

## Research Article

# A Time and Location Correlation Incentive Scheme for Deep Data Gathering in Crowdsourcing Networks

Fulong Ma,<sup>1</sup> Xiao Liu,<sup>1</sup> Anfeng Liu ,<sup>1</sup> Ming Zhao,<sup>2</sup> Changqin Huang,<sup>3</sup> and Tian Wang <sup>4</sup>

<sup>1</sup>School of Information Science and Engineering, Central South University, Changsha 410083, China

<sup>2</sup>School of Software, Central South University, Changsha 410083, China

<sup>3</sup>School of Information Technology in Education, South China Normal University, Guangzhou 510631, China

<sup>4</sup>College of Computer Science & Technology, Huaqiao University, Xiamen, Fujian Province 361021, China

Correspondence should be addressed to Anfeng Liu; [afengliu@mail.csu.edu.cn](mailto:afengliu@mail.csu.edu.cn)

Received 4 September 2017; Accepted 14 November 2017; Published 9 January 2018

Academic Editor: Zhaolong Ning

Copyright © 2018 Fulong Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To tackle the issue in deep crowd sensing, a Time and Location Correlation Incentive (TLCI) scheme is proposed for deep data gathering in crowdsourcing networks. In TLCI scheme, a metric named “Quality of Information Satisfaction Degree” (QoISD) is to quantify how much collected sensing data can satisfy the application’s QoI requirements mainly in terms of data quantity and data coverage. Two incentive algorithms are proposed to satisfy QoISD with different view. The first algorithm is to ensure that the application gets the specified sensing data to maximize the QoISD. Thus, in the first incentive algorithm, the reward for data sensing is to maximize the QoISD. The second algorithm is to minimize the cost of the system while meeting the sensing data requirement and maximizing the QoISD. Thus, in the second incentive algorithm, the reward for data sensing is to maximize the QoISD per unit of reward. Finally, we compare our proposed scheme with existing schemes via extensive simulations. Extensive simulation results well justify the effectiveness of our scheme. The QoISD can be optimized by 81.92%, and the total cost can be reduced by 31.38%.

## 1. Introduction

Internet of Things (IoT) [1–5] and cloud computing [6–13] take advantage of the ubiquity of smart sensor-equipped devices such as smartphones, iPad, and vehicle sensor devices. All of these are to collect information with low cost and provide a new paradigm for solving the complex data sensing based applications from the significant demands of critical infrastructure such as surveillance systems [14–16], remote patient care systems in healthcare, intelligent traffic management, and automated vehicles in transportation environmental [17–19] and weather monitoring systems. In such applications, due to the need to collect a wide range and large amount of data even in a long-term continuous period, the cost of traditional method deploying sensing devices or specialized employee to collect the data is so high limiting the development of applications [20, 21]. Hence, a new method of data collection named crowd sensing (or participatory sensing) is adopted to collect data [21–26]. First, there are a

large number of users who can participate in data collection. Because of the development of electrical devices, for example, the smartphones have been very popular in the past few years [27]. In the final quarter of 2010, it is the first time that smartphone sales passed PCs [28]. Smartphones are equipped with rich sensors which are cheap but useful, such as accelerometer, digital compass, GPS, microphone, gyroscope, and camera. These sensors can monitor a diverse range of human activities and collect large amounts of useful information [23, 28]. These can sense data through their smartphones or smart sensing devices which are collectively referred to as users. Second, the data in this way to collect has a broad range and at a low cost, which can provide effective data for many applications. This data collection is generally carried out as shown in Figure 1. In such a crowdsourcing sensing system, there are four main components. (a) users (also known as crowdsourcing, or participants): they represent the smart sensing devices or people who hold sensing devices, which are to sense data samples, report to applications, and

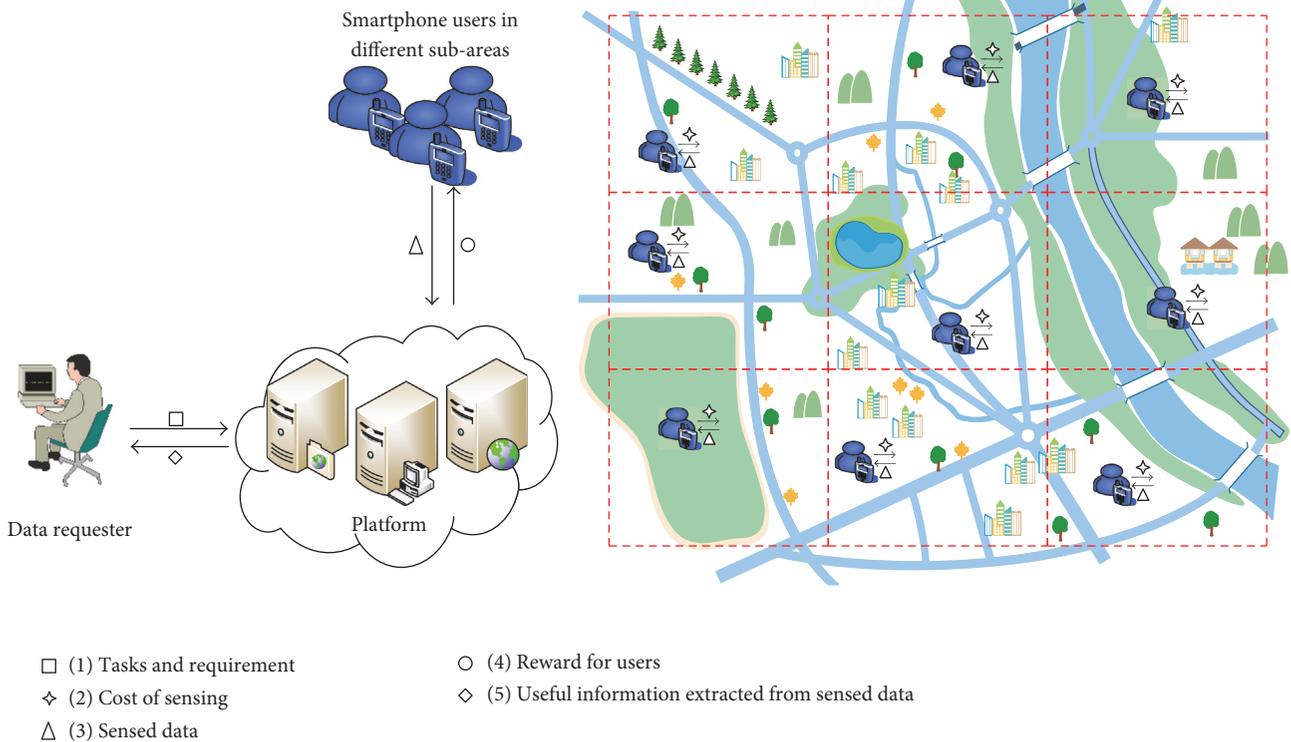


FIGURE 1: A crowdsourcing sensing system.

gain the reward, also known as reporter; (b) applications (or task publisher): application is the demander of data, which publishes the requirements for data and pays a certain reward to users, motivating the users to collect sensing data, and after users collect data and report to applications, applications provide the customers for advanced services based on the collected sensing data and charge a certain fee to make up for the data contribution; (c) customers: the people who use the application's services and pay a certain fee to applications; (d) the third-party process platform: the interaction between users, applications, and customers is not certainly direct and may be carried out through the third-party trusted platform and the third-party trusted interactive platform helping the interaction between the parties becomes more effective and convenient and it has been widely used in many researches and practices. The main body of this paper is the incentive mechanism and interaction between users and applications [23]. The interactive process of crowdsourcing sensing system is shown as follows: the applications publish the attributes, such as area, time period, storage quality, and the reward for reporting data samples [22]. The users in target area collect the sensing data, report the sensed data to applications, and gain the reward. Applications process the collected data, form the services, and provide the customers for services. Because a large number of users are involved in collecting the sensing data, it is fast and of low cost and high quality to complete the applications which cannot be achieved in the past. For example, in the wide-area bird ecology system, in order to obtain the biological habits of the entire bird ecosystem, migration, and the relationship between the food chains, a

large amount of data must be collected in a wide location and for long time to provide effective service. In these large data services, due to the large amount and range, long time period, and even high data acquisition frequency of data collection, the cost is very high if the applications deploy and maintain the data acquisition equipment (or employee) by itself. Thus, after the applications published needed data elements of bird data samples, time, locations, and the reward for the data, a large number of users volunteered to collect the data and report to applications, which can form the services which are of low cost and high quality and have rich sampling data compared to other methods.

It is foreseeable that, with the development of smart sensing devices, the sensing devices will become more powerful and have a wider sensing scope. And crowd sensing based applications will play an increasingly important role in every aspect of mankind in the future, so as to provide more convenient and better user experience for human. However, it depends on the high-quality data samples that the crowd sensing processes the data, provides customers for advanced services, and gets the payoff which can maintain a positive input to the data collection to ensure the sustainable development of applications. For data collection, it is necessary that rewards are introduced into the crowdsourcing sensing system because (1) it is possible that participating in a sensing task may incur monetary costs, the usage of network bandwidth, and the consumption of the power of the smartphones. Thus, the rewards could motivate them to tolerate these costs and contribute to the sensing task. (2) The smartphones are personal which are not controlled by others.

We cannot decide when and how to collect sensor data and it is necessary that rewards can be used to affect their behavior. As a general incentive mechanism, the users who participate in sensing will get monetary reward since they provide the sensed data.

There are many existing studies [29–32] that focus on the incentive mechanism. This paper summarizes the existing research into the following categories.

(1) The incentive mechanisms which are for the purpose of certain number of data samples: the main objective of this type of incentive mechanism is to obtain a sufficient number of data samples. There are many ways in such strategy, such as game based mechanism and demand and supply model based market incentive mechanism.

However, in this type of study, the main consideration is to collect a sufficient number of data samples with the least cost, but without considering or rarely considering the quality of data. It obviously no longer meets the current rapid development of the application demands.

(2) Incentive mechanism based on quality: this kind of research not only considers the quantity and cost of data collection but also takes quality of data collection into account. For example, Tham and Luo [30] propose a metric called Quality of Contributed Service (QCS), and the metric named Quality of Information (QoI) is also proposed by Reddy et al. [31] to evaluate the quality of the data samples. This type of incentive mechanism achieves efficient data collection through selecting high QCS reporters. Data coverage is also studied by Song et al. [32], which recruits the most matching participants to maximize coverage. But the actual performance of these strategies is not stratifying due to three reasons as follows. First of all, these studies tend to think that the participants can be controlled by the applications so that the applications can select or recruit the certain participants to meet the QoS, QoI, or coverage of applications. However, whether users participate in sensing task depends on themselves in practice. Hence, the participants cannot be chosen by the application. Secondly, it always assumes that the number of users involved in sensing data is very large in these applications so that the applications could only consider how to select the certain users meeting the needs of application, which is impossible in many cases. Even if in the downtown area where there are rich sensing devices, if there is no sufficient incentive ability, the users who can participate in sensing data are so few that there are little participants which can be chosen by applications. In this case, the focus of the study is to study how to make a suitable reward to motivate users to participate in sensing task rather than considering how to select certain users. Last but not least, in these studies, even if selecting the data with Quality of Information (QoI) in high priority, it cannot obtain high-quality applications from the overall perspective. The reason is that the quality of data is high from the partial perspective, but, in the overall perspective, there is no improvement on quality of applications while collecting data samples in the area where the collected data is enough, even if the data is of high quality. At the time, the data in the area where the collected data is very few can efficiently improve the quality of applications, even if the data is of low quality. Therefore,

there is local optimization but no overall optimization if data collection only depends on QoI.

In deep crowd sensing (DCS), there are more detailed requests for the data collection [33–38]. First of all, in this aspect of data quality, the demand of sensing data in target area is  $D$ , which means that the application is of high quality if the collected data is not less than  $D$ . Thus, the extra sensing data that is more than  $D$  is redundant. Hence, the system should collect the certain amount of sensor data meeting the demands and it is important that the collected data have a well coverage while the total amount of the collected data is  $D$ . Coverage generally refers to the case where data is distributed evenly in the data collection range. If the target area of data collection is divided into many subareas and every subarea is not large, the best coverage is that the amount of data collected by each subarea is equal. When the total amount of collected data is certain, if the data collected in some subareas is more than needed, it leads to the lacking of data collected in other subareas. As a result, the QoI is low in the subareas that the collected data is less than needed. As the task mentioned above for monitoring bird migration, data collection is needed in specified monitoring area. If there is no collected data in some subareas, it means that the monitoring of bird migration is not comprehensive. And as the monitoring of haze, if there is no collected data in some subareas, it leads to the disappearance of these subareas of haze in the data report.

On the other hand, the data demand  $D$  of system is changing over the time. In order to reduce the cost of data collection, making sure the data demand  $D$  is to ensure that the collected data can reflect the actual situation. Thus, the data demand  $D$  is often changing in different cases. As the haze monitoring mentioned above, if the weather is stable, the haze of the whole city is stable, and the decrease of actual total collected data has little impact on the high-quality services provided by applications, but it can reduce a lot of cost. And in the case of large changes in the weather, it needs data collection with high frequency and large amount of data to ensure that applications meet the requirements of customers. Obviously, this issue also exists in other applications and it lacks the consideration of these cases in previous studies.

For users, it is different between the users in different subareas, which seriously affect the effectiveness of previous studies. Specifically, there are significant differences in the number of users in different subareas, which leads to a serious shortage of incentive mechanisms that provide all users with the same incentive reward. The main reason is two points as follows: (a) the impact on application quality. Generally speaking, the probability of users involved sensing data is higher when gaining more reward, and, on the contrast, it is the same. Thus, in the case of gaining the same reward, the less subareas where the users are, the less available data samples are, which cannot meet the requirements of applications. And data samples in the subareas where there are more users are so many leading to the data redundancy and it even seriously affects the quality of services provided by applications, which has a negative impact on the development of applications. (b) Because the cost in different subareas is different, it has more damage to the quality of services provided by applications.

In practice, the cost in different subareas is different. The general rule is that the cost of sensing data is low in the subareas where there are adequate infrastructures because the convenience of communication bandwidth, net speed, communication costs, and physical traffic reduce the cost of system in these subareas. On the contrary, at the edge of the subareas where network, communication, network, and traffic are inconvenient, the cost of system is high. At the time, combined with the previous reasons, the number of users is large in the low-cost subareas. Thus, in the same incentive reward strategy, it will result in the rich is the richer, and the poor is the poorer, which further deteriorates the performance of the applications. It can be seen that the previous studies cannot solve the problem better, which results in deep crowd sensing.

To tackle those issues in deep crowd sensing, a Time and Location Correlation Incentive (TLCI) scheme is proposed for deep data gathering in crowdsourcing networks in this paper. The goal of TLCI scheme is to collect the appropriate data and to satisfy the Quality of Information (QoI) requirements of applications with minimum budget as much as possible. In TLCI scheme, the data sensing is fine-grained and the reward for data samples is dynamically changing based on the time and location of data sensing, which can effectively improve the system performance. The main contributions of this paper are summarized as follows.

