

Article

Participant Service Ability Aware Data Collecting Mechanism for Mobile Crowd Sensing

Jing Yang^{1,2,3} and Jialiang Xu^{1,2,3,*} 

¹ School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; yangjingb@cqupt.edu.cn

² Key Laboratory of Optical Communication and Networks, Chongqing 400065, China

³ Key Laboratory of Ubiquitous Sensing and Networking, Chongqing 400065, China

* Correspondence: s160131178@stu.cqupt.edu.cn; Tel.: +86-130-0832-6168

Received: 12 October 2018; Accepted: 23 November 2018; Published: 1 December 2018



Abstract: To collect data efficiently and reliably in Mobile Crowd Sensing (MCS), a Participant Service Ability Aware (PSAA) data collecting mechanism is proposed. First, participants select the best sensing task according to the task complexity and desired reward in the multitasking scenario. Second, the Stackelberg Game model is established based on the mutual choice of participants and platform to maximize their utilities to evaluate the service ability of participants. Finally, participants transmit data to platform directly or indirectly through the best relay and the sensing data from the participants with better service ability is selected to complete sensing tasks accurately and efficiently with the minimum overall reward expense. The numerical results show that the proposed data collection mechanism can maximize the utility of participants and platform, efficiently accomplish sensing tasks and significantly reduce the overall reward expense.

Keywords: Mobile Crowd Sensing; data collecting; Stackelberg game; service ability

1. Introduction

In recent years, the rapid proliferation of mobile devices such as smartphones and tablets with powerful sensing ability enables MCS to become a mobile application hotspot for the collaborative work between participants and platform [1,2]. In MCS, participants exploit the sensors embedded in their carried mobile devices (e.g., light sensor and magnetic force sensor) to collect and share data including the ambient brightness and magnetic field [3–5]. Then, the platform gathers sensing data from participants and provides real time and accurate services based on the sensing results [6]. Because MCS collects large-scale sensing data efficiently and flexibly, it can be applied to various projects. OrganiCity [7] is an example project with a very specific purpose to engage people in the development of future smart cities and provides an Experimentation as a Service (EaaS) platform to provide data streams from diverse sources inside a smart city to consumers, while recruiting participants to collect data for updating diverse sources. SmartSantander [8] is based on the largest Future Internet (FI) infrastructure and creates an experimental test facility for the research of Internet of Things (IoT) architecture, services and applications. Participants can integrate smartphones into the SmartSantander infrastructure to not only extend the infrastructure capabilities but also expand the sensing coverage and ubiquity. Besides, SmartSantander can use these smartphones for event detection. For example, SmartSantander can study mobility patterns in city streets and simultaneously monitor environmental parameters. YouSense [9] is a model-based platform created to manage participatory sensing activities and actively notify participants with questions when they are in the optimal location to collect sensing data. While avoiding bothering participants with unnecessary requests, YouSense

motivates more participants to actively participate in sensing activities to increase the quantity and quality of sensing data.

In MCS, the platform is responsible for publishing sensing tasks and providing certain rewards to motivate participants to actively perform sensing tasks. Participants are responsible for exploiting sensors embedded in mobile devices to collect and upload data to the platform [10]. Compared with the platform, participants are service providers, so the ability of participants to complete the sensing task accurately and efficiently can be defined as service ability, and the high service ability indicates that they can complete sensing tasks more efficiently. Therefore, the service ability is an important factor for participants to complete sensing tasks accurately and efficiently. However, the service ability evaluation and participant selection become challenging problems.

The existing MCS data collection research has two application modes according to service objects: participant-centric and platform-centric modes [11,12]. In the participant-centric mode, the relevant sensing data are collected by mobile devices to record participant activities. In contrast, the platform-centric mode focuses on collecting environmental data and publishing sensing results online to meet the public demands after analyzing and processing the sensing data. Although the above modes can collect some sensing data, they only consider the number of participants in sensing tasks or the total volume of sensing data collected by the platform. The larger number of participants and greater total volume of sensing data signify the higher probability to complete sensing tasks, but these two modes ignore the participant service ability. However, in the multitasking scenario, there are differences among participants in several factors, such as the geographical location, interest and residual energy, resulting in the differences in task duration and overall reward expense [13–15]. Therefore, how to exploit participants with better service ability in sensing tasks, how to select sensing tasks according to the service ability of participants and how to complete sensing tasks accurately and in real time with the minimum overall reward expense become key technical challenges.

In view of the above problems, a PSAA data collection mechanism is proposed. First, participants measure the task complexity according to their willingness, energy consumption rate and task duration, and then select the best sensing task based on the desired reward. Meanwhile, to maximize the utility of platform and participants, the platform evaluates the service ability of participants according to the timeliness and desired reward of participants and complexity of sensing tasks to selectively collect sensing data. To complete sensing tasks efficiently, participants determine the transmission mode of sensing data according to their residual energy and task duration, and then transmit data to the platform directly or indirectly through the best relay, so that sensing tasks are accomplished accurately and in real time with the minimum overall reward expense.

The main contributions of this paper are as follows:

- (1) The complexity of sensing task is evaluated by the participant willingness, energy consumption rate and task duration to effectively improve the energy utilization of mobile devices performing sensing tasks. Consequently, participants select the best task according to the task complexity and desired reward.
- (2) Aiming at the mutual choice between participants and platform, the Stackelberg game model is established to maximize the utility of participants and platform. Furthermore, the participant service ability is evaluated according to the participant's utility, timeliness, desired reward and task complexity, and employed by the platform to selectively receive sensing data.
- (3) To collect sensing data in time and provide reliable service, participants determine whether the sensing data are transmitted directly or indirectly to the platform based on the residual energy, energy consumption rate and task duration. If participants forward sensing data indirectly to the platform, the best relay is selected based on the participant service ability, sensing task similarity and intimacy degree between participants.

The rest of the paper is organized as follows. The related work are introduced in Section 2. Section 3 describes the system model. The participant service ability is evaluated in Section 4.

Section 5 designs a PSAA data collection mechanism. The numerical results are given in detail in Section 6. Finally, Section 7 concludes this paper.

2. Related Work

We now introduce related work on MCS data collecting mechanisms.

