LI T, LIU A F, XIONG N N, et al. A trustworthiness-based vehicular recruitment scheme for information collections in distributed networked systems [J]. Information sciences, 2021, 545: 65 – 81. DOI:10.1016/j.ins.2020.07.052

[23] HU L, LIU A F, XIE M D, et al. UAVs joint vehicles as data mules for fast codes dissemination for edge networking in smart city [J]. Peer-to-peer networking and applications, 2019, 12(6): 1550 – 1574. DOI: 10.1007/s12083-019-00752-0

[24] OUYANG Y, ZENG Z W, LI X, et al. A verifiable trust evaluation mechanism for ultra-reliable applications in 5G and beyond networks [J]. Computer standards & interfaces, 2021, 77: 103519. DOI: 10.1016/j.csi.2021.103519

[25] ZHU X Y, LUO Y Y, LIU A F, et al. A deep learning-based mobile crowdsensing scheme by predicting vehicle mobility [J]. IEEE transactions on intelligent transportation systems, 2021, 22(7): 4648 – 4659. DOI: 10.1109/TITS.2020.3023446

[26] HUANG W, OTA K, DONG M X, et al. Result return aware offloading scheme in vehicular edge networks for IoT [J]. Computer communications, 2020, 164: 201 – 214. DOI: 10.1016/j.comcom.2020.10.019

[27] SHEN M Q, LIU A F, HUANG G S, et al. ATTDC: an active and traceable trust data collection scheme for industrial security in smart cities [J]. IEEE Internet of Things journal, 2021, 8(8): 6437 – 6453. DOI: 10.1109/JIOT.2021.3049173

[28] WANG T, LUO H, ZHENG X, et al. Crowdsourcing mechanism for trust evaluation in CPCS based on intelligent mobile edge computing [J]. ACM transactions on intelligent systems and technology, 2019, 10(6): 1 – 19. DOI: 10.1145/3324926

[29] BAEK D, CHEN J, CHOI B J. Small profits and quick returns: An incentive mechanism design for crowdsourcing under continuous platform competition [J]. IEEE Internet of Things journal, 2020, 7(1): 349 – 362. DOI: 10.1109/JIOT.2019.2953278

[30] LIU Y X, DONG M X, OTA K, et al. ActiveTrust: secure and trustable routing in wireless sensor networks [J]. IEEE transactions on information forensics and security, 2016, 11(9): 2013 – 2027. DOI: 10.1109/TIFS.2016.2570740

[31] WAGGONER B and CHEN Y L. Output agreement mechanisms and common knowledge [C]//Second AAAI Conference on Human Computation & Crowdsourcing (HCOMP). Pittsburgh, USA: AAAI, 2014

[32] HUANG C, YU H R, BERRY R A, et al. Using truth detection to incentivize workers in mobile crowdsourcing [J]. IEEE transactions on mobile computing, 2020, early access. DOI: 10.1109/TMC.2020.3034590

[33] KIM T K, SEO H S. A trust model using fuzzy logic in wireless sensor network [J]. World academy of science, engineering and technology, 2018, 42: 63 – 66

[34] FUANG W D, ZHANG C L, SHI Z D, et al. BTRES: beta-based trust and reputation evaluation system for wireless sensor networks [J]. Journal of network and computer applications, 2016, 59: 88 – 94. DOI: 10.1016/j.jnca.2015.06.013

[35] YAO Z Y, KIM D Y, DOH Y M. PLUS: parameterized and localized trust management scheme for sensor networks security [C]//IEEE International Conference on Mobile Adhoc and Sensor Systems. Vancouver, Canada, 2006: 437 – 446. DOI: 10.1109/MOBHOC.2006.278584

[36] BALAKRISHNAN V, VARADHARAJAN V, TUPAKULA U. Subjective logic based trust model for mobile ad hoc networks [C]//4th International Conference on Security and Privacy in Communication Netowrks. Istanbul, Turkey: ACM, 2008: 1 – 11. DOI: 10.1145/1460877.1460916

Biographies

LI Xiuxian is currently pursuing his master's degree at School of Computer Science and School of Cyberspace Science from Xiangtan University, China. His research interests include mobile crowding sensing, IoT devices, and edge computing.

LI Zhetao (liztchina@hotmail.com) is a professor with the College of Computer, Xiangtan University, China. He received his B.Eng. degree in electrical information engineering from Xiangtan University in 2002, the M.Eng. degree in pattern recognition and intelligent system from Beihang University, China in 2005, and the Ph.D. degree in computer application technology from Hunan University, China in 2010. From December 2013 to December 2014, he was a postdoc in wireless network at Stony Brook University, USA. He is a member of IEEE and CCF.

OUYANG Yan is currently a postgraduate student with the School of Computer Science and Engineering, Central South University, China. Her research interests include crowd sensing networks and wireless sensor networks.

DUAN Haohua received his bachelor's degree in computer science and technology from Jilin University, China in 2020. He is currently pursuing his Ph.D. degree in computer science, Shanghai Jiao Tong University, China. His research interests include security and privacy in machine learning and blockchain.

XIANG Liyao received her B.Eng. degree in electrical and computer engineering from Shanghai Jiao Tong University, China in 2012, and Ph.D. degree in computer engineering from the University of Toronto, Canada in 2018. She is currently an assistant professor with Shanghai Jiao Tong University. Her research interests include security and privacy, privacy analysis in data mining, and mobile computing.