(1) A Time and Location Correlation Incentive (TLCI) scheme is proposed for deep data gathering in crowdsourcing networks. The TLCI scheme is a fine-grained data gathering mechanism which is more efficient than previous strategies to provide high-quality data gathering with lower cost, because our scheme takes the number of users in different time and locations, the cost for sensing data, and the willingness of participating in sensing data into consideration.

(2) A concept named “QoI satisfaction degree” (QoISD) is introduced in this paper to quantify the degree of how collected sensing data can satisfy whole requirements of applications mainly in terms of data quantity and data coverage. The applications of our proposed algorithm show “QoI satisfaction degree” is more effective in evaluating the overall performance contribution of the data samples to the system.

(3) Two incentive algorithms are proposed to satisfy QoISD with different view. The first algorithm is to ensure that the application gets the specified sensing data to maximize the QoISD. So in the first incentive algorithm, the reward for data sensing is to maximize the QoISD. The second algorithm is to minimize the cost of the system while meeting the sensing data requirement and maximizing the QoISD. So the reward for data sensing is to maximize the QoISD per unit of reward in the second incentive algorithm.

(4) Finally, we compare our proposed scheme with existing schemes via extensive simulations. Experimental results show that our distributed incentive mechanism can successfully attain our aim in this work, which is more suitable in real world. Extensive simulation results well justify the effectiveness of our scheme. The QoISD can be optimized by 81.92%, and the total cost can be reduced by 31.38%.

The rest of the paper is organized as follows. We review related work in Section 2. In Section 3, we describe the system model and formulate the problem of our incentive mechanism. Section 4 present the details of our distributed algorithm for achieving the aim. We evaluate the proposed algorithm via simulations in Section 5. Our paper is concluded in Section 6.

## 2. Related Work

Crowd sensing (CS) refers to sensing a wide range of human activities and their surrounding environment by kinds of sensing devices such as smartphones, iPad, or sensor nodes [32, 39–42]. Because the sensing devices is so vast, it is given a new name crowdsourcing networks (CN) [43]. Due to quick development of mobile sensing devices, collecting the data by these devices is so rich, real-time, and low cost that the traditional way of data collection has a fundamental change [29–32, 43]. As a result, there is a huge change in the network [43]. The center of the network computing migrates from the network center to the network edge [44], which is from cloud computing to edge computing [44], then to big data network [45]. The root of these technological developments lies in the collection of high-quality data [29–32]. Thus, how to obtain high-quality data becomes the focus of research.

Due to the cost of sensing, process, and report data, such as storage, energy, and communication, the incentive mechanisms are a key part of these systems and have been studied extensively [29–32, 43]. The incentive mechanism has been applied to crowd sensing with some specialties that are not available in other applications [43]. First, in CS, the number of reporter is large, and most of them do not have identity authentication features. Thus, it does not have the verification functionality that is available in most interactive systems. In traditional crowdsourcing, such as Amazon Mechanical Turk [44], the crowdsourcing platform publishes their tasks and the reward for subtask (or microtask). The worker is required to submit proof of the task after completing the task and crowdsourcing platform will pay the worker for the promised reward. In this way, through the large number of workers it can achieve the huge task which can be broken down into many microtasks. However, this method is difficult to be implemented in the CS because of the following reason. In CS, the task for data collection is called microtask, in which each data sample is so micro so that if using a complex authentication mechanism, its cost and time consumption are far more than the cost of the task itself. Thus, it is not suitable for CS adopting such certification mechanism. Therefore, an effective incentive mechanism to motivate the users to collect data samples is very important for CS [29].

There are many incentive mechanisms that are designed to attract users to participate in sensing task. This paper divides it into the following categories.

(1) The incentive mechanisms which are for the purpose of certain number of data samples: the incentive mechanism is to motivate the users to collect the data through adjusting the reward. The general process is as follows: we assume the amount of data expected to be collected is  $D$ . It will decrease the reward for each data sample to reduce the cost

of applications while the actual collected data is more than the expected. On the contrary, if the current collected data is less than  $D$ , then the current incentive reward is not enough and the enthusiasm of the users to sense data is not high. At the time, the system improves the enthusiasm and makes the collected data reach  $D$  through increasing the reward for data samples. There are many differences for the specific applications. A game based on services price decision (GSPD) model to depict the process of price competition between Services Organizers (SOs, i.e., applications) and entities (users) as well as internal entities (users), which leads to a Pareto-optimal equilibrium point. The auction incentive mechanism is another incentive mechanism which is applied in CS. A reverse auction-based dynamic price (RADP) incentive mechanism was proposed by Lee and Hoh [46]. And a subset of users with lowest reward (prices) are chosen by the service provider that needs sensing data, and their sensor data are purchased at their bid prices. It shows that the dynamic price incentive mechanism can reduce the incentive cost compared with the fixed price. Another reverse auction-based incentive mechanism was designed by Jaimes et al. [47] which considers the locations of the users, the budget constraints, and the sensed coverage. Their incentive schemes can improve the covered area. The Stackelberg game joint auction-based mechanism is also proposed by Yang et al. [48].

(2) Incentive mechanism based on quality: this kind of research not only considers the quantity and cost of data collection but also takes quality of data collection into account. In this type of research, some evaluation norms of quality of data samples are proposed to evaluate the quality of data samples. In some studies to collect data samples for service composition, the concept of Quality of Contributed Service (QCS) is proposed by Tham and Luo [30] to measure the contribution of collected data to the combined service. In addition, in some applications, such as noise mapping, traffic condition reporting, and environmental impact monitoring, the metric named Quality of Information (QoI) is proposed by Wang et al. [29] to evaluate the quality of the data samples. In such studies, it is generally considered that the number of participants involved sensing data is very large. Hence, the goal of the incentive strategy design is to select the participants with highest level of performance assessment (e.g., QCS and QoI) from these users. At the time, these studies focus on minimizing the overall cost, profit, or energy consumption while selecting the high-quality participants, shown by Liu et al. [49–52].

(3) In addition to the above two types of research, recent studies think that only the reward based incentive mechanism does not necessarily achieve good results. One important reason for this is as follows: these studies did not consider the users and collected data credibility. Thus, the recent researchers have proposed a comprehensive consideration of the QoI of data and the collected data credibility. These studies think that the users' reputation represents the collected data credibility. Therefore, the applications selecting users with high reputation on priority can improve the performance of the system. Such a reputation-based incentive

mechanism can be found in Zhang and Van Der Schaar [53] as well as Wang et al. [54].

### 3. The System Model and Problem Statement

*3.1. System Model.* It is considered that a crowdsourcing sensor data gathering system is made up of a platform in the cloud and a large number of smartphone users connecting to the platform. The applications (data requesters) publish the task on the platform to collect the sensor data. In addition, the smartphone users  $U$  who are possibly to participate in the sensing task are denoted as follows:

$$U = \{u_1, u_2, \dots, u_n\}. \quad (1)$$

The participants  $\mathcal{U}$  represent the users who are satisfied with reward for their sensed data contribution and participating in the sensing task. So the set  $\mathcal{U}$  is the subset of the set  $U$ .

$$\mathcal{U} = \{u_1, u_2, \dots, u_n\}. \quad (2)$$

The participants take a cost when they use their smartphone to collect the data, such as the electricity consumption when opening the sensor or their consumed time, which have an impact on whether to participate in task. In order to formulate the users' behavior in the reality, it is necessary to stimulate the user's cost and the probability of participating in sensing task.

*Definition 1* (participants cost). The sensing time  $t_i$  represents the consumed time of participants  $u_i$  for participating in the sensing task. In the sensing time, participants collect data continuously. However, the cost is different in different circumstance and different time. The cost of the users is involved in the consumption of time and power of phone. For example, the cost is simpler in cities than mountains and daytime in the night. And the cost is different in the same location and time even if the sensing time is same. It is mainly because the difference of cost of consumption of phone power is due to the difference of smartphones and sensors. Thus, the cost function of participants is involved in the time and sensing circumstance.

We assume that we need the sensor data in a place  $L$  and in a time period  $T$ . If the sensed data is evenly distributed on the target area and time, it makes sensed data representative and is of high quality. The targeted area  $L$  and the task's period  $T$  will be divided equally, such as  $L = \{L_1, L_2, \dots, L_n\}$ ,  $T = \{T_1, T_2, \dots, T_m\}$ , which form many grids. The size of the grid is related to applications. In general, the number of grids should be enough, and the grid area should be small enough. In this way, the collected data from participants  $u_i$  could represent the data situation for the grid.  $T_i$  represents a short period of time during which the monitored value of the object is stationary. The data collected by sensor nodes during this time period can represent the physical values for this time period. The granularity of time  $T_i$  division is also related to applications. If the application changes substantially, then it requires that  $T_i$  is small. And if the application changes within

a narrow range, then  $T_i$  can take a larger value. For example, in the case of smooth weather, the temperature in an hour has almost no change; then the time granularity of sensing temperature can take a larger value. In this period of time, any data collected during this time period can represent the temperature during this time period. And when the weather changes substantially,  $T_i$  should take a smaller value so that the temperature value sensed in a very short time period is basically equal and can represent the temperature in this time period. If taking a larger value, the difference of the sensed value is very large in different time. Thus, the sensed value in different time cannot represent the temperature in the time period.

And we use  $c_i^j$  to represent the cost in subperiod  $T_j$  and in subarea  $L_i$  caused by circumstance involved in location and time.  $\mathcal{C}_i^j$  represents the cost matrix in the entire sensing time and area.

$$\mathcal{C}_i^j = \begin{bmatrix} c_1^1 & c_1^2 & \cdots & c_1^m \\ c_2^1 & c_2^2 & \cdots & c_2^m \\ \vdots & \vdots & \ddots & \vdots \\ c_n^1 & c_n^2 & \cdots & c_n^m \end{bmatrix} \quad i \in [1, n], j \in [1, m]. \quad (3)$$

We assume that the cost involved in power consumption has a linear relationship with the sensing time, denoted as  $\alpha_i \tau$ , and  $\tau$  is sensing time. In addition,  $\alpha_i$  is user parameter, which depends on the phone of participant and different between different users. Hence, the cost function  $C_i(\tau)$  of sensing time  $\tau$  for participant  $u_i$  is formulated as

$$C_i(\tau) = \alpha_i \tau + c_i^j. \quad (4)$$

*Definition 2* (the probability of participation). The probability of participation  $\delta_i(p)$  of user  $u_i$  represents the probability that the user  $u_i$  participates in the sensing task when the reward is  $r_i$  for the sensing time  $\tau$ . We assume that the users are rational and strategic so that whether they participate in the task depends on the reward and cost. The probability of participation is increasing as the ratio of economic benefits increases. We use the power function  $\delta(r_i)$  of ratio of economic benefits to represent the probability that user  $u_i$  participate in data collection when the reward is  $r_i$ , which is shown in the following:

$$\delta(r_i) = \begin{cases} 0, & r_i \leq C_i(\tau) \\ \left( \frac{r_i - C_i(\tau)}{C_i(\tau)} \right)^\beta - 1, & C_i(\tau) < r_i \leq 2C_i(\tau) \\ 1, & r_i > 2C_i(\tau). \end{cases} \quad (5)$$

And  $\beta$  is the parameter to be measured.

The set of users in subarea  $L_i$  and subperiod  $T_j$  is denoted as follows:

$$U_i^j = \begin{bmatrix} u_1^1 & u_1^2 & \cdots & u_1^m \\ u_2^1 & u_2^2 & \cdots & u_2^m \\ \vdots & \vdots & \ddots & \vdots \\ u_n^1 & u_n^2 & \cdots & u_n^m \end{bmatrix} \quad i \in [1, n], j \in [1, m], \quad (6)$$

where  $u_i^j = \{u_1, u_2, \dots, u_z\} \mid u_k$  is the smartphone user. And  $u_i^j$  represent the set of users in the subperiod  $T_j$  and subarea  $L_i$ . In addition, we use  $|u_i^j|$  to represent the number of users in the subperiod  $T_j$  and subarea  $L_i$ .

$U_i^j$  in (6) represent the set of users in the subperiod  $T_j$  and subarea  $L_i$ . But all users in a grid not necessarily participate in sensing task. Hence, we use  $\mathcal{U}_i^j$  to represent the users who actually participate in the sensing as follows:

$$\mathcal{U}_i^j = \begin{bmatrix} u_1^1 & u_1^2 & \cdots & u_1^m \\ u_2^1 & u_2^2 & \cdots & u_2^m \\ \vdots & \vdots & \ddots & \vdots \\ u_n^1 & u_n^2 & \cdots & u_n^m \end{bmatrix} \quad i \in [1, n], j \in [1, m], \quad (7)$$

where  $u_i^j = \{u_1, u_2, \dots, u_z\} \mid u_k$  is the participants.  $u_i^j$  represent the set of participants in the subperiod  $T_j$  and subarea  $L_i$ . In addition, we use  $|u_i^j|$  to represent the number of participants in the subperiod  $T_j$  and subarea  $L_i$ .

We assume the reward  $R_i^j$  provided by applications is the metric shown in

$$R_i^j = \begin{bmatrix} r_1^1 & r_1^2 & \cdots & r_1^m \\ r_2^1 & r_2^2 & \cdots & r_2^m \\ \vdots & \vdots & \ddots & \vdots \\ r_n^1 & r_n^2 & \cdots & r_n^m \end{bmatrix} \quad i \in [1, n], j \in [1, m], \quad (8)$$

where  $r_i^j$  represent the reward of participants in the subperiod  $T_j$  and subarea  $L_i$ .

Equation (6) shows the users situation in the subperiod  $T_j$  and subarea  $L_i$ . But all users in a grid not necessarily participate in sensing task. In the incentive reward mechanism shown in (8), according to (5), the evaluated participation situation is shown in

$$\mathcal{W}_i^j = \begin{bmatrix} w_1^1 & w_1^2 & \cdots & w_1^m \\ w_2^1 & w_2^2 & \cdots & w_2^m \\ \vdots & \vdots & \ddots & \vdots \\ w_n^1 & w_n^2 & \cdots & w_n^m \end{bmatrix} \quad i \in [1, n], j \in [1, m], \quad (9)$$

where  $w_i^j = \delta(r_i^j) \times u_i^j$ .