Many research efforts have been dedicated to MCS data collecting mechanisms. An incentive-aware time-sensitive data collection mechanism is designed in [16]. The proposed mechanism selects relay user by applying the Nash bargain solution, and participants forward sensing data through relay users to ensure the timely data collection. Although the above mechanism can motivate participants to collect sensing data actively, it ignores the overall reward expense of the platform. In addition, Shah-Mansouri et al. [17] proposed a Profit Maximizing Truthful (ProMoT) auction mechanism. In ProMoT, the platform releases sensing tasks to participants, and then participants submit task bids to the platform. Then, the platform selects a subset of participants based on their bids and provides proper payments for them. Although the ProMoT auction mechanism can maximize the platform utility, reduce the overall reward expense and motivate participants to truthfully participate in the auction, it ignores the data quality of participants and cannot guarantee the accuracy and reliability of sensing data. Sun et al. [18] designed an incentive scheme based on heterogeneous trust values for joint social states and real-time throughput. Participants are selected according to state attributes. Nevertheless, the proposed mechanisms do not take the reliability of sensing data into account in [16–18]. Because participants are affected by their attributes, social behaviors and other factors in MCS, when participants collect low quality sensing data, the sensing tasks cannot be completed accurately and reliably.

To reliably collect sensing data in MCS, Gao et al. [19] designed a data quality prediction mechanism with Poisson distribution based on the assumption that participants transmitted sensing data one by one. The proposed mechanism ensures the high quality of data collection and completed sensing tasks accurately. Dai et al. [20] proposed the Integrated Incentive Mechanism (IIM) to motivate participants to provide high quality sensing data, where the platform updates the reputation of participants based on their corresponding behaviors. Wen et al. [21] designed a quality-driven auction-based incentive mechanism to evaluate the reliability of sensing data according to a probabilistic model and verified the rationality of the proposed mechanism in the indoor positioning application scenario. Krontiris et al. [22] proposed a multi-attribute auction mechanism based on the traditional auction mechanism, which comprehensively considered factors such as the rewards of participants, durations of sensing tasks and total amount of collected sensing data to effectively improve the reliability of sensing data. Although the above mechanisms can ensure the reliability of sensing data, they ignore the timeliness of sensing tasks and cannot meet the flexibility and real time requirements of MCS.

In addition, only three factors are considered by the above-mentioned methods, namely the timeliness of sensing tasks, reliability of sensing data and reward expense of platform, but the complexity of sensing tasks should also be considered. In MCS, if the complexity of sensing tasks exceeds the sensing ability of participants, they cannot complete the sensing tasks efficiently and accurately. Given this kind of situation, a data collecting mechanism is designed in this paper to balance the complexity, timeliness of sensing tasks and platform utility to ensure the timely accomplished sensing tasks with the minimum overall reward expense.

3. System Model

The data collection process is shown in Figure 1. When users need to collect some GPS data, they send the task request to platform, and then the platform analyzes and sends sensing tasks to participants. In the multitasking scenario, when participants receive the information of sensing tasks in the task area, they select the best task by evaluating the complexity of sensing tasks and then collect sensing data correspondingly. At the same time, through the comprehensive analysis of the utility of

participants and platform, the timeliness and other factors are combined to evaluate the participant service ability, and then the platform chooses participants with high sensing data quality and low rewards. In the participant selecting process, the best sensing task and participants with better service ability are chosen, where the utility of participants and platform can be maximized simultaneously by the equilibrium solution. After sensing data are transmitted directly or indirectly to the platform, the platform decides whether to receive the sensing data according to the service ability, and therefore can complete sensing tasks reliably and in time with the minimum overall reward expense.

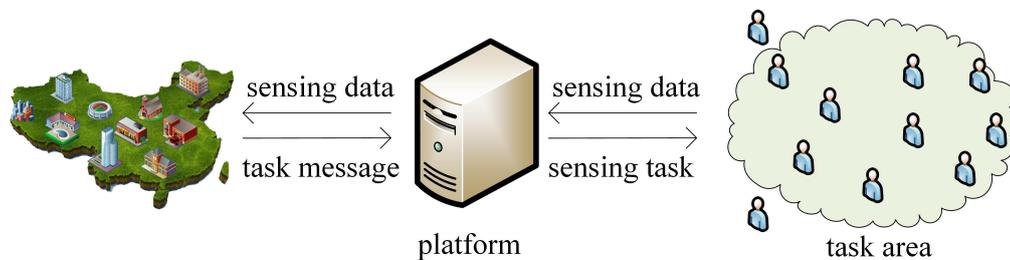


Figure 1. Data collection process.

In the above process, the GPS data must differ among distinct locations. To tackle the spatial inconsistencies, the task message is divided into many sensing tasks. We assume that the set of sensing tasks is denoted by $F = \{f_1, f_2, \dots, f_n\}$, each sensing task has its specified area and time [23]. Considering that some participants may require some extra effort for performing sensing tasks, i.e., participants need to position themselves at the specified area. However, this paper focuses on how participants choose sensing tasks and how the platform selectively receives sensing data based on the service ability of participants when they are in the task area. Therefore, $P = \{p_1, p_2, \dots, p_m\}$ is denoted by the set of participants within the task area, the participant p_i can measure the complexity of sensing tasks and select the best sensing task. At the same time, when p_i performs the sensing task, the ongoing work or battery life of p_i may be affected. To avoid the above situation, p_i is free to choose the time to perform the sensing task, but he must complete the sensing tasks within the task deadline. In addition, we note that, if the sensing data of p_i are constantly rejected by the platform, his willingness to perform the sensing task will be reduced. To actively motivate p_i participations in MCS, if the service ability of p_i is the same as others, the platform will receive his sensing data preferentially.

4. Evaluation of Service Ability

First, participants measure the complexity of a sensing task based on the willingness of participants, energy consumption rate of mobile devices and duration of sensing tasks, and then the best task can be selected according to the desired reward. Second, the Stackelberg game model is established to analyze the utility of participants and platform to maximize the reward of participants and the overall reward expense utilization of the platform in the equilibrium solution. Finally, by evaluating the service ability of participants, the platform selects participants with better service ability and participants choose the best task to complete accurately and reliably.

4.1. The Best Task Decision

Participants only execute one sensing task under the multitasking scenario. Because participants choose sensing tasks according to their willingness, if the willingness of participants is low, the probability of refusing a sensing task is high, resulting in the high complexity of the sensing task. In addition, since the collection of sensing data mainly depends on the sensors embedded in mobile devices, different types of sensing tasks should be processed. When collecting sensing data, mobile devices carried by participants suffer a certain energy loss. If a participant consumes more energy for a given sensing task, its complexity is higher. When participants perform a sensing task with long duration, they not only consume more energy in the processes of collecting, storing and

transmitting sensing data, but also affect the real time performance of tasks, and therefore the task complexity is high. Based on the above analysis, the complexity of sensing tasks should be defined by the participant willingness, energy consumption rate and task duration.

We assume the sensing task ultimately chosen by participant p_i is f_j . If the time when the platform allows participants to start f_j is $t_{f_j}^s$ and the time when sensing data collection ends is $t_{f_j}^e$, $t_{f_j}^e - t_{f_j}^s$ is the task duration. Besides, we suppose the time when p_i receives the information of f_j is t_{ij}^a and the time when p_i begins to collect sensing data is t_{ij}^m . Obviously, the value range of t_{ij}^m is $[t_{ij}^a, t_{f_j}^e]$, participating decision time B_{ij} of p_i to f_j is $t_{ij}^m - t_{ij}^a$. If participating decision time B_{ij} is short, p_i will participate in sensing task f_j actively, and the willingness of p_i is high. In addition, We define that the total energy of participant p_i is E_i^a , his residual energy is E_i^r . Since p_i may affect the use of mobile devices, the high the value of E_i^r signifies the active participation of p_i in f_j , namely the high willingness of p_i . In addition, if the decision time B_{ij} is equal to 0, it can be judged that the willingness of p_i is equal to 1, which means that the sensing task is performed immediately when p_i receives the information of f_j . Based on the above analysis, the willingness W_{ij} of p_i can be measured by decision time B_{ij} and residual energy E_i^r , as calculated by Equation (1).

$$W_{ij} = \begin{cases} \frac{E_i^r}{E_i^a} \cdot \frac{\sum_{i=1}^v B_{ij}/v}{B_{ij}}, & B_{ij} \neq 0 \\ 1, & B_{ij} = 0 \end{cases} \quad (1)$$

where v is the total number of participants selecting task f_j and $\sum_{i=1}^v B_{ij}/v$ is the average decision time for f_j .

The willingness of p_i can be effectively measured by Equation (1). Furthermore, the energy consumption per unit time of p_i is assumed to be γ_{ij} in sensing task f_j , so that complexity D_{ij} of sensing task f_j measured by p_i is obtained by Equation (2).

$$D_{ij} = 1 - W_{ij} + \frac{\gamma_{ij} \cdot (t_{f_j}^e - t_{ij}^m)}{E_i^r} \quad (2)$$

To maximize the actual rewards with the minimum energy consumption, participants choose the best task. Assuming the desired reward of p_i is c_{ij} according to D_{ij} , p_i selects the best task based on Equation (3):

$$\begin{cases} f_j = \arg \max_{\forall f_j \in F} (x_{ij} \cdot c_{ij}/D_{ij}) \\ x_{ij} \cdot (1 - x_{ij}) = 0 \\ \sum_{\forall f_j \in F} x_{ij} = 1 \end{cases} \quad (3)$$

where c_{ij}/D_{ij} is the degree that p_i participates in f_j . When desired reward c_{ij} of p_i is high, the complexity D_{ij} of f_j is low. To maximize the reward with the minimum energy consumption, a large c_{ij}/D_{ij} signifies a high probability of p_i participating in f_j . Moreover, x_{ij} is a binary Boolean variable and x_{ij} is set to 1 when p_i selects task f_j . Otherwise, x_{ij} is set to 0. $\sum_{\forall f_j \in F} x_{ij} = 1$ ensures the total number of sensing tasks performed by p_i at the same time is at most 1.

4.2. Utility Analysis

When participants perform the best task, there is an energy loss in their mobile devices during the sensing data collecting process. Therefore, when participants perform sensing tasks successfully, the platform needs to pay some rewards to compensate for the energy loss to motivate participants to perform sensing tasks actively. However, the overall reward expense of each sensing task is limited,

so participants with the high sensing data quality and low actual rewards must be selected to minimize the overall reward expense of the platform.

4.2.1. Data Quality

Sensing data quality is determined based on two factors: completeness and accuracy. Because the more complete are the sensing data, the clearer is the information transmitted to the platform by participants, data completeness is conducive to the accurate completion of sensing tasks. In addition, the sensing data accuracy can reflect the ground truth, which is also beneficial to the accurate completion of the sensing task.

We suppose that the total volume of sensing data acquired by p_i in the sensing task f_j is l_{ij} . Theoretically, the total volume of data obtained by p_i is l_{ij}^{total} . Obviously, the completeness of sensing data is high when l_{ij} approaches l_{ij}^{total} and $\max(l_{ij}/l_{ij}^{total})$ signifies the maximum approximation among all participants selecting task f_j , so the completeness h_{ij1} of sensing data can be calculated by:

$$h_{ij1} = \frac{l_{ij}/l_{ij}^{total}}{\max(l_{ij}/l_{ij}^{total})} \quad (4)$$

For the accuracy of sensing data, we assume that, when the platform ends collecting sensing data, the sensing dataset collected by the participants selecting task f_j is $X_j = \{x_{1j}, x_{2j}, \dots, x_{vj}\}$ and the platform will find a sensing data y_j with the highest similarity from X_j . When x_{ij} is close to y_j , the similarity between x_{ij} and y_j is high, indicating that the values of x_{ij} and y_j are close and the accuracy of x_{ij} is high. Therefore, the accuracy h_{ij2} of x_{ij} is measured according to the similarity between x_{ij} and y_j , as shown in Equation (5).

$$h_{ij2} = \frac{\frac{\theta}{(y_j - x_{ij})^2}}{\sum_{z=1}^m \frac{\theta}{(y_j - x_{zj})^2}} \quad (5)$$

where the value of $(y_j - x_{ij})^2$ signifies the similarity between x_{ij} and y_j , a small value indicates a high similarity, θ is the sum of similarity deviations, and $\theta = \sum_{i=1}^v (y_j - x_{ij})^2$.

In the calculation of y_j , the accuracy of each sensing data is initialized to $1/v$ to measure the weight of y_j . Thus, the most similar sensing data y_j can be obtained by Equation (6).

$$y_j = \arg \min \sum_{i=1}^v (h_{ij2} \cdot (y_j - x_{ij})^2) \quad (6)$$

where h_{ij2} is obtained by the iterative calculation. First, h_{ij2} is initialized to $1/v$, and y_j is calculated according to Equation (6). Then, h_{ij2} is updated by Equation (5). Finally, when the iteration converges, the update stops and the latest updated value is the accuracy of sensing data.

Through the above analysis, data quality q_{ij} of p_i can be quantified from completeness h_{ij1} and accuracy h_{ij2} . Since h_{ij1} and h_{ij2} have different effects on q_{ij} , data quality of p_i can be quantified by $q_{ij} = \alpha_{j1}h_{ij1} + \alpha_{j2}h_{ij2}$, where $\alpha_{j1} + \alpha_{j2} = 1$.