The amount of data expected to collect data is shown in

$$D_i^j = \begin{bmatrix} d_1^1 & d_1^2 & \cdots & d_1^m \\ d_2^1 & d_2^2 & \cdots & d_2^m \\ \vdots & \vdots & \ddots & \vdots \\ d_n^1 & d_n^2 & \cdots & d_n^m \end{bmatrix} \quad i \in [1, n], \quad j \in [1, m], \quad (10)$$

where  $d_i^j = |u_i^j| \tau$

The actual amount of data obtained is shown in

$$\mathbb{D}_i^j = \begin{bmatrix} dl_1^1 & dl_1^2 & \cdots & dl_1^m \\ dl_2^1 & dl_2^2 & \cdots & dl_2^m \\ \vdots & \vdots & \ddots & \vdots \\ dl_n^1 & dl_n^2 & \cdots & dl_n^m \end{bmatrix} \quad i \in [1, n], \quad j \in [1, m], \quad (11)$$

where  $d_i^j = |u_i^j| \tau$

**Definition 3** (QoI satisfaction degree, QoISD). In the crowd-sourcing gathering system, the data requester (i.e., applications) publishes the task on the platform to collect the sensor data. Thus, the quality of sensed is very important for data requester. In order to make the sensed data of high quality, it is obvious that the data should have a good distribution on locations and time so that the data is more representative.

In this paper, a metric named ‘‘QoI satisfaction degree’’ (QoISD) is to quantify how much collected sensing data can satisfy the application’s QoI requirements mainly in terms of data coverage.

We assume that the data demand of application is shown in (12). If the actual data  $d_i^j$  is exactly the same as the data demand  $\mathfrak{D}_i^j$ , it shows that the incentive mechanism is efficient, which means there is no redundant data. It is the best situation, and, at the time, the QoISD is highest.

$$\mathfrak{D}_i^j = \begin{bmatrix} \mathfrak{D}_1^1 & \mathfrak{D}_1^2 & \cdots & \mathfrak{D}_1^m \\ \mathfrak{D}_2^1 & \mathfrak{D}_2^2 & \cdots & \mathfrak{D}_2^m \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{D}_n^1 & \mathfrak{D}_n^2 & \cdots & \mathfrak{D}_n^m \end{bmatrix} \quad i \in [1, n], \quad j \in [1, m]. \quad (12)$$

However, in practice, the incentive mechanism is difficult to make  $d_i^j = \mathfrak{D}_i^j \mid \forall i \in [1, n], \quad j \in [1, m]$ . Therefore, in this paper, we use  $\mathfrak{E}^j$  in (13) to represent the difference between the collected data and the data demand, which is QoISD. Obviously, when QoISD  $\mathfrak{E}^j \rightarrow 0$ , it is the best situation.

$$\mathfrak{E}^j = \sum_{k=1}^n (d_k^j - \mathfrak{D}_k^j)^2. \quad (13)$$

If the application requires the same amount of data for each grid, which is a constant  $\mathfrak{D}$ , then (13) is transformed into the following:

$$\mathfrak{E}^j = \sum_{k=1}^n (d_k^j - \mathfrak{D})^2. \quad (14)$$

### 3.2. Problem Statement

(1) *Sensor Data to Meet the Requirement.* Sensing data to meet the applications requirement contains two aspects of meaning: (a) considering the total data requirement is  $D_{\text{total}}^j$  in subperiod  $T_j$ . Therefore, the total amount of data collected in the target area must be greater than  $D_{\text{total}}^j$ , as in (16):

$$\sum_{i=1}^n d_i^j \geq D_{\text{total}}^j \mid \forall j \in [1, m]. \quad (15)$$

(b) The amount of data collected for each grid  $L_i$  should be greater than the amount of data required for application, which is shown in

$$d_i^j \geq \mathfrak{D}_i^j \mid \forall i \in [1, n], \quad j \in [1, m]. \quad (16)$$

(2) *Maximize the Quality of Data.* As mentioned above, in order to get high quality of sensor data, we should make the data’s distribution best based on time and locations, which means optimizing the QoISD, as follows:

$$\min \quad (\mathfrak{Q}^j) = \min \left( \sum_{k=1}^n (d_k^j - \mathfrak{D}_k^j)^2 \right) \mid \forall j \in [1, m]. \quad (17)$$

(3) *Minimize the Cost for Collection Data.* The aim of incentive mechanism is to minimize the cost of system and optimize the quality of the collected data. Therefore, an important goal of Time and Location Correlation Incentive (TLCI) scheme is to minimize the cost for data collection.

$$\min \quad (\mathfrak{C}^j) = \min \left( \sum_{k=1}^n (r_k^j u_k^j) \right) \mid \forall j \in [1, m]. \quad (18)$$

Thus, the overall goal of TLCI scheme is as follows:

$$\begin{aligned} \min \quad (\mathfrak{Q}^j) &= \min \left( \sum_{k=1}^n (d_k^j - \mathfrak{D}_k^j)^2 \right) \mid \forall j \in [1, m], \\ \min \quad (\mathfrak{C}^j) &= \min \left( \sum_{k=1}^n (r_k^j u_k^j) \right) \mid \forall j \in [1, m], \end{aligned} \quad (19)$$

$$\text{s.t.} \quad \sum_{i=1}^n d_i^j \geq D_{\text{total}}^j \mid \forall j \in [1, m],$$

$$d_i^j \geq \mathfrak{D}_i^j \mid \forall i \in [1, n], \quad j \in [1, m].$$

## 4. Incentive Mechanism Design

In order to state the parameter of this paper clearly, the main notions introduced in this paper can be found in Parameters’ Description.

As shown in Figure 2, we propose an incentive mechanism in this work to get enough sensor data which is to optimize the QoISD. In our incentive mechanism, the platform sends the task’s information to every user, and

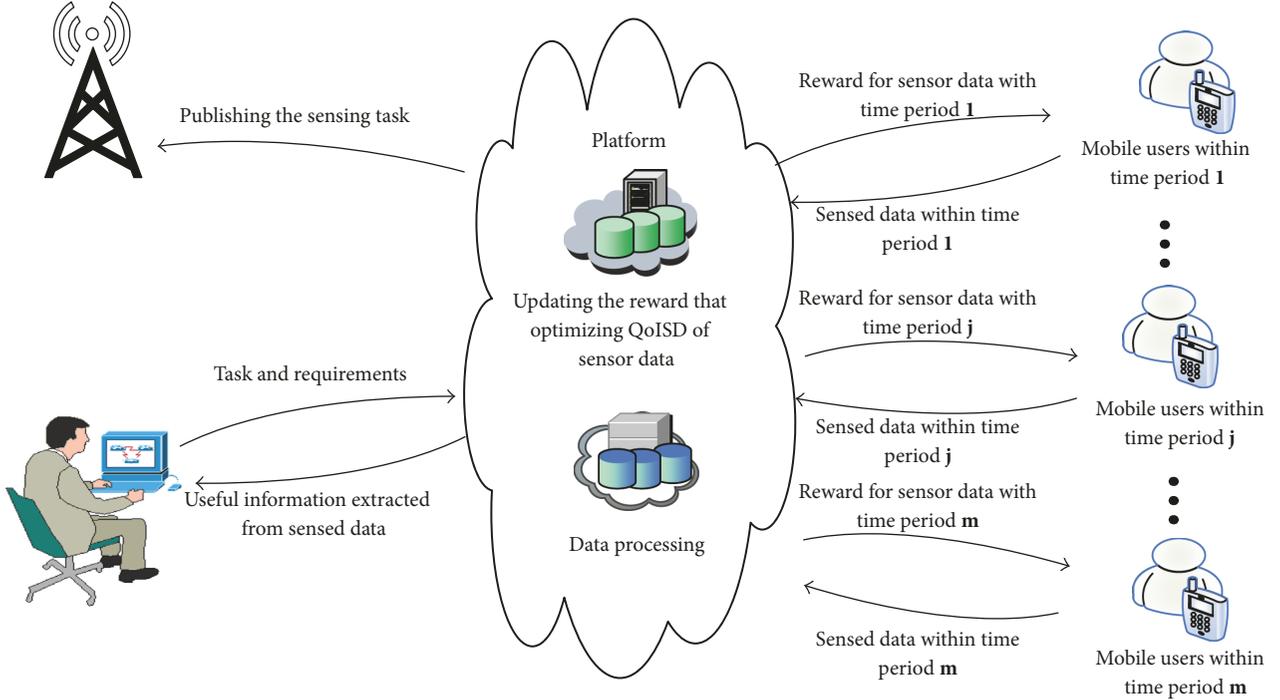


FIGURE 2: An overview of our incentive mechanism.

whether users participate in the sensing task depends on their own cost and the published reward. We assume that the total lasting time period of sensing task is divided into some time periods. And in every subperiod, the target area where data is expected to be obtained is divided into some subareas and the platform publishes a new reward for participants' contribution, which is different for participants in different subareas to optimize the QoISD. It is repeated until the sensing task is completed.

In the incentive mechanism, we use two reward allocation algorithms to optimize the QoISD and reduce the cost of system and both of the algorithms are based on greedy strategies. One algorithm is to solve the problem of optimizing the QoISD when attaining the task's requirement. The other algorithm is improved on the basis of the first algorithm, which is to take the less cost on rewarding the participants and has a better speed than the former. The two incentive algorithms are discussed as follows.

**4.1. Distributed Algorithm for Optimizing QoISD.** The aim in this work is to find a better way of reward allocation so that the system could gain enough high-quality sensed data. The aim in first algorithm is to optimize the QoISD while meeting data requirements that the sensed data in subperiod  $T^j$  reach the data demand  $D_{\text{total}}^j$ , which is formulated as

$$\begin{aligned} \min \quad & (\mathcal{E}^j) = \min \left( \sum_{k=1}^n (d_k^j - \mathfrak{D}_k^j)^2 \right) \mid \forall j \in [1, m] \\ \text{s.t.} \quad & \sum_{i=1}^n d_i^j \geq D_{\text{total}}^j \mid \forall j \in [1, m]. \end{aligned} \quad (20)$$

The algorithm mainly motivates the smartphone users to participate in sensing task through adjusting the reward allocation for users in every subarea  $R^j = \{r_1^j, r_2^j, \dots, r_n^j\}$ . We first analyze the optimal results of sensed data. If the total amount of sensor data needed for applications is  $D_{\text{total}}^j$ , and the amount of sensed data in every subarea is equal which is formulated as  $\mathcal{A}^j = \{\mathfrak{D}_1^j, \mathfrak{D}_2^j, \dots, \mathfrak{D}_n^j\}$ , it requires  $\mathfrak{D}_i^j \triangleq \mathfrak{D}_k^j \mid \forall i, k \in [1, n]$ . In this case, the optimal incentive mechanism is to adjust the reward vector  $R^j$  so that making the amount of sensed data in every subarea is  $D_{\text{total}}^j/n$  exact and the value of QoISD is zero which is optimal. Similarly, if there is no requirement that  $\mathfrak{D}_i^j \triangleq \mathfrak{D}_k^j \mid \forall i, k \in [1, n]$ , and the data demand is  $\mathcal{A}^j = \{\mathfrak{D}_1^j, \mathfrak{D}_2^j, \dots, \mathfrak{D}_n^j\}$ , we assume the actual sensed data is  $D^j = \{d_1^j, d_2^j, \dots, d_n^j\}$ . The optimal incentive strategy is to adjust the reward vector  $R^j$  to optimize the QoISD the same as above. Obviously, it should increase the reward in the subarea where sensed data is less than demand and the difference of both is rather large to motivate the users' participation and makes the QoISD optimal. On the contrary, it optimizes the QoISD decreasing the reward in the subarea where sensed data is more than demand. As shown in (21),  $\sigma_i^j$  represents that deviation rate between actual sensed data and demand in subarea  $L_i$  and subperiod  $T_j$ .

$$\sigma_i^j = \frac{(\mathfrak{D}_i^j - d_i^j)}{\mathfrak{D}_i^j}. \quad (21)$$

Obviously, when  $\sigma_i^j$  is positive it means the actual sensed data is less than demand and it needs to raise the reward  $r_i^j$ . In

addition, the larger  $\sigma_i^j$ , the more amount the lacked data and the increased reward  $r_i^j$ . On the contrary, when  $\sigma_i^j$  is negative, it means the actual sensed data is more than demand and it should decrease the reward  $r_i^j$  to reduce the cost of the system.

Our incentive algorithm is to find the optimal reward allocation for every subarea. The mechanism algorithm iteratively raises the reward to the users in the subarea that could optimize QoISD  $\mathfrak{C}^j$  until the sensed data is enough to the sensing system's demand and the QoISD is optimal. The detailed descriptions are provided as follows.

*Step 1.* First, according the information of sensing task, we initialize the reward of every subarea to  $r_0$  and calculate the new amount of sensor data  $D_{\text{real}}^j$ .

*Step 2.* Second, we iteratively adjust the reward vector to optimize the QoISD  $\mathfrak{C}^j$ . In each iteration, the algorithm adjusts the reward in a subarea which makes QoISD optimal. We use  $\xi$  to mark the target subarea and  $r_\xi$  to mark the increased reward, which is negative when the actual sensed data in target area is more than demand. In addition,  $\varepsilon_{\min}$  represents the optimal QoISD currently and the algorithm updates the value of  $\varepsilon_{\min}$  in every iteration. Hence, in the end of every iteration, it updates the reward  $r_\xi^j$  in target subarea and the amount of sensed data. In addition, the algorithm ignores the subarea where there is no increased amount of sensor data, as shown in Steps (19)-(20).

*Step 3.* The sensed data  $D_{\text{real}}^j$  is updated in every iteration, and the algorithm ends when  $D_{\text{real}}^j$  reaches the data demand  $D_{\text{total}}^j$ . In the end of this incentive algorithm, it returns the reward vector, which makes the QoISD optimal.

The proposed algorithm is given in Algorithm 1.

*4.2. Distributed Algorithm for Optimizing the QoISD and Reducing the Cost.* Based on Algorithm 1, in this subsection, we propose another algorithm for taking the platform's cost and QoISD into account. The aim in Algorithm 2 is to optimize the QoISD and reduce the cost of system while meeting data requirements that the sensed data within subperiod  $T^j$  reaches the data demand  $D_{\text{total}}^j$ , which is shown in

$$\begin{aligned} \min \quad & (\mathfrak{C}^j) = \min \left( \sum_{k=1}^n (d_k^j - \mathfrak{D}_k^j)^2 \right) \mid \forall j \in [1, m] \\ \min \quad & (\mathfrak{C}^j) = \min \left( \sum_{k=1}^n (r_k^j \mathbf{u}_k^j) \right) \mid \forall j \in [1, m] \\ \text{s.t.} \quad & \sum_{i=1}^n d_i^j \geq D_{\text{total}}^j \mid \forall j \in [1, m] \\ & d_i^j \geq \mathfrak{D}_i^j \mid \forall i \in [1, n], j \in [1, m]. \end{aligned} \quad (22)$$

It is apparent that we could solve it by weighted sum method since it is a multiobjective optimization problem. By

selecting scalar weights parameters  $\lambda$  and  $1-\lambda$  for  $\mathfrak{C}^j$  and  $\mathfrak{C}^j$ , the target function is shown in the following:

$$\begin{aligned} \min \quad & (\lambda \mathfrak{C}^j + (1-\lambda) \mathfrak{C}^j) \mid \forall j \in [1, m] \\ \text{s.t.} \quad & \sum_{i=1}^n d_i^j \geq D_{\text{total}}^j \mid \forall j \in [1, m] \\ & d_i^j \geq \mathfrak{D}_i^j \mid \forall i \in [1, n], j \in [1, m], \end{aligned} \quad (23)$$

where  $0 \leq \lambda \leq 1$ . And we could adjust the value of  $\lambda$  to achieve our purpose.