To avoid subjective factors leading to inaccurate results of α_{j1} and α_{j2} , we use the entropy weight method to determine their values [24]. The data dimension of h_{ij1} and h_{ij2} may introduce errors, so h_{ij1} and h_{ij2} are standardized, as shown in Equation (7).

$$\hat{h}_{ijL} = \frac{h_{ijL} - \mu_{jL}}{o_{jL}}, L = 1, 2 \quad (7)$$

where μ_{jL} and o_{jL} represent the mean and standard deviation of h_{ijL} , respectively.

We can obtain the information entropy H_{jL} of h_{ijL} according to Equations (7) and (8).

$$H_{jL} = -\frac{1}{\ln v} \cdot \sum_{i=1}^v \vartheta_{ijL} \ln \vartheta_{ijL} \quad (8)$$

where ϑ_{ijL} is the proportion of \hat{h}_{ijL} and $\vartheta_{ijL} = \hat{h}_{ijL} / \sum_{i=1}^v \hat{h}_{ijL}$. Finally, the normalized weight α_{jL} of h_{ijL} can be calculated by:

$$\alpha_{jL} = \alpha'_{jL} / \sum_{L=1}^2 \alpha'_{jL} \quad (9)$$

where α'_{jL} is the entropy weight of H_{jL} and $\alpha'_{jL} = (1 - H_{jL}) / \sum_{L=1}^2 (1 - H_{jL})$.

4.2.2. Utility Analysis

The platform can analyze the utility of participants based on the data quality and actual rewards, and then select participants with high utilities to maximize the platform utility. We assume that the overall reward expense of platform for sensing task f_j is R_j , and the energy cost per data quality is o_j in sensing task f_j . If the data quality of p_i is q_{ij} , the utility $u_j^{p_i}(q_{ij}, m_p)$ of p_i to f_j can be calculated by:

$$u_j^{p_i}(q_{ij}, m_p) = m_p \cdot (q_{ij} / \sum_{p_i \in P^j} q_{ij}) R_j - o_j \cdot q_{ij} \quad (10)$$

where m_p is the parameter determining the sensing data value, P^j is the set of participants performing sensing task f_j , and $m_p \cdot (q_{ij} / \sum_{p_i \in P^j} q_{ij}) R_j$ indicates the actual reward of p_i .

In situations of few participants in a sensing task, the sensing data quality of participants is crucial for the platform to accomplish the sensing task accurately and reliably. Because of the limited participant range, the platform utility grows with the increasing sensing data quality. With the increasing number of participants, the platform aims to achieve the minimum overall reward expense. By not only considering the sensing data quality of participants but also analyzing the utility of participants, the platform prefers to select participants with high utility. Therefore, when the number of participants is large, the sensing data quality is no longer the only factor affecting the accurate and reliable completion of sensing tasks, and the platform utility grows slowly with the increasing sensing data quality. Based on characteristics of platform utility and its change with sensing data quality, platform utility $u_j^s(q_{ij}, m_p)$ for sensing task f_j can be calculated by Equation (11).

$$u_j^s(q_{ij}, m_p) = m_s \cdot \log_2(1 + \sum_{p_i \in P^j} q_{ij}) - m_p \cdot R_j \quad (11)$$

where m_s is the parameter determining the value of the received sensing data and $\log_2(1 + \sum_{p_i \in P^j} q_{ij})$ reflects the overall trend of platform utility with the increase of data quality.

Since participants expect to maximize rewards with the minimal energy consumption, it is necessary to maximize the utility of participants by solving the optimization problem in Equation (12).

$$\begin{aligned} \max_{q_{ij}} \quad & m_p \cdot \frac{q_{ij}}{\sum_{p_i \in P^j} q_{ij}} R_j - o_j \cdot q_{ij} \\ \text{s.t.} \quad & q_{ij} \geq 0 \end{aligned} \quad (12)$$

Correspondingly, the platform selects participants with high sensing data quality and low actual rewards to minimize the overall reward expense. Therefore, it is necessary to maximize the platform utility by addressing the optimization problem in Equation (13).

$$\begin{aligned} \max_{m_p} \quad & m_s \cdot \log_2(1 + \sum_{p_i \in P^j} q_{ij}) - m_p \cdot R_j \\ \text{s.t.} \quad & m_p \geq 0 \end{aligned} \quad (13)$$

Obviously, Equations (12) and (13) constitute a Stackelberg game [25,26]. Since Stackelberg equilibrium belongs to subgame Nash equilibrium, the unilateral decision changes of participants or platform cannot further improve their utilities. Consequently, the equilibrium must be obtained to find a stable utility equilibrium between participants and platform, so that the utilities of both are maximized. According to the definition of equilibrium solution [27], for any $q_{ij} \geq 0$ and $m_p \geq 0$, if (q_{ij}^*, m_p^*) satisfies the conditions in Equation (14), (q_{ij}^*, m_p^*) is the equilibrium solution of the Stackelberg game.

$$\begin{cases} u_j^{p_i}(q_{ij}^*, m_p^*) \geq u_j^{p_i}(q_{ij}, m_p^*) \\ u_j^s(q_{ij}^*, m_p^*) \geq u_j^s(q_{ij}, m_p) \end{cases} \quad (14)$$

According to Equations (10) and (13) and the number of participants in the sensing task, Equation (15) can be obtained.

$$q_{ij}^* = m_p \cdot \frac{|P^j| - 1}{|P^j|^2 \cdot o_j} R_j \quad (15)$$

where q_{ij}^* is an expression of parameter m_p and $|P^j|$ represents the total number of participants in the set of P^j . According to Equations (11) and (15), the partial derivatives of platform utility u_j^s can be exploited to find the maximum value, as shown in Equation (16).

$$\begin{cases} \frac{\partial u_j^s}{\partial m_p} = \frac{m_s}{\ln 2} \cdot \frac{(|P^j|-1)R_j}{|P^j| \cdot o_j + m_p \cdot (|P^j|-1)R_j} - R_j \\ \frac{\partial^2 u_j^s}{\partial m_p^2} = -\frac{m_s}{\ln 2} \cdot \frac{(|P^j|-1)^2 R_j^2}{[|P^j| \cdot o_j + m_p \cdot (|P^j|-1)R_j]^2} \end{cases} \quad (16)$$

When m_p is equal to zero, obviously $u_j^s(q_{ij}, m_p)$ is also zero, so platform utility $u_j^s(q_{ij}, m_p)$ is equal to zero according to Equation (11). Because $u_j^s(q_{ij}, m_p)$ is increasing slowly, when $u_j^s(q_{ij}, m_p)|_{m_p=0} = 0$, $\frac{\partial u_j^s}{\partial m_p}|_{m_p=0} > 0$ can be obtained and we can determine the value range of m_s , as shown in Equation (17).

$$m_s > \ln 2 \cdot \frac{|P^j| \cdot o_j}{|P^j| - 1} \quad (17)$$

Since Equation (16) shows that $\frac{\partial^2 u_j^s}{\partial m_p^2} \leq 0$ always holds, parameter m_p satisfying $\frac{\partial u_j^s}{\partial m_p} = 0$ can maximize $u_j^s(q_{ij}, m_p)$, where m_p^* can be calculated by:

$$m_p^* = \frac{m_s}{\ln 2 \cdot R_j} - \frac{|P^j| \cdot o_j}{(|P^j| - 1)R_j} \quad (18)$$

Thus, the equilibrium solution can maximize the utility of participants and platform, as shown in Equation (19).

$$\begin{cases} u_j^{p_i}(q_{ij}^*, m_p^*) = \frac{m_p^* \cdot R_j}{|P^j|^2} \\ u_j^s(q_{ij}^*, m_p^*) = m_s \cdot \log_2\left(\frac{|P^j|-1}{|P^j| \cdot o_j} \cdot \frac{m_s}{\ln 2}\right) - \frac{m_s}{\ln 2} + \frac{|P^j| \cdot o_j}{|P^j|-1} \end{cases} \quad (19)$$

According to Equation (19), as the total number of participants $|P^j|$ increases, the utility of participants decreases sharply, because, when $|P^j|$ is large enough, the platform will selectively receive sensing data to provide accurate and reliable services. Therefore, to incite the platform to receive their sensing data, participants will inevitably improve sensing data quality or reduce actual rewards. However, when participants change their strategies unilaterally, they cannot improve their own utilities in the Stackelberg game, so that the utility of participants decreases as $|P^j|$ increases. Moreover, as shown in Equation (19), parameter m_s in the platform utility $u_j^s(q_{ij}^*, m_p^*)$ are not all positive, because m_s

determines the value of the received sensing data. As the quality of sensing data is greatly improved, the platform utility is not sharply increased, so m_s are not all positive in $u_j^s(q_{ij}^*, m_p^*)$. Based on the above analysis, the game result determined by the equilibrium solution is accurate, which can maximize the utility of participants and platform.

4.3. Evaluation of Service Ability

The utilities of participants and platform are maximized to select the best task in the multitasking scenario. However, due to the significant features of a large number of participants in MCS, the only goal of participants is to maximize their actual rewards in the sensing data collecting process and they ignore the final rewards of other participants performing the same task. At the same time, the only goal of the platform is to provide accurate and reliable services with the minimum overall reward expense. Therefore, the maximum utility of a single participant or the platform cannot unilaterally achieve the optimal overall utility. In addition, because participants have random mobility, they may be interfered by factors such as the geographical environment in the sensing data collecting process, causing the sensing data to be inconsistent with task requirements or to be tampered with by malicious participants. The above situation seriously affects the sensing results of the platform, and therefore not only participants should choose the best task, but also the platform should screen all participants performing the same sensing task to provide accurate and reliable services.

Participants have the great flexibility in choosing a sensing task at any time, the willingness of participants varies, and the energy consumption of sensing tasks is also different, indicating participants independently measure the complexity of sensing tasks. Moreover, to maximize the actual rewards of participants, minimize the overall reward expense of platform, and maximize the utilities of participants and platform as a whole, the platform needs to analyze the service ability of participants to select the best participants. The service ability is crucial for the platform to measure whether to receive the sensing data of participants; a stronger service ability signifies a higher probability of sensing data being received by the platform.

The platform limits the duration of sensing tasks due to the real time requirements of services. Obviously, the less time a participant takes to make a decision about a sensing task signifies the better timeliness and the higher probability of the participant choosing the task, and the platform tends to select this participant in order to complete the task accurately and reliably. However, participant selecting cannot be only based on the timeliness. Because the platform utility is affected by the participant utility due to the important factor of data quality, the high participant utility signifies the accurate sensing data and high benefits for the platform. Therefore, it is necessary to analyze the impact of participant utility on service ability. In terms of the task complexity and the desired rewards of participants, if the task complexity and desired rewards evaluated by participants are low, the overall reward utilization rate of the platform is high and the participant service ability is also strong. According to the above analysis, the participant service ability can be measured by the timeliness, participant utility, task complexity and desired rewards.

We know participation decision time B_{ij} of p_i to f_j is $t_{ij}^m - t_{ij}^a$ in Section 4.1, and its maximum value is $t_{f_j}^e - t_{ij}^a$. If participation decision time B_{ij} is short, p_i chooses the best sensing task f_j timely. Therefore, timeliness of participants M_{ij} can be analyzed by B_{ij} . Obviously, when B_{ij} is equal to 0, timeliness M_{ij} is the maximum and M_{ij} must decrease with the growing B_{ij} . Moreover, when the value of B_{ij} is small, the task duration is relatively long and the value of M_{ij} changes slightly around the maximum value. Otherwise, when the value of B_{ij} is large, the sensing time allowed by the platform is limited and the value of M_{ij} will drop sharply. Although the above-mentioned nonlinear trend is consistent with $1 - e^{-e^{-B_{ij}}}$, for the convenience of description and comparison, it is defined that the value of M_{ij} is set to 1 when B_{ij} is set to 0, and the value of M_{ij} is set to 0 when B_{ij} is equal to $t_{f_j}^e - t_{ij}^a$.

Due to the non-linear variation trend, M_{ij} decreases slowly first and then sharply decreases and the quantification of timeliness M_{ij} is shown in Equation (20).

$$M_{ij} = \begin{cases} 0 & , t_{ij}^a = t_{f_j}^e \\ \frac{T_{ij}(B_{ij}+t_{ij}^a)-T_{ij}(t_{f_j}^e)}{T_{ij}(t_{ij}^a)-T_{ij}(t_{f_j}^e)} & , t_{ij}^a \neq t_{f_j}^e \end{cases} \quad (20)$$

where $T_{ij}(t) = 1 - e^{-e^{-t}}$ describes the overall trend of timeliness with participation decision time.