The idea of Algorithm 1 is to raise the reward to the users in the subarea that could optimize QoISD  $\mathfrak{C}^j$  iteratively until the sensed data is enough to the sensing system's demand and the QoISD is optimal. In order to reduce the cost of system, the reward should be allocated to the cost-effective subarea, where system could gain more sensor data with low cost. For each subarea  $L_k$ , we use the  $\Upsilon_k$  to represent how cost-effective it is, which is the ratio of increased data to increased cost and formulated as

$$\Upsilon_k = \frac{(\delta(r_k + r_\nabla) - \delta(r_k)) \mathbf{u}_k^j \tau}{\mathbf{u}_k^j \times (\delta(r_k + r_\nabla) \times (r_k + r_\nabla) - \delta(r_k) x r_k)}. \quad (24)$$

Hence, there is a main alternation in Algorithm 1:

$$\varepsilon_{\min} = -\lambda \Upsilon_k + (1-\lambda) \sum_{i=1}^n (d_i^j - \mathfrak{D}_i^j)^2. \quad (25)$$

The proposed algorithm is given in Algorithm 2.

## 5. Performance Analysis and Optimization

*5.1. Methodology and Setup.* In this section, we compare our two algorithms, Time and Location Correlation Incentive Algorithm 1 (TLCI1) and Time and Location Correlation Incentive Algorithm 2 (TLCI2) with the pricing based fixed price algorithm (PPA). The main character of this algorithm (PPA) is that the participants get the same reward for the same amount of sensor data. We assume the smartphone users are rational and strategic so that whether they participate in sensing task depends on their cost and reward. In order to show how they are similar or different, the parameters of these are the same.

We first study the situation of sense data and the cost of system for rewarding the participants at different locations in 5.2. Then in 5.3, we study the sensed data and the cost in different time in two situations and make a performance comparison of the three algorithms. In addition to the performance comparison based on different time and locations, we also study the quality of sensed data and the stability of algorithms in 5.4 and 5.5, respectively.

The default values of all parameters in our model are set as follows. The default units of reward and sensing time are dollar and minute, respectively. And the user-specific parameters  $\alpha_i$  are uniformly distributed random values within (1, 3). In addition, according to the data obtained from the

```

Input: The data demand  $\mathcal{D}^j = \{\mathfrak{D}_1^j, \mathfrak{D}_2^j, \dots, \mathfrak{D}_n^j\}$ ;
Output: The new vector of reward  $R^j = \{r_1^j, r_2^j, \dots, r_n^j\}$ .
(1) For each  $r_i^j \mid i \in [1, n]$ 
(2)    $r_i^j = r_0$  // to initialize the reward of every sub-area to  $r_0$ 
(3) End for
(4)  $D_{\text{real}}^j = \sum_{k=1}^n d_k^j$ 
(5) While  $D_{\text{real}}^j < D_{\text{total}}^j$  Do
(6)    $\mathcal{C}_u^j = \sum_{k=1}^n (d_k^j - \mathfrak{D}_k^j)^2$ ;
      //  $\mathcal{C}_u^j$  is the QoISD before the update
(7)    $\xi = 0$  //to mark the selected grid which is not determined
(8)    $\varepsilon_{\min} = +\infty$  //to store the minimum QoISD
(9)   For  $k = 1$  to  $n$  Do //each grid  $L_k$ 
(10)    If  $(d_k^j - \mathfrak{D}_k^j) < 0$ 
(11)      $r_{\nabla} = +\zeta$ 
      //to increase the reward when the sensed data doses not meet the demand
(12)    Else
(13)      $r_{\nabla} = -\zeta$ 
(14)     //to decrease the reward when the sensed data doses not meet the demand
(15)    End if
(16)     $\overline{d}_k^j = d_k^j + (\delta(r_k + r_{\nabla}) - \delta(r_k)) u_k^j \tau$ 
      //the sensed data in grid  $L_k$  after updating the reward
(17)     $d_k^j \leftarrow \overline{d}_k^j$ 
(18)     $\mathcal{C}_{\text{new}}^j = \sum_{i=1}^n (d_i^j - \mathfrak{D}_i^j)^2$ 
      //  $\mathcal{C}_{\text{new}}^j$  is the updated QoISD
(19)    If  $\mathcal{C}_{\text{new}}^j = \mathcal{C}_u^j$  //if there is no improved QoISD
(20)     Goto Step (10)
(21)    End if
(22)    If  $\mathcal{C}_{\text{new}}^j < \varepsilon_{\min}$ 
(23)      $\varepsilon_{\min} = \mathcal{C}_{\text{new}}^j$ 
(24)      $\xi = k$ ; //to select the  $L_k$  that makes QoISD optimal
(25)      $r_{\xi} = r_{\nabla}$ 
(26)    End if
(27)  End for
(28)   $r_{\xi}^j = r_{\xi}^j + r_{\xi}$ 
(29)   $D_{\text{real}}^j = \sum_{k=1}^n d_k^j$  //to update the sensed data
(30) End while
(31) Return the new reward  $R^j$ 

```

ALGORITHM 1: The QoISD optimization algorithm.

questionnaire survey, the probability parameter  $\beta$  is 1.745. All simulations under the same setting are repeated ten times to get the average values.

**5.2. The Performance of Sensed Data and Cost at Different Locations.** We first study the performance of the three algorithms under different situations. We mainly study two scenarios, the active and the inactive period of users. In the active period, there are large number of users distributed in the target area. On the contrary, fewer people are involved in the sensing task in the inactive period. The study on two scenarios is shown as follows.

**5.2.1. The Active Period of Users.** In the active period of users, there are large number of users that could participate in the sensing task and have plenty of free time to collect the sensor data. We assume that the platform's demand is 4000 (Mb) and

the number of users of every subarea is uniformly distributed random values within (50, 100). The results are as follows. As shown in Figures 3, 4, and 5, the range of sensed data in the subareas in TLCI1 and TLCI2 are similar and the range of sensed data in our two TLCI algorithms varying from 45 to 55 is less than that in PPA, which is varying from 40 to 75 and has a negative impact on the quality of sensed data. The result shows that the variance of sensed data can be reduced by 68.85% and 52.50% in TLCI1 and TLCI2, respectively. We also stimulate the participants' reward in Figures 6 and 7 about our two TLCI algorithms. Compared to the same reward for every user in PPA, the reward varying from 3.7 to 7 is different for each user in our two TLCI algorithms.

**5.2.2. The Inactive Period of Users.** In the inactive period of users, the users do not have enough time to participate in the sensing task. Thus, the number of users is less than the

```

Input: The data demand  $\mathcal{A}^j = \{\mathfrak{D}_1^j, \mathfrak{D}_2^j, \dots, \mathfrak{D}_n^j\}$ ;
Output: The new vector of reward  $R^j = \{r_1^j, r_2^j, \dots, r_n^j\}$ .
(1) For each  $r_i^j \mid i \in [1, n]$ 
(2)    $r_i^j = r_0$  // to initialize the reward of every sub-area to  $r_0$ 
(3) End for
(4)  $D_{\text{real}}^j = \sum_{k=1}^n d_k^j$ 
(5) While  $D_{\text{real}}^j < D_{\text{total}}^j$  Do
(6)    $\mathfrak{C}_u^j = \sum_{k=1}^n (d_k^j - \mathfrak{D}_k^j)^2$ ;
      //  $\mathfrak{C}_u^j$  is the QoISD before the update
(7)    $\xi = 0$  //to mark the selected grid which is not determined
(8)    $\varepsilon_{\text{min}} = +\infty$  //to store the minimum QoISD
(9)   For  $k = 1$  to  $n$  Do //each grid  $L_k$ 
(10)    If  $(d_k^j - \mathfrak{D}_k^j) < 0$ 
(11)      $r_{\nabla} = +\varsigma$ 
           //to increase the reward when the sensed data doses not meet the demand
(12)    Else
(13)      $r_{\nabla} = -\varsigma$ 
           //to decrease the reward when the sensed data doses not meet the demand
(14)    End if
(15)     $\overline{d}_k^j = d_k^j + (\delta(r_k + r_{\nabla}) - \delta(r_k)) u_k^j \tau$ 
           //the sensed data in grid  $L_k$  after updating the reward
(16)     $d_k^j \leftarrow \overline{d}_k^j$ 
(17)     $\mathfrak{C}_{\text{new}}^j = \sum_{i=1}^n (d_i^j - \mathfrak{D}_i^j)^2$ 
           //  $\mathfrak{C}_{\text{new}}^j$  is the updated QoISD
(18)    If  $\mathfrak{C}_{\text{new}}^j = \mathfrak{C}_u^j$  //if there is no improved QoISD
(19)     Goto Step (10)
(20)    End if
(21)    If  $\mathfrak{C}_{\text{new}}^j < \varepsilon_{\text{min}}$ 
(22)      $\varepsilon_{\text{min}} = -\lambda Y_k + (1 - \lambda) \mathfrak{C}_{\text{new}}^j$ 
(23)      $\xi = k$ ; //to select the  $L_k$  that make QoISD optimal
(24)      $r_{\xi} = r_{\nabla}$ 
(25)    End if
(26)    End for
(27)     $r_{\xi}^j = r_{\xi}^j + r_{\xi}$ 
(28)     $D_{\text{real}}^j = \sum_{k=1}^n d_k^j$  //to update the sensed data
(29) End while
(30) Return the new reward  $R^j$ 

```

ALGORITHM 2: The QoISD optimization and minimizing the cost algorithm.

number of users in the active period. We assume that the number of users in every subarea is uniformly distributed random values within (30, 60). As shown in Figures 8, 9, and 10, we could draw the same conclusion as that in active period. In the inactive period, the variance of sensed data can be reduced by 90.85% and 79.06% in TLCI1 and TLCI2, respectively. In addition, the situation of reward allocation is shown in Figures 11 and 12 that the reward allocation of two algorithms is similar.

*5.3. The Performance of Sensed Data and Cost at Different Time.* It is considered that the variation of the number of users in the whole day has an objective law, which is shown as Figure 13. The users prefer to participate in the sensing task in the daytime and the number of users in daylight is far less

than in daytime. For a more comprehensive consideration, we consider the whole QoISD that involves the data at different times rather than the QoISD at a time period. Therefore, when the impact of incentive reward on the whole QoISD is too little, it will reduce the data demand and increase to other time periods according to the number of users. It is to ensure that the collected data of the whole day meet the demand.

In this subsection, we study the performance of three algorithms in two situations: the random distribution of users and the concentrated distribution of users. In situation of the random distribution of users, the users are evenly distributed at different subareas. We assume that the total users at a time is  $n_r$ , and the users at a subarea are uniformly distributed random values within  $(n_r/m-30)$  and  $(n_r/m+30)$  when there are  $m$  subregions at the time. In the concentrated distribution of users, the 50% of users focus on the 30% of central area.

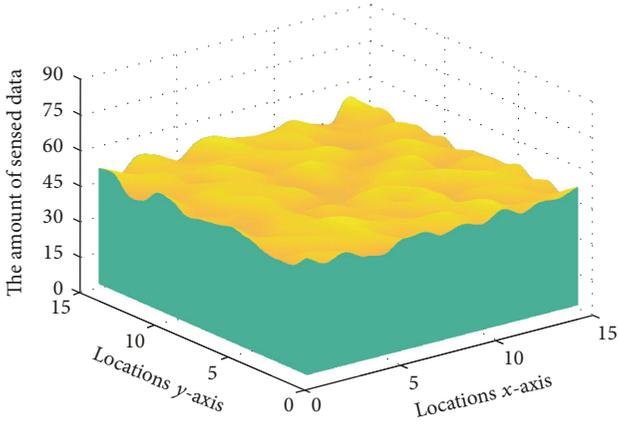


FIGURE 3: The distribution of sensed data of TLCI1 in active period of users.

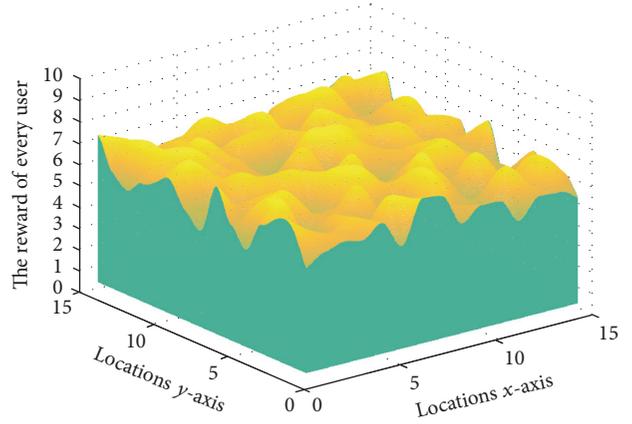


FIGURE 6: The distribution of users' reward of TLCI1 in active period of users.

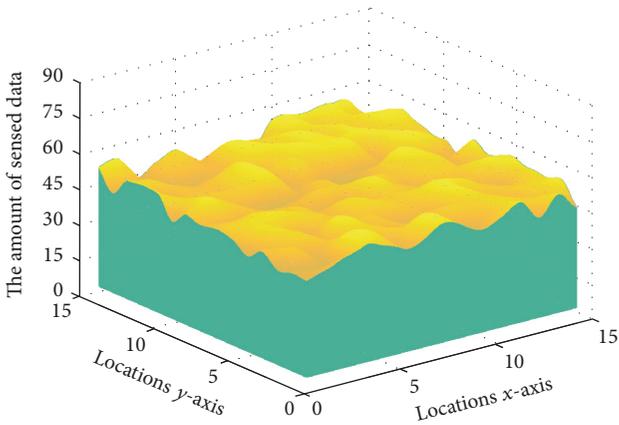


FIGURE 4: The distribution of sensed data of TLCI2 in active period of users.

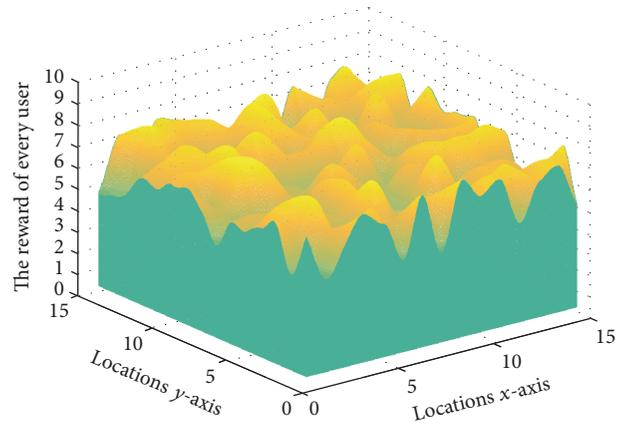


FIGURE 7: The distribution of users' reward of TLCI2 in active period of users.

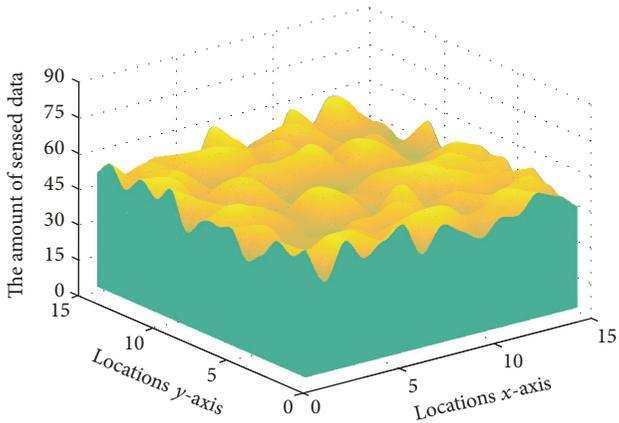


FIGURE 5: The distribution of sensed data of PPA in active period of users.

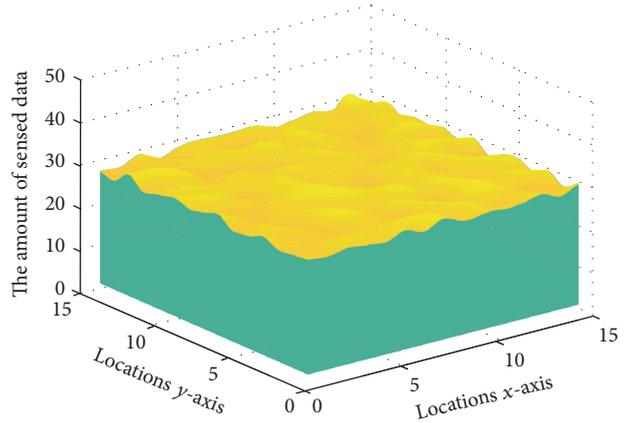


FIGURE 8: The distribution of sensed data of TLCI1 in inactive period of users.

5.3.1. *The Random Distribution of Users.* In the random distribution of users, the users are evenly distributed at the locations at a time. In this subsection, we stimulate the average reward for users and total reward budget for users' contribution at different time, as shown in Figures 14 and 16. It is obvious that, compared to fixed reward in PPA, our

two TLCI algorithms have a better adaptability, adjusting the user's reward according to the number of users. The overall trend is that the average of reward is declining with more users as shown in Figure 14, which can be interpreted as the payment in exchange for participation. As a result, the number of participants and the total reward at different time

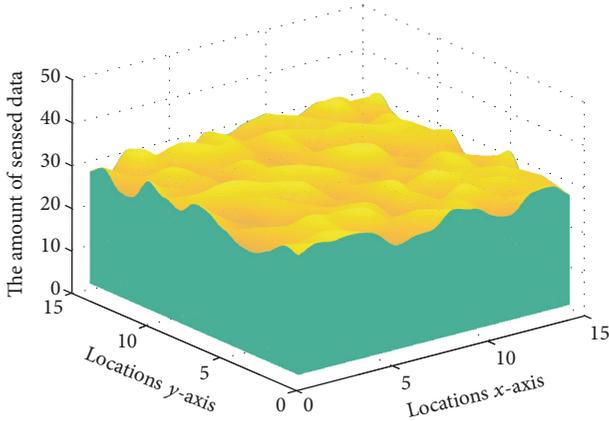


FIGURE 9: The distribution of sensed data of TLCI2 in inactive period of users.

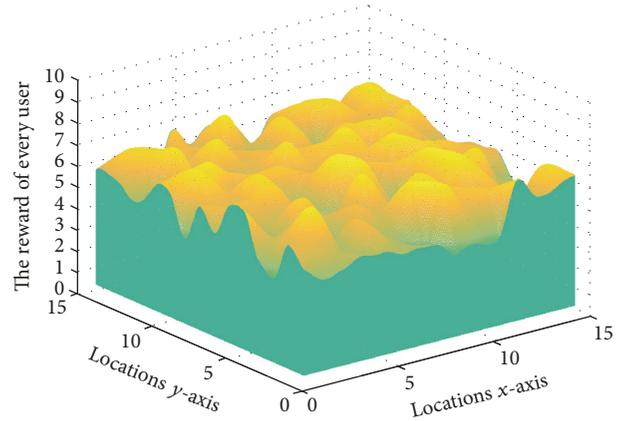


FIGURE 12: The distribution of users' reward of TLCI2 in inactive period of users.

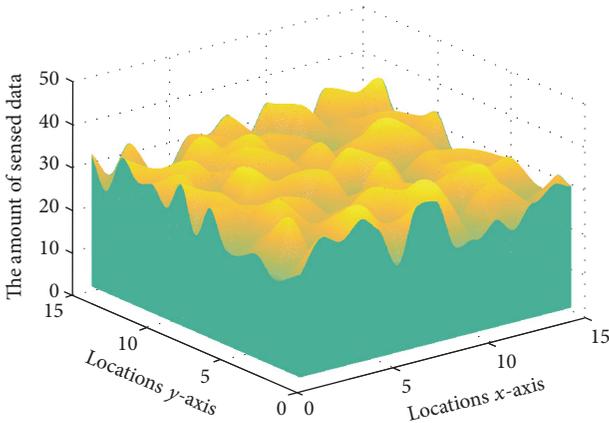
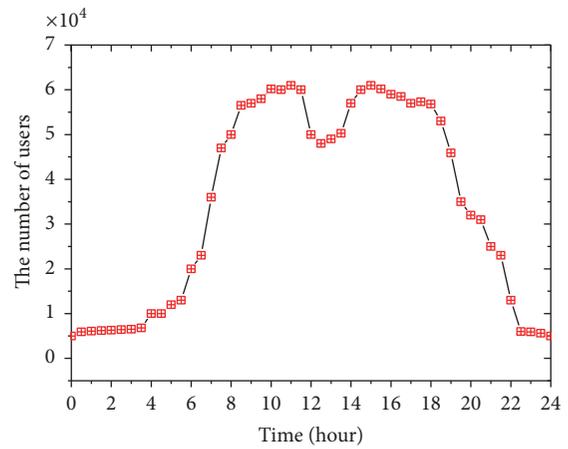


FIGURE 10: The distribution of sensed data of PPA in inactive period of users.



—■— The variation of users

FIGURE 13: The number of users at different time.

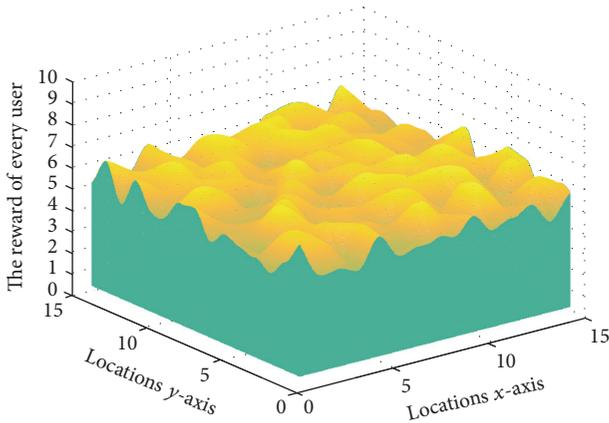
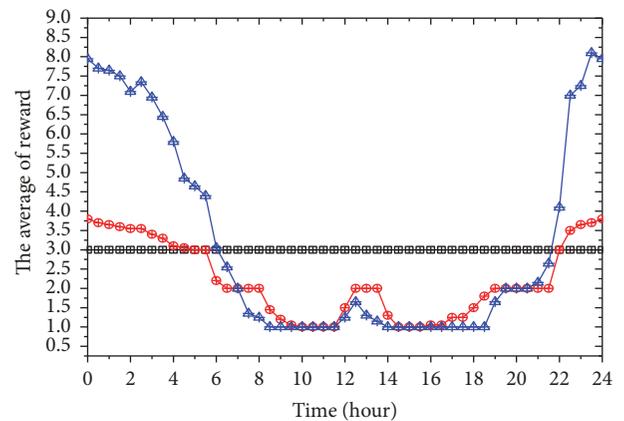


FIGURE 11: The distribution of users' reward of TLCI1 in inactive period of users.



—■— PPA  
—●— GCMA2  
—▲— GCMA1

FIGURE 14: The average of reward for a user at different time.

are shown in Figures 15 and 16, in which total reward and total participants grow up in number when there are few users so that it can collect more sensor data. In addition, we can see the total cost of system is reduced by 19.97% and 42.79% in TCL11 and TCL12, respectively. Figure 17 shows the cumulative total reward over the time, which proves the

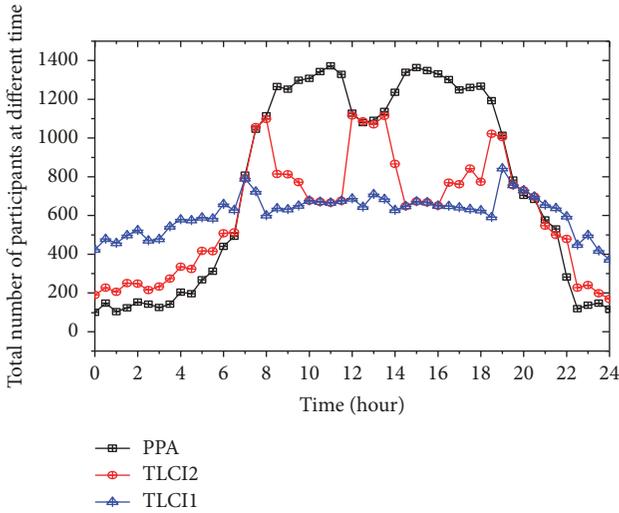


FIGURE 15: The number of participants at different time.

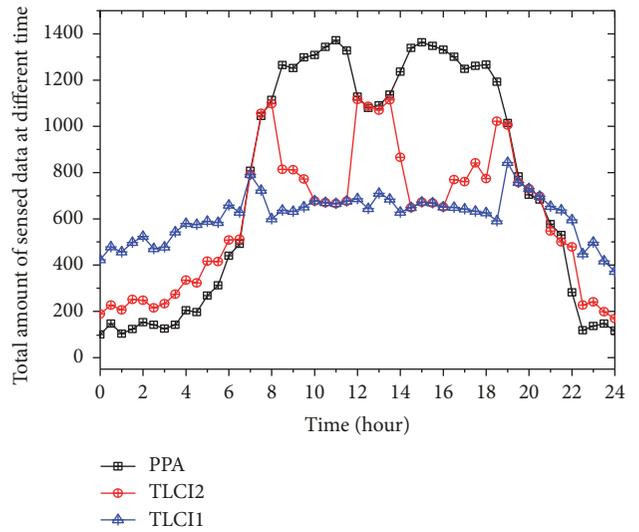


FIGURE 18: The total amount of sensed data.

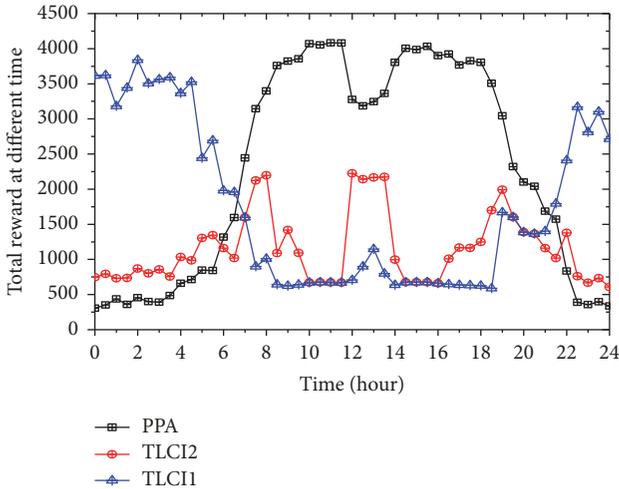


FIGURE 16: The total reward at different time.

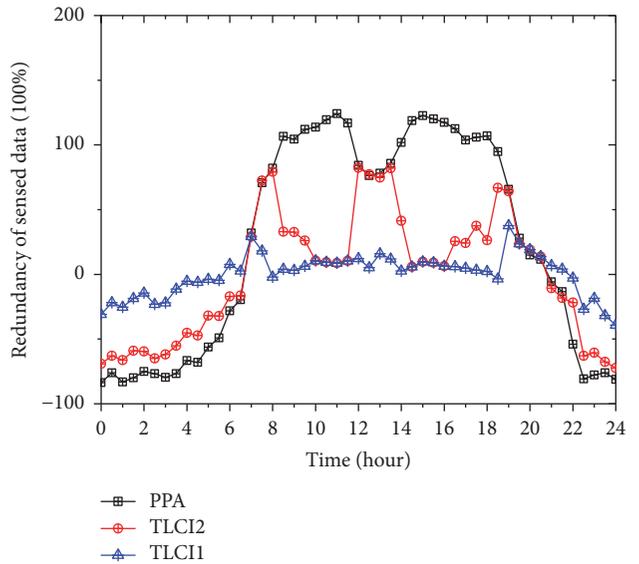


FIGURE 19: The redundancy of sensed data.

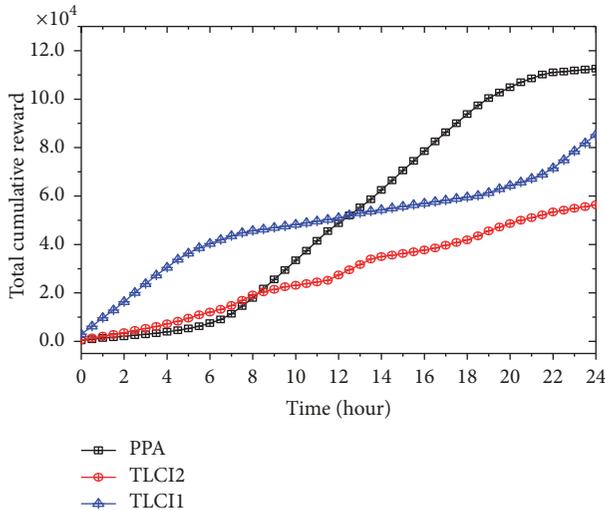


FIGURE 17: The total cumulative reward over time.

effectiveness of our two algorithms and the TLCI2 could reduce the cost. In addition to the total reward, we also study the sensed data over the time in Figure 18. Figure 19 shows that the redundancy of sensed data over time that the percentage of the difference between actual and expected sensed data to expected data. It is similar to the sensed data over time. And in Figure 20, we calculate the cumulative sensed data over time and we can see the sensed data in PPA exceeds the demand. It shows the PPA algorithms have a bad adaptability compared with TLCI algorithms.