The platform determines whether to receive sensing data according to the participant service ability. When continuing to participate in sensing tasks, the service ability of participants is updated constantly, and the current service ability depends on not only the next sensing task but also the utility of the participants who completed the last sensing task. Although participants select the best sensing task at the initial stage, the historical task list is empty, which means that there is no participant who completed the last sensing task. Therefore, the initial service ability is evaluated according to the timeliness, task complexity and desired rewards. The range of service ability is defined in $[0,1]$, and the value of the strongest service ability is 1. Assuming the initial best sensing task of participant p_i is f_j , we can obtain initial service ability s_i^j of p_i by Equation (21).

$$s_i^j = M_{ij} \cdot \frac{\min(c_{kj})}{c_{ij}} \cdot \frac{\max(D_{kj}) - D_{ij}}{\max(D_{kj}) - \min(D_{kj})}, \forall p_k \in P^j \quad (21)$$

When the historical task list is not empty, if participant p_i completes the last sensing task significantly better than other participants, the platform will choose to receive the sensing data from p_i , which means p_i has the strongest service ability to perform the next sensing task. With the increase of participant utility, a new equilibrium is reached between participants and platform, the more advantageous to platform completing the sensing task reliably, so the service ability of participants grows faster and faster. Therefore, the utility of p_i completing the last sensing task f_b is assumed to be $u_b^{p_i}$. According to the initial service ability quantification method and the value range of service ability, if p_i selects f_h as the next sensing task, next service ability s_i^h can be obtained by Equation (22).

$$s_i^h = \min(1, s_i^j \cdot e^{(u_b^{p_i}-1)}) \quad (22)$$

In Equation (22), $u_b^{p_i}$ is equal to $\frac{u_b^{p_i}}{\sum u_b^{p_i} / |P^b|}$, where $\sum u_b^{p_i} / |P^b|$ is the average utility of participants, so $u_b^{p_i}$ representing the extent to which utility $u_b^{p_i}$ of p_i is significantly higher than those of other participants for last sensing task f_b .

5. Data Collection Mechanism

Due to the limited residual energy of mobile devices and task duration, the sensing data may not be directly transmitted to the platform, so the transmission modes of sensing data can be switched according to the relationship between the theoretical transmission delay and the maximum transmission delay. Besides, participants choose the best relay according to the next service ability, similarity of sensing tasks and intimacy degree of participants. The participant service ability aware data collection mechanism is designed for the multitasking scenario.

5.1. Transmission Mode Determination

In the task duration, for the platform to collect sensing data in time and provide reliable services, the sensing data are transmitted directly or indirectly to the platform based on factors such as the energy consumption rate of mobile devices [28]. Therefore, it is necessary to discriminate the transmission modes of sensing data.

Constrained by the residual energy of mobile devices and sensing task duration, if participants choose to transmit sensing data directly, the energy loss of data transmission must be lower than the residual energy and the sensing data should be successfully transmitted to the platform. Therefore, under the above two conditions, the low energy consumption of transmitting the sensing data signifies the small actual transmission delay and the great probability of transmitting the sensing data directly. Otherwise, the sensing data will be transmitted indirectly through the best relay.

To discriminate the transmission modes of sensing data, we assume that all participants theoretically analyze the direct transmission. Once the above conditions are not met, participants choose the indirect data forwarding. For sensing task f_j , participant p_i is assumed to start transmitting sensing data at time t_{ij}^s . If the sensing data transmission can be completed theoretically at time t_{ij}^o , the total volume of sensing data collected successfully is S_{ij} . In addition, the average data transmission rate of p_i is assumed to be η_{ij} , so the above conditions are shown in Equation (23).

$$\begin{cases} \gamma_{ij}(t_{ij}^o - t_{ij}^s) \leq E_i^r \\ \eta_{ij}(t_{ij}^o - t_{ij}^s) \geq S_{ij} \\ 0 < t_{ij}^o - t_{ij}^s \leq t_{f_j}^e - t_{ij}^s \end{cases} \quad (23)$$

In Equation (23), $t_{f_j}^e - t_{ij}^s$ represents the maximum transmission delay allowed by the platform to collect the sensing data from p_i and $t_{ij}^o - t_{ij}^s$ is the theoretical transmission delay. When the transmission delay is too large to meet the conditions in Equation (23), indirect forwarding is employed to complete the sensing data transmission. If the theoretical transmission delay satisfies the above equation, the sensing data can be directly transmitted to the platform. However, since the actual scenario is complicated, when the actual transmission delay is larger, participants consume more energy, which affects the timeliness and efficiency of sensing data collection. According to Equation (23), the sensing data are transmitted directly when $(t_{ij}^o - t_{ij}^s) / (t_{f_j}^e - t_{ij}^s) \in (0, 0.5]$. Otherwise, the sensing data are transmitted in an indirect forwarding manner.

After discriminating the transmission modes of sensing data, $P_d = \{p_d^1, p_d^2, \dots\}$ and $P_{nd} = \{p_{nd}^1, p_{nd}^2, \dots\}$ are assumed to be the participants with direct transmission and indirect transmission respectively. Participant p_d^α transmits sensing data to the platform directly, whereas participant p_{nd}^β selects the best relay from $P_d = \{p_d^1, p_d^2, \dots\}$ to complete the sensing data transmission. If the next service ability of p_d^α is strong, the probability that the utilities of the participant and platform are maximized is high, and the platform will be highly likely to receive sensing data from p_d^α , and therefore the probability of p_{nd}^β choosing p_d^α for data transmission is high. However, in the multitasking scenario, many sensing tasks that can be selected by participants at the same time, when the best tasks selected by p_d^α and p_{nd}^β are different, because the next service ability of p_d^α is stronger, p_d^α can successfully transmit the sensing data of p_{nd}^β to the platform, but the platform cannot identify that the sensing data of p_d^α and p_{nd}^β belong to different sensing task, the accuracy of sensing results is affected. Therefore, p_{nd}^β tends to select the participant in the same sensing task for data transmission. In other words, the high similarity of sensing tasks selected by p_d^α and p_{nd}^β signifies the great probability of p_{nd}^β selecting p_d^α for indirect data forwarding. In addition, when p_d^α and p_{nd}^β choose the same sensing task, p_d^α is not necessarily willing to help p_{nd}^β forward the sensing data, because they may distrust each other and their mobile devices have the energy consumption in the process of data transmission. Therefore, p_d^α is not necessarily the best relay for p_{nd}^β . Due to the social attributes of participants, when p_{nd}^β chooses the best relay or p_d^α helps the data forwarding, they have certain preferences [29–31]. If the social relationship between p_d^α and p_{nd}^β is close, they will trust each other and p_{nd}^β will be highly likely to select p_d^α as the best relay. Eventually, the best relay can be selected according to the next service ability of participants, similarity of sensing tasks and intimacy degree between participants.