**5.3.2. The Concentrated Distribution of Users.** In the situation of the concentrated distribution of users, the 50% of users focus on the 30% of central area. In this subsection, we stimulate the average reward of users and total reward budget for users' contribution at different time, as shown in Figures

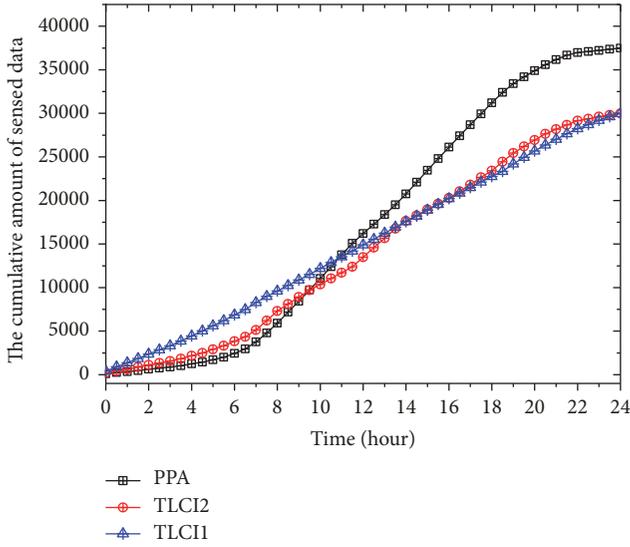


FIGURE 20: The cumulative amount of sensed data.

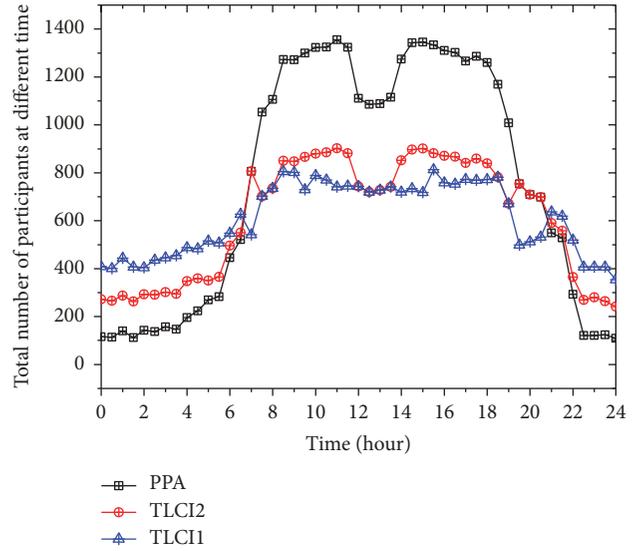


FIGURE 22: The number of participants at different time.

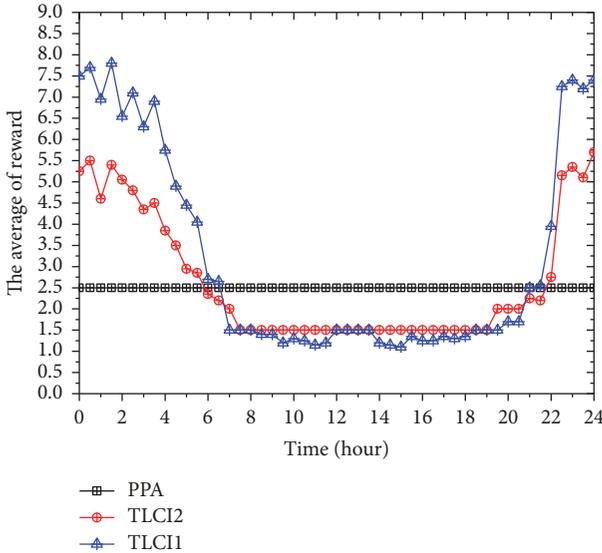


FIGURE 21: The average of reward for a user at different time.

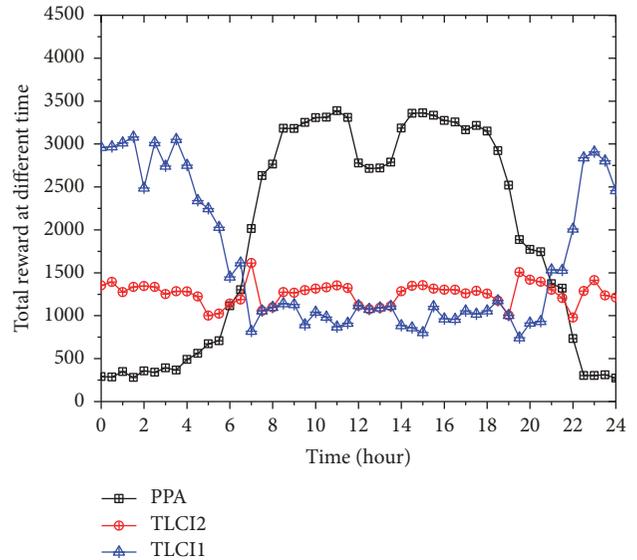


FIGURE 23: The total reward at different times.

21 and 23. It is obvious that compared to fixed reward in PPA, our two TLCI algorithms have a better adaptability, adjusting the user’s reward according to the number of users. The overall trend is that the average of reward is declining with the more users as shown in Figure 21, which can be interpreted as the payment in exchange for participation. As a result, the number of participants and the total reward at different time are shown in Figures 22 and 23, in which total reward and total participants grow up in number when there are few users so that it can collect more sensor data. In addition, we can see the total cost of system is reduced by 34.20% and 14.89% in TLCI1 and TLCI2, respectively. Figure 24 shows the cumulative total reward over the time, which proves the effectiveness of our two algorithms and the TLCI2 could reduce the cost. In addition to the total reward,

we also study the sensed data over the time in Figure 25. Figure 26 shows that the redundancy of sensed data over time that the percentage of the difference between actual and expected sensed data to expected data. It is similar to the sensed data over time. And in Figure 27, we calculate the cumulative sensed data over time and we can see the sensed data in PPA exceeds the demand because the fixed reward in PPA has a bad adaptability compared with our two TLCI algorithms.

5.3.3. *The Comparison of Three Algorithms.* In this subsection, we compare the performance of three algorithms in two situations mentioned above. The results are shown as follows.

We use situation A and situation B to represent the situation of the random distribution and the concentrated

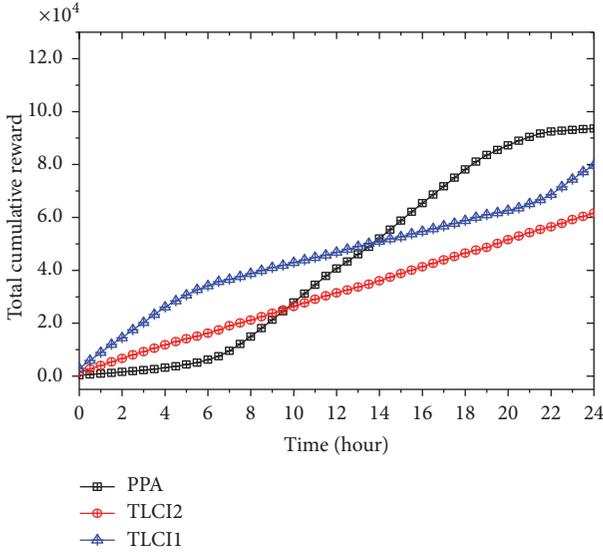


FIGURE 24: The total cumulative reward over time.

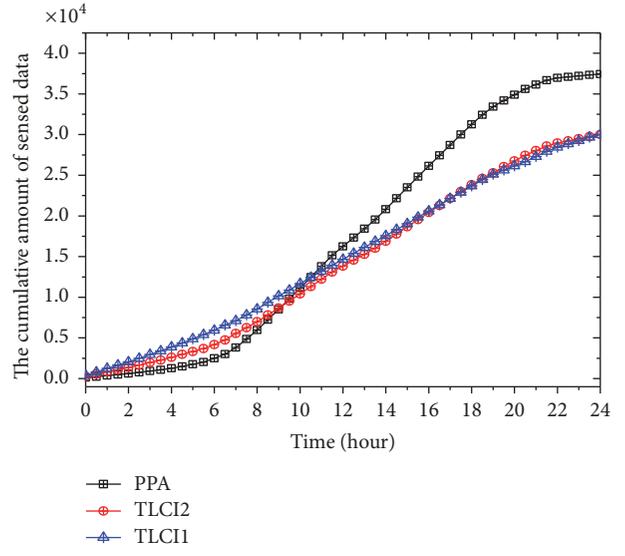


FIGURE 27: The cumulative amount of sensed data.

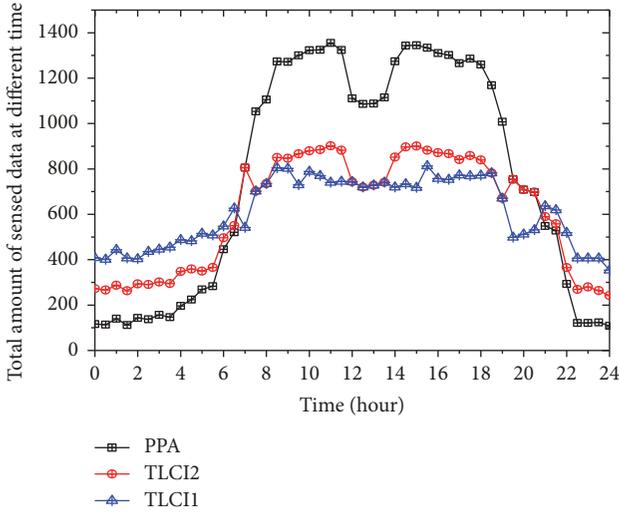


FIGURE 25: The total amount of sensed data.

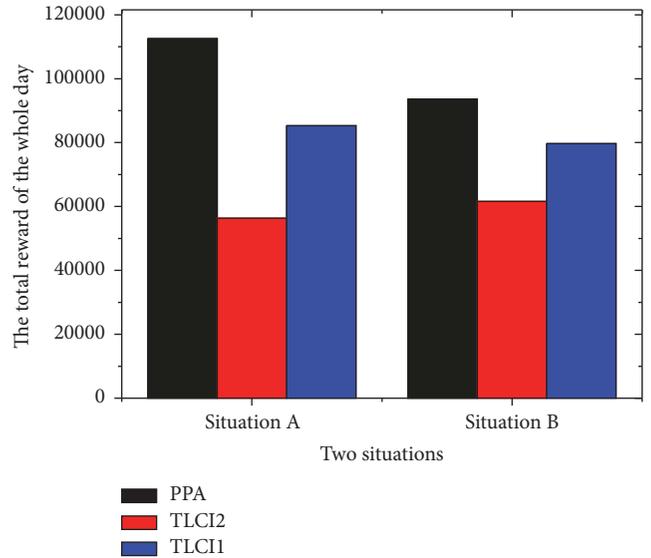


FIGURE 28: The comparison of total reward of the whole day in two situations.

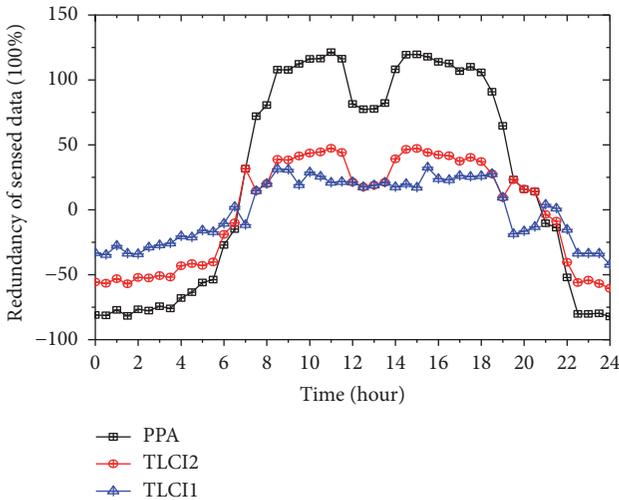


FIGURE 26: The redundancy of sensed data.

distribution, respectively. We first calculate the total reward and total sensed data in the whole day and make a comparison in Figures 28 and 29, respectively. According to Figure 28, we can see that our two TLCI algorithms can reduce the cost for rewarding the participants compared with fixed reward and TLCI2 has a better performance. On average of cost in two TLCI algorithms, it can reduce the cost of system by 37.07% in situation A and 24.54% in situation B, respectively. And Figure 30 shows that the TLCI algorithms can make the cost valuable rather than collecting redundant data.

Then, we make a comparison of total redundancy of sensed data of whole day in two situations and the result is shown in Figure 30. It shows the effectiveness of our two TLCI algorithms and the TLCI1 has a better performance

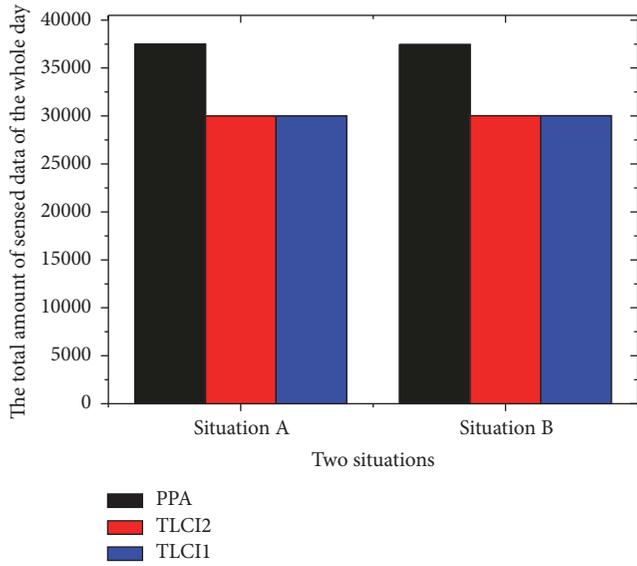


FIGURE 29: The comparison of total sensed data of the whole day in two situations.

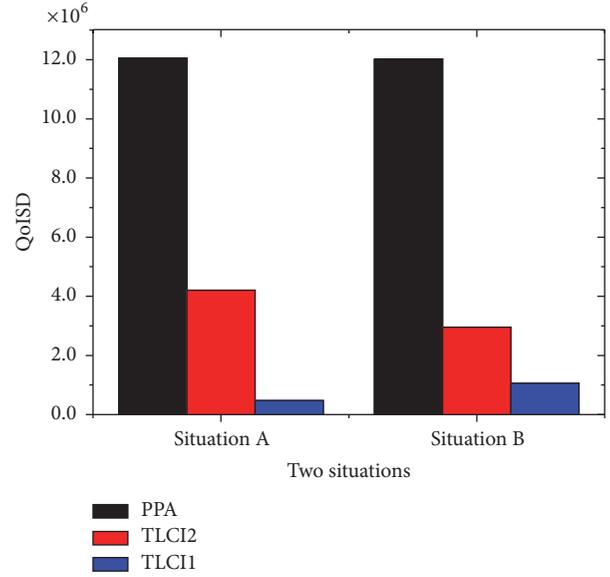


FIGURE 31: The comparison of QoISD of the whole day in two situations.

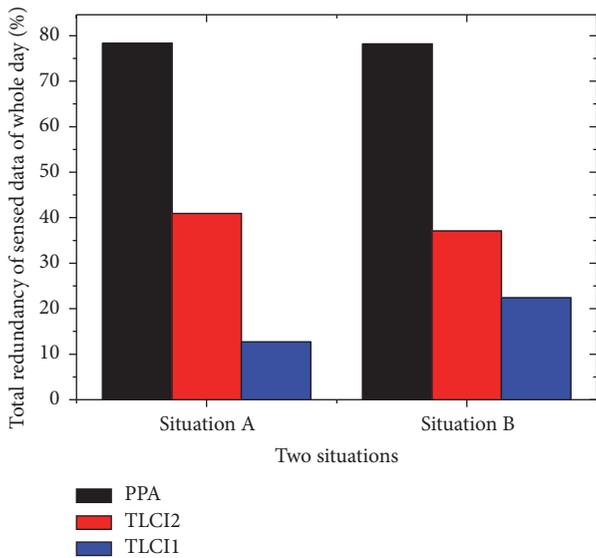


FIGURE 30: The comparison of total redundancy of sensed data of whole day in two situations.

compared with TLCI1. In addition, we make a comparison of performance on optimizing the QoISD in three algorithms, and the result is shown in Figure 31. On the average of QoISD in two situations, the QoISD can be optimized by 81.92%, and the total cost can be reduced by 31.38% with considering two TLCI algorithms.

**5.4. The Quality of Sensed Data.** In this subsection, we use the difference between the amount of actual and expected sensed data to evaluate the performance of our two algorithms (TLCI) and PPA. As above, we evaluate it in two situations, the active and inactive period of users, and the results are shown as follows. Figures 32, 33, and 34 show the difference

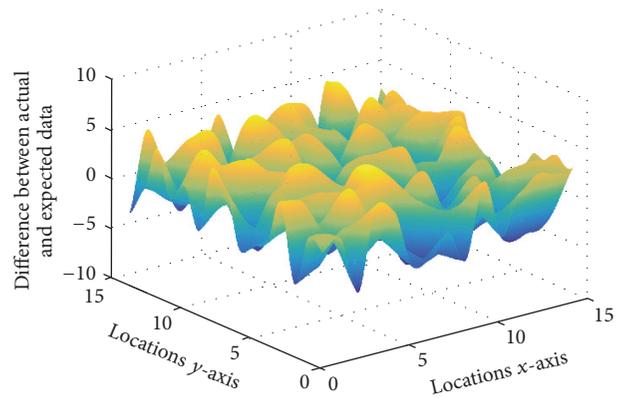


FIGURE 32: The difference between actual and expected sensed data in active period in TLCI1.

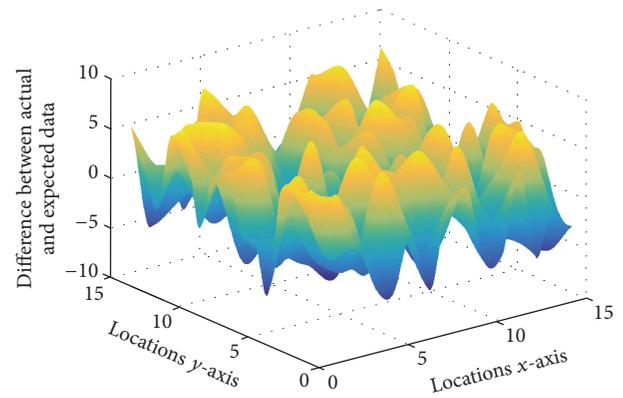


FIGURE 33: The difference between actual and expected sensed data in active period in TLCI2.

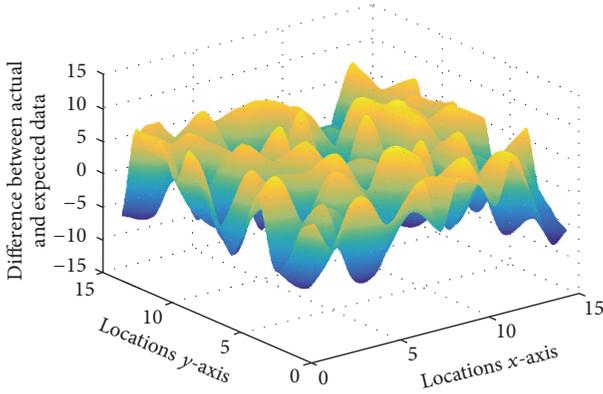


FIGURE 34: The difference between actual and expected sensed data in active period in PPA.

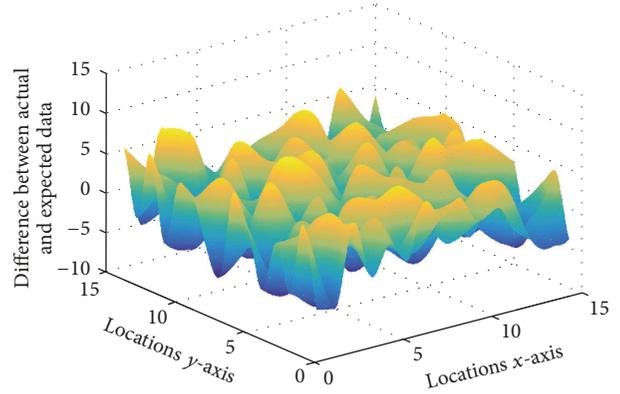


FIGURE 37: The difference between actual and expected sensed data in inactive period in PPA.

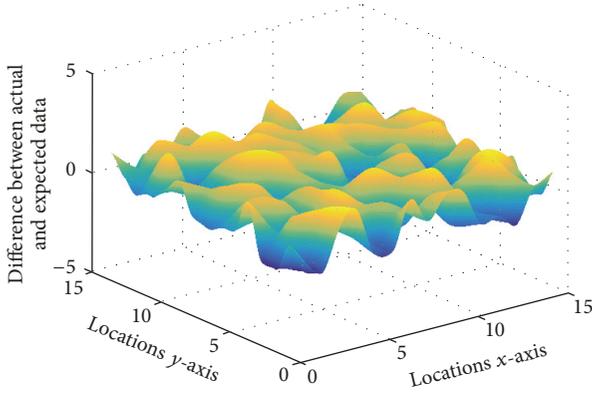


FIGURE 35: The difference between actual and expected sensed data in inactive period in TLCI1.

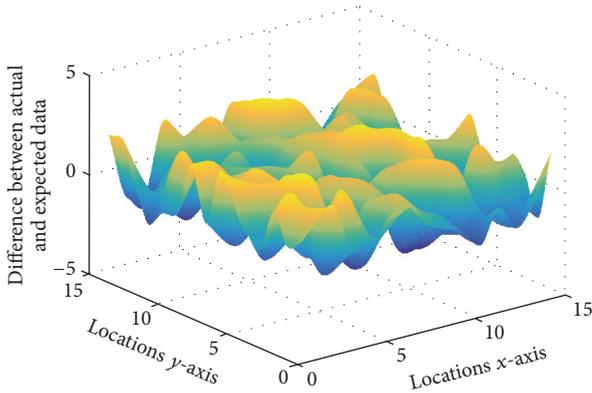
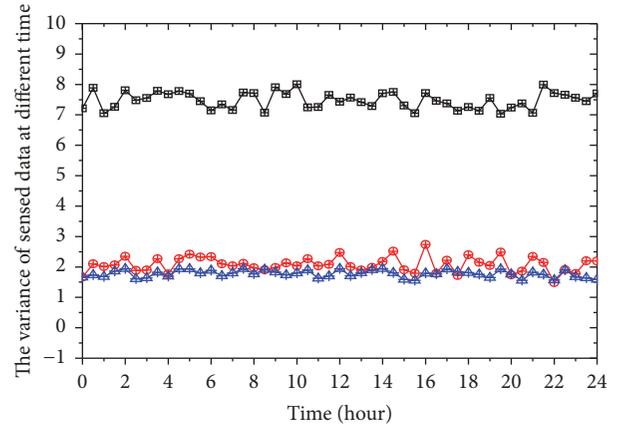


FIGURE 36: The difference between actual and expected sensed data in inactive period in TLCI2.

between actual and expected sensor data in three algorithms in active period. According to the figures, we can see that the difference between actual and expected sensed data in two TLCI algorithms is less than that in PPA algorithm and the performance of two TLCI algorithms is similar, where the difference is varying from  $-2.1$  to  $2.45$  in TLCI1, varying from  $-3.34$  to  $3.12$  in TLCI2, and varying from  $-9.12$  to  $9.27$  in PPA. And Figures 35, 36, and 37 show the same



—■— PPA  
—●— TLCI2  
—▲— TLCI1

FIGURE 38: The variance of sensed data over time.

conclusion as above. At last in this subsection, we calculate the variance of sensed data at different time in the situation of the distribution of users, which is shown in Figure 38. The variance of sensed data in two TLCI algorithms is less than that in PPA algorithms, where two TLCI algorithms have the similar performance on optimizing the quality of data. We can see the advantage of our two TLCI algorithms, making the sensed data evenly distributed at different locations.

**5.5. The Stability of TLCI Algorithms.** In this subsection, we mainly study the stability of TLCI algorithms. We first study the number of iterations by changing the value of  $\lambda$ . The result is shown in Figure 39, and we can see the number of iterations is declining with the increasing  $\lambda$  in TLCI2, which proves the better performance in speed in TLCI2 compared with TLCI1. Then we study the stability of TLCI algorithms. It is known that our algorithms will reach a steady state after limited number of iterations.

In Figures 40 and 41, we compare the total reward and the sensed data over the iteration in our two algorithms and it proves that the TLCI2 has the better performance in terms

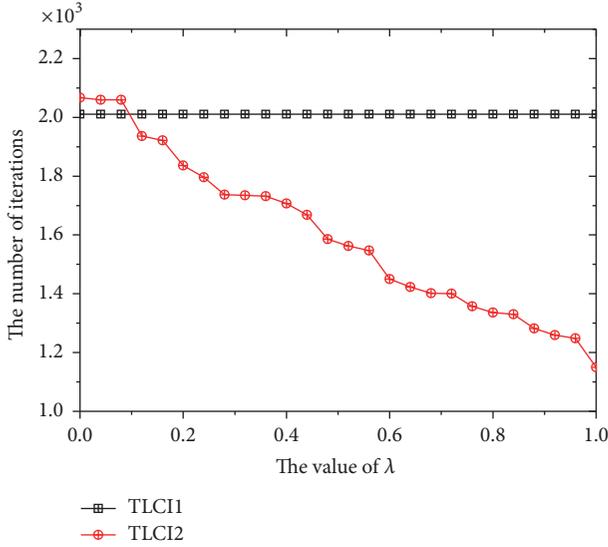


FIGURE 39: The number of iterations by changing the value of  $\lambda$ .

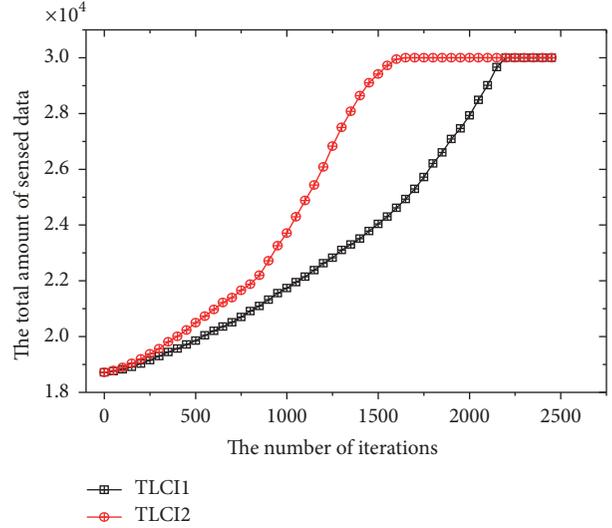


FIGURE 41: The total amount of sensed data over the iteration.

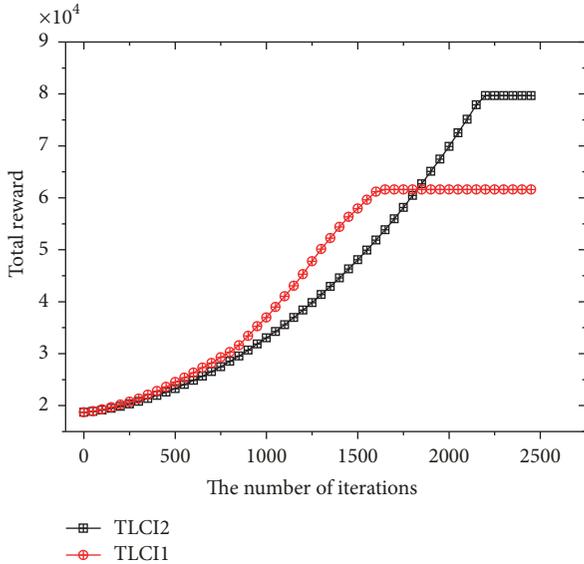


FIGURE 40: The total cumulative reward over the iteration.

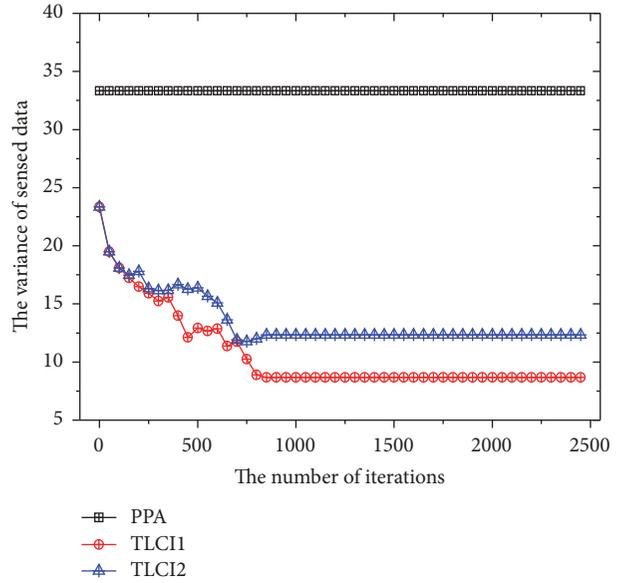


FIGURE 42: The variance of sensed data over the iteration.

of speed of algorithm and cost reduction, respectively. And we can see that the total cost of system after stabilization in TLCI2 is lower than that in TLCI1. In the end of this subsection, we calculate the variance of sensed data over the iteration as shown in Figure 42. The variance of sensed data at different subregions will stabilize after limited iterations and the stable variance of sensed data in TLCI2 is less than that in TLCI1, which means the same conclusion as the previous experiment.