According to the evaluation method in Equation (22), the next service ability of p_d^α is assumed to be $s_{d_\alpha}^h$. In terms of sensing tasks similarity, participants only consider the task complexity and desired reward from their own perspectives when selecting the best sensing task, and they ignore the choices of other participants. In other words, participants choose the best sensing task without distinguishing the transmission mode. Therefore, they cannot analyze the similarity between sensing tasks of all other participants within the participation decision time. To avoid the above problem, participants perform sensing tasks continuously in the sensing area. When the similarity between historical task lists of participants is high, they are interested in many sensing tasks and have high similarity in the next sensing task. When analyzing the best relay, we should measure the similarity of sensing tasks according to historical task lists. The method measuring similarity $r_{\alpha\beta}$ between sensing tasks of p_d^α and p_{nd}^β is shown in Equation (24).

$$r_{\alpha\beta} = \frac{|F^\alpha \cap F^\beta|}{|F^\alpha \cup F^\beta|} \quad (24)$$

where F^α and F^β represent the historical sensing tasks performed by p_d^α and p_{nd}^β , respectively, $|F^\alpha \cap F^\beta|$ and $|F^\alpha \cup F^\beta|$ represent the total numbers of p_d^α and p_{nd}^β performing the same historical sensing tasks, and $r_{\alpha\beta}$ has value range [0,1].

In addition, the intimacy degree between participants can be analyzed through the interaction time, because the long interaction time between participants signifies the high intimacy degree between them. The intimacy degree between participants grows fast when their early interactions are established. As they continue to interact, the intimacy degree gradually increases to the maximum. When the intimacy degree is stable, its value remains almost unchanging. Therefore, the range of intimacy degree is defined in [0,1], and the intimacy degree between p_d^α and p_{nd}^β can be obtained according to the interaction time, as shown in Equation (25).

$$N_{\alpha\beta} = \frac{n_{\alpha\beta} / \sum_{p_d^\sigma \in P_d} n_{\sigma\beta}}{1 + e^{-t}} \quad (25)$$

where $n_{\alpha\beta}$ represents the times that p_d^α helps p_{nd}^β complete the sensing data transmission, and $\sum_{p_d^\sigma \in P_d} n_{\sigma\beta}$ represents the times of the participants in set P_d forwarding the sensing data for p_{nd}^β . To accurately determine the initial intimacy degree between participants, when p_d^α and p_{nd}^β establish the first interaction, their intimacy degree is set to $n_{\alpha\beta} / \sum_{p_d^\sigma \in P_d} n_{\sigma\beta}$.

Since the next service ability of p_d^α is $s_{d_\alpha}^h$, similarity of sensing tasks $r_{\alpha\beta}$ and intimacy degree $N_{\alpha\beta}$ between p_d^α and p_{nd}^β are described by Equations (24) and (25), respectively. Finally, p_{nd}^β chooses the best relay according to Equation (26).

$$p_d^\alpha = \arg \max_{\forall p_d^\alpha \in P_d} (s_{d_\alpha}^h \cdot r_{\alpha\beta} \cdot N_{\alpha\beta}) \quad (26)$$

The platform determines whether to receive the sensing data of participants according to the service ability. If the best relay p_d^α is selected by p_{nd}^β according to Equation (26), p_d^α not only transmits its own sensing data, but also helps p_{nd}^β forward sensing data. Therefore, after receiving the sensing data from p_{nd}^β , p_d^α marks it for the platform to collect the sensing data correctly.

5.2. Sensing Data Collection

In the multitasking scenario, participants select the best sensing task according to the sensing task complexity and desired rewards to maximize their utilities with the minimum energy consumption. However, the platform determines whether to receive the sensing data to maximize its utility and complete sensing tasks efficiently and accurately by evaluating the service ability of participants. The next sensing task that participant p_i selects to perform is assumed to be f_h and the platform selects participants according to Equation (27).

$$\begin{aligned}
 & \begin{cases} \max u_h^s(q_{ih}, m_p) \\ \max u_h^{p_i}(q_{ih}, m_p) \\ \max (\min s_i^h) \end{cases} \\
 \text{s.t.} & \begin{cases} x_{ih} = 1 \\ \sum_{\forall p_i \in P^h} c_{ih}^s \leq R_h, \forall p_i \in P^h \\ B_{ih} < t_{f_h}^e - t_{ih}^a \end{cases} \quad (27)
 \end{aligned}$$

In Equation (27), objective function $\max u_h^s(q_{ih}, m_p)$ and $\max u_h^{p_i}(q_{ih}, m_p)$ are employed to maximize the utilities of participants and platform, when participants select the best sensing task and platform chooses the participants with strong service ability. Besides, $\max (\min s_i^h)$ represents the requirements of the platform for participant service ability. The constraint conditions mean that, when participant p_i chooses f_h as the next sensing task, the participation decision time of p_i must meet the real time requirement of f_h and cannot exceed the maximum participation decision time of f_h . Furthermore, the total rewards of all participants who choosing f_h must not exceed the total reward cost of f_h . Participants and platform choose each other to transmit sensing data directly or indirectly, and then the platform chooses to receive the sensing data from participants with strong service ability. The algorithm of proposed PSAA data collecting mechanism is shown below (Algorithm 1).

Algorithm 1 PSAA data collecting mechanism.

```

1: BEGIN
2: for all  $p_i \in P$  do
3:   choose task from  $F$  of  $p_i$  according to Equation (2)
4: end for
5: for all  $p_\alpha^j \in P^j$  do
6:   determine sending method of  $p_\alpha^j$  according to Equation (23)
7: end for
8: for all  $p_{nd}^\gamma \in P_{nd}$  do
9:   for all  $p_d^\lambda \in P_d$  do
10:    if satisfy Equation (26) then
11:       $p_{nd}^\gamma$  sends data to  $p_d^\lambda$ 
12:    else
13:      Go to Line 9
14:    end if
15:  end for
16: end for
17: for all  $p_\alpha^j \in P^j$  do
18:   if satisfy Equation (27) then
19:     the server accepts sensing data
20:   else
21:     the server refuses to receive
22:   end if
23: end for
24: END

```

6. Numerical Results

We evaluated the proposed mechanism on a real dataset: T-drive [32,33]. T-drive contains the GPS traces of 10,357 taxis in Beijing for one week, from 2 February to 8 February 2008. Each taxi is equipped with the GPS sensor, and the average retrieving the GPS position is about 177 s. In T-drive, each trace includes the taxi ID, timestamp and the latitude and longitude position of the taxi. We took each taxi as a participant equipped with built-in sensors and assumed that the data collected by the participants are their GPS location information.