## 6. Conclusion

This paper has focused on the problem of optimizing the quality of sensor data and collecting enough sensor data in a sensor data gathering system. In such a system, the cost

of each user depends on his circumstance, which has an impact on users whether to participate in the sensing task. The platform makes different reward for task participants in different region and time period so that it could make the distribution of sensed data on location and time best. Such a sensing system has been developed to realize such applications, such as NoiseTube [55].

In this paper, we have proposed two effective incentive mechanism algorithms for motivating smartphone users to participate in the smartphone sensing. Both of the algorithms are based on greedy strategies. One algorithm is to solve the problem of optimizing the QoISD. The other algorithm is improved on the basis of the first algorithm, which is to take the less cost on rewarding the participants. Extensive simulations have been performed and the results have confirmed

that the two algorithms can achieve our goal and the latter has a better speed than the former. QoISD can be optimized by 81.92%, and the total cost can be reduced by 31.38%.

## Parameters' Description

|                        |   |
|------------------------|---|
| $\alpha_i, c_i^j$ :    | The parameter in the cost   |
| $u_n$ :                | The smartphone users  |
| $u_n$ :                | The participants in sensing task  |
| $C_i(\tau)$ :          | The cost of user $i$ when sensing time is $\tau$  |
| $r_i$ :                | The reward of participant $i$   |
| $\delta(r_i)$ :        | The participation probability of a user participating in sensing task                         |
| $u_i^j$ :              | The number of smartphone users in subperiod $T_j$ and subarea $L_i$                           |
| $u_i^j$ :              | The number of task participants in subperiod $T_j$ and subarea $L_i$                          |
| $r_i^j$ :              | The reward for every participant in subperiod $T_j$ and subarea $L_i$                         |
| $u_i^j$ :              | The number of participants in subperiod $T_j$ and subarea $L_i$ based on probability estimate |
| $d_i^j$ :              | The estimated amount of sensor data in subperiod $T_j$ and subarea $L_i$                      |
| $d_n^2$ :              | The actual amount of sensor data in subperiod $T_j$ and subarea $L_i$                         |
| $\mathfrak{D}_i^j$ :   | The actual demand of sensor data in subperiod $T_j$ and subarea $L_i$                         |
| $\mathfrak{E}^j$ :     | QoISD in subperiod $T_j$  |
| $C^j$ :                | The cost of system for rewarding participants in subperiod $T_j$                              |
| $D_{\text{total}}^j$ : | The total demand of sensor data in subperiod $T_j$  |
| $\lambda$ :            | The parameter in the target function  |
| $\varepsilon_k$ :      | The cost-efficiency of subarea $k$ .  |

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61772554, 61370229, 61572526, and 61572528), the National Basic Research Program of China (973 Program) (2014CB046305), the Science and Technology Projects of Guangdong Province, China (2016B010109008 and 2016B030305004), and the Science and Technology Projects of Guangzhou Municipality, China (201604010003 and 201604010054).

## References

- [1] K. Zhang, X. Liang, R. Lu, and X. Shen, "Sybil attacks and their defenses in the internet of things," *IEEE Internet of Things Journal*, vol. 1, no. 5, pp. 372–383, 2014.
- [2] Z. Sheng, H. Wang, C. Yin, X. Hu, S. Yang, and V. C. M. Leung, "Lightweight management of resource-constrained sensor devices in internet of things," *IEEE Internet of Things Journal*, vol. 2, no. 5, pp. 402–411, 2015.
- [3] T. Qiu, R. Qiao, and D. Wu, "EABS: an event-aware backpressure scheduling scheme for emergency internet of things," *IEEE Transactions on Mobile Computing*, 2017.
- [4] T. Qiu, Y. Zhang, D. Qiao, X. Zhang, M. L. Wymore, and A. K. Sangaiah, "A robust time synchronization scheme for industrial internet of things," *IEEE Transactions on Industrial Informatics*.
- [5] A. Liu, Z. Chen, and N. N. Xiong, "An adaptive virtual relaying set scheme for loss-and-delay sensitive WSNs," *Information Sciences*, vol. 424, pp. 118–136, 2018.
- [6] Z. Ning, F. Xia, N. Ullah, X. Kong, and X. Hu, "Vehicular social networks: enabling smart mobility," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 16–55, 2017.
- [7] X. Hu, V. C. Leung, Y. Kwok et al., "SAFeD): a crowd-cloud co-design approach to situation-aware music delivery for drivers," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 12, no. 1, 2015.
- [8] H. Li, D. Liu, Y. Dai, T. H. Luan, and S. Yu, "Personalized search over encrypted data with efficient and secure updates in mobile clouds," *IEEE Transactions on Emerging Topics in Computing*, vol. PP, no. 99, 2016.
- [9] T. Wang, J. Zeng, M. Z. Bhuiyan et al., "Trajectory privacy preservation based on a fog structure for cloud location services," *IEEE Access*, vol. 5, pp. 7692–7701, 2017.
- [10] J. Gao, P. Liu, X. Kang, L. Zhang, and J. Wang, "PRS: parallel relaxation simulation for massive graphs," *The Computer Journal*, vol. 59, no. 6, pp. 848–860, 2016.
- [11] K. Zhang, X. Liang, R. Lu, and X. Shen, "PIF: a personalized fine-grained spam filtering scheme with privacy preservation in mobile social networks," *IEEE Transactions on Computational Social Systems*, vol. 2, no. 3, pp. 41–52, 2015.
- [12] J. Xu, X. Liu, M. Ma, A. Liu, T. Wang, and C. Huang, "Intelligent aggregation based on content routing scheme for cloud computing," *Symmetry*, vol. 9, no. 10, 2017.
- [13] Z. Su, Q. Xu, M. Fei, and M. Dong, "Game theoretic resource allocation in media cloud with mobile social users," *IEEE Transactions on Multimedia*, vol. 18, no. 8, pp. 1650–1660, 2016.
- [14] K. Zhang, K. Yang, X. Liang, Z. Su, X. Shen, and H. H. Luo, "Security and privacy for mobile healthcare networks: from a quality of protection perspective," *IEEE Wireless Communications Magazine*, vol. 22, no. 4, pp. 104–112, 2015.
- [15] D. Zeng, L. Gu, L. Lian, S. Guo, H. Yao, and J. Hu, "On cost-efficient sensor placement for contaminant detection in water distribution systems," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2177–2185, 2016.
- [16] C. Hu, M. Li, D. Zeng, and S. Guo, "A survey on sensor placement for contamination detection in water distribution systems," *Wireless Networks*, 2016.
- [17] J. Gui and K. Zhou, "Flexible adjustments between energy and capacity for topology control in heterogeneous wireless multi-hop networks," *Journal of Network and Systems Management*, vol. 24, no. 4, pp. 789–812, 2016.
- [18] Z. Ning, X. Hu, Z. Chen et al., "A cooperative quality-aware service access system for social internet of vehicles," *IEEE Internet of Things Journal*.
- [19] M. Zhou, M. Zhao, A. Liu, M. Ma, T. Wang, and C. Huang, "Fast and efficient data forwarding scheme for tracking mobile targets in sensor networks," *Symmetry*, vol. 9, no. 11, p. 269, 2017.

- [20] T. Li, Y. Liu, L. Gao, and A. Liu, "A cooperative-based model for smart-sensing tasks in fog computing," *IEEE Access*, vol. 5, pp. 21296–21311, 2017.
- [21] M. Marjanović, L. Skorin-Kapov, K. Pripuzić, A. Antonić, and I. P. Žarko, "Energy-aware and quality-driven sensor management for green mobile crowd sensing," *Journal of Network and Computer Applications*, vol. 59, pp. 95–108, 2016.
- [22] A. Antonić, M. Marjanović, K. Pripuzić, and I. P. Žarko, "A mobile crowd sensing ecosystem enabled by CUPUS: cloud-based publish/subscribe middleware for the internet of things," *Future Generation Computer Systems*, vol. 56, pp. 607–622, 2016.
- [23] Y. Gao, Y. Chen, and K. J. R. Liu, "On cost-effective incentive mechanisms in microtask crowdsourcing," *IEEE Transactions on Computational Intelligence and AI in Games*, vol. 7, no. 1, pp. 3–15, 2015.
- [24] T. Wang, J. Zeng, Y. Lai et al., "Data collection from WSNs to the cloud based on mobile Fog elements," *Future Generation Computer Systems*, 2017.
- [25] T. Wang, Y. Li, G. Wang, J. Cao, M. Z. Bhuiyan, and W. Jia, "Sustainable and efficient data collection from WSNs to cloud," *IEEE Transactions on Sustainable Computing*, 2017.
- [26] H. Gao, C. H. Liu, W. Wang et al., "A survey of incentive mechanisms for participatory sensing," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 918–943, 2015.
- [27] X. Hu, T. H. S. Chu, V. C. M. Leung, E. C.-H. Ngai, P. Kruchten, and H. C. B. Chan, "A survey on mobile social networks: applications, platforms, system architectures, and future research directions," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1557–1581, 2014.
- [28] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 140–150, 2010.
- [29] Y. Wang, Z. Cai, G. Yin, Y. Gao, X. Tong, and G. Wu, "An incentive mechanism with privacy protection in mobile crowdsourcing systems," *Computer Networks*, vol. 102, pp. 157–171, June 2016.
- [30] C.-K. Tham and T. Luo, "Quality of contributed service and market equilibrium for participatory sensing," *IEEE Transactions on Mobile Computing*, vol. 14, no. 4, pp. 829–842, 2015.
- [31] S. Reddy, D. Estrin, and M. Srivastava, "Recruitment framework for participatory sensing data collections," in *Pervasive Computing*, vol. 6030 of *Lecture Notes in Computer Science*, pp. 138–155, Springer, Berlin, Germany, 2010.
- [32] Z. Song, C. H. Liu, J. Wu, J. Ma, and W. Wang, "QoI-aware multitask-oriented dynamic participant selection with budget constraints," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4618–4632, 2014.
- [33] X. Liu, "A novel transmission range adjustment strategy for energy hole avoiding in wireless sensor networks," *Journal of Network and Computer Applications*, vol. 67, pp. 43–52, 2016.
- [34] H. Xin and X. Liu, "Energy-balanced transmission with accurate distances for strip-based wireless sensor networks," *IEEE Access*, vol. 5, pp. 16193–16204, 2017.
- [35] J. Gui, L. Hui, and N. Xiong, "A game-based localized multi-objective topology control scheme in heterogeneous wireless networks," *IEEE Access*, vol. 5, pp. 2396–2416, 2017.
- [36] J. Gui and J. Deng, "A topology control approach reducing construction cost for lossy wireless sensor networks," *Wireless Personal Communications*, vol. 3, pp. 2173–2202, 2017.
- [37] D. Zeng, P. Li, S. Guo, T. Miyazaki, J. Hu, and Y. Xiang, "Energy Minimization in Multi-Task Software-Defined Sensor Networks," *IEEE Transactions on Computers*, vol. 64, no. 11, pp. 3128–3139, 2015.
- [38] T. Qiu, A. Zhao, F. Xia, W. Si, and D. O. Wu, "ROSE: robustness strategy for scale-free wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 5, pp. 2944–2959, 2017.
- [39] X. Hu, T. H. S. Chu, H. C. B. Chan, and V. C. M. Leung, "Vita: a crowdsensing-oriented mobile cyber-physical system," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 1, pp. 148–165, 2013.
- [40] G. Yang, S. He, and Z. Shi, "Leveraging crowdsourcing for efficient malicious users detection in large-scale social networks," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 330–339, 2017.
- [41] S. He, D. Shin, J. Zhang, J. Chen, and P. Lin, "An exchange market approach to mobile crowdsensing: pricing, task allocation, and walrasian equilibrium," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 4, pp. 921–934, 2017.
- [42] X. Duan, C. Zhao, S. He, P. Cheng, and J. Zhang, "Distributed algorithms to compute walrasian equilibrium in mobile crowdsensing," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 5, pp. 4048–4057, 2017.
- [43] A. Kittur, E. H. Chi, and B. Suh, "Crowdsourcing user studies with Mechanical Turk," in *Proceedings of the 26th Annual CHI Conference on Human Factors in Computing Systems, (CHI '08)*, pp. 453–456, Florence, Italy, April 2008.
- [44] K. Lu, S. Wang, L. Xie, Z. Wang, and M. Li, "A dynamic reward-based incentive mechanism: Reducing the cost of P2P systems," *Knowledge-Based Systems*, vol. 112, pp. 105–113, 2017.
- [45] Z. Chen, M. Ma, X. Liu, A. Liu, and M. Zhao, "Reliability improved cooperative communication over wireless sensor networks," *Symmetry*, vol. 9, no. 10, p. 209, 2017.
- [46] J.-S. Lee and B. Hoh, "Sell your experiences: a market mechanism based incentive for participatory sensing," in *Proceedings of the 8th IEEE International Conference on Pervasive Computing and Communications, (PerCom '10)*, pp. 60–68, April 2010.
- [47] L. G. Jaimes, I. Vergara-Laurens, and M. A. Labrador, "A location-based incentive mechanism for participatory sensing systems with budget constraints," in *Proceedings of the 10th IEEE International Conference on Pervasive Computing and Communications, (PerCom '12)*, pp. 103–108, March 2012.
- [48] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (MobiCom '12)*, pp. 173–184, August 2012.
- [49] C. H. Liu, B. Zhang, X. Su, J. Ma, W. Wang, and K. K. Leung, "Energy-aware participant selection for smartphone-enabled mobile crowd sensing," *IEEE Systems Journal*, vol. PP, no. 99, 2015.
- [50] C. H. Liu, J. Fan, P. Hui, J. Wu, and K. K. Leung, "Toward QoI and energy efficiency in participatory crowdsourcing," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4684–4700, 2014.
- [51] C. H. Liu, J. Fan, P. Hui, J. Wu, and K. K. Leung, "Toward QoI and energy efficiency in participatory crowdsourcing," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 10, pp. 4684–4700, 2015.
- [52] C. H. Liu, B. Zhang, X. Su, J. Ma, W. Wang, and K. K. Leung, "Energy-aware participant selection for smartphone-enabled mobile crowd sensing," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1435–1446, 2017.

- [53] Y. Zhang and M. Van Der Schaar, "Robust reputation protocol design for online communities: a stochastic stability analysis," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 5, pp. 907–920, 2013.
- [54] W. Wang, H. Gao, C. H. Liu, and K. K. Leung, "Credible and energy-aware participant selection with limited task budget for mobile crowd sensing," *Ad Hoc Networks*, vol. 43, pp. 56–70, 2016.
- [55] N. Maisonneuve, M. Stevens, M. E. Niessen, and L. Steels, "NoiseTube: measuring and mapping noise pollution with mobile phones," in *Information Technologies in Environmental Engineering*, Environmental Science and Engineering, pp. 215–228, Springer, Berlin, Germany, 2009.