We chose the region around the Fourth Ring Road in Beijing as the sensing region, the latitude range of the sensing region is (39.84002° N–39.99397° N) and the longitude range is (116.27621° E–116.49424° E). According to the location of the sensing region, we selected 900 participants whose trajectories were distributed in the sensing region. In addition, for the convenience of analysis, a dataset containing the GPS traces of 900 participants for one day was extracted from T-drive. Meanwhile, a false dataset containing 100 participants uploading the wrong GPS location information was constructed to simulate the dishonesty behaviors of participants collecting low-quality sensing data. We selected another 100 participants in the sensing region from T-drive. Some of their GPS position were modified during a certain period, and the modified GPS position was randomly generated. Therefore, we used the real dataset and the false dataset to build the simulation environment. Considering that a participant's reward could be achieved in different formats in practice, such as money or bonus points, we used dimensionless units to represent both the participants' reward and the maximum total reward of sensing task. The main simulation parameters are given in Table 1.

Table 1. Simulation parameter settings.

Parameter	Value
simulation duration/s	34,200
number of participants	[100, 1000]
number of tasks	[1, 15]
task duration/s	[900, 3600]
total energy of mobile device/(mA·h)	2000
residual energy of mobile device/(mA·h)	[400, 1900]
maximum total reward of sensing task	[1000, 2000]
reward for a single participant	[1, 5]
average data transmission rate/kbps	400

As shown in Table 1, the residual energy of participants and the reward of a single participant were assumed to be generated randomly, and all simulations were run on MATLAB, which was deployed on a PC with 1.70 GHz CPU and 4GB RAM. In addition, the proposed PSAA was compared with the ProMoT [17] and IIM [20] mechanisms. First, we randomly selected some participants to observe the change of their service ability. Then, the utilities of participants and platform were analyzed to validate PSAA. Furthermore, in the multitasking scenario, PSAA, ProMoT and IIM mechanisms were analyzed under different numbers of participants and different overall reward expenses.

6.1. Utility Analysis of Participants and Platform

To effectively analyze the service ability of participants, we selected participants No. 36, No. 125, No. 321 and No. 763 who performed the same sensing tasks, and their service abilities were clearly differentiated. The changes in the service ability are shown in Figure 2. Obviously, the range of service ability is [0,1] according to its definition. We note a difference in the initial service ability of participants, which may be due to factors such as the residual energy or desired reward. Besides, the service ability of participant No. 36 increased gradually from 0.61 to 0.93, indicating that he chose the best sensing task and collected high-quality sensing data, so his service ability increased. The service ability of participant No. 763 reduced gradually from 0.32 to 0.01, which may be caused by dishonest behavior and unreasonable desired reward. The service ability of participant No. 125 increased from

0.53 to 0.7 and then decreased to 0.55. It is possible that the residual energy of participant No. 125 was insufficient to perform more sensing tasks, resulting in the increased complexity of the sensing task. The service ability of participant No. 321 reduced from 0.47 to 0.23 and then increased to 0.51, because it may be difficult to select the best sensing task when the number of sensing tasks is small, resulting in his low willingness. As the number of tasks increases, he can choose the best sensing task to increase the service ability.

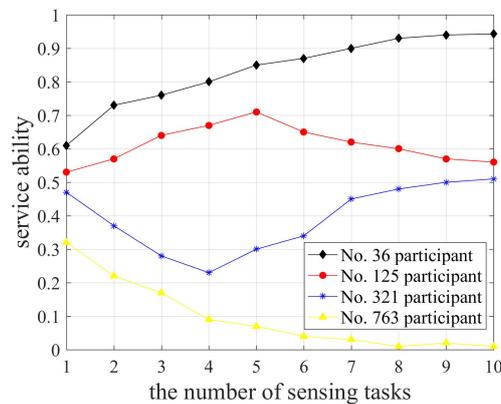


Figure 2. Changes in the service ability of participants.

The utility of participants is shown in Figure 3. It can be seen that the utility of participants in PSAA was the highest because participants chose the best task and platform selectively collects sensing data. In the bilateral selection process, the Stackelberg game was formed with an equilibrium solution to provide participants with the highest utility. At the same time, compared to ProMoT, IIM had higher utility of participants, because the platform selectively received the sensing data according to the reverse auction principle and participants collected high quality sensing data to obtain as high a reward as possible.

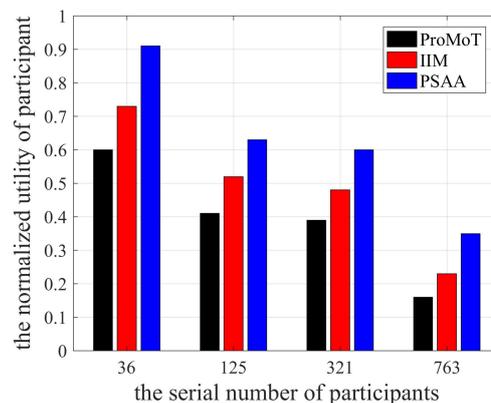


Figure 3. The utility of participants.

The platform utility is shown in Figure 4. As the number of sensing tasks increased, the platform utility increased gradually and then remained relatively stable, because the platform needed to collect, analyze, and process the sensing data. In addition, PSAA stabilized before IIM and ProMoT. Because PSAA takes into account the timeliness of participants and data transmission delay, it consumes the least time to complete the sensing task and its platform utility also stabilizes quickly. Since the platform selectively receives sensing data according to the participant service ability, the utilities of platform and participants of PSAA are both the highest due to the equilibrium solution. Compared to IIM, the platform in ProMoT only considers the total volume of sensing data, which has the lowest utility.

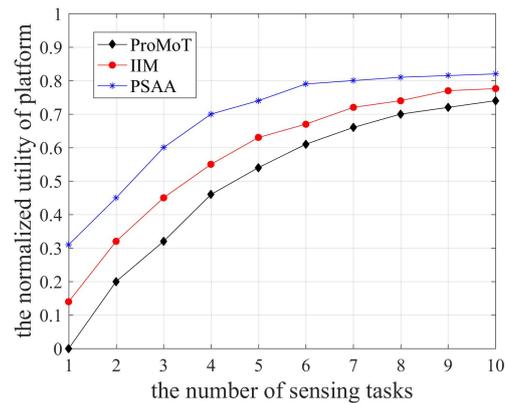


Figure 4. The utility of platform.

6.2. Performance Analysis in the Multitasking Scenario

We assumed the total number of sensing tasks is 5 in the multitasking scenario, as shown in Table 2. PSAA, ProMoT and IIM are verified based on the total volume of sensing data, overall sensing time, total number of participants in each task and overall reward expenses.

Table 2. Sensing task information.

Task	Starting Time	Ending Time	Task Duration/min
1	08:00:00	08:50:00	50
2	08:07:00	08:47:00	40
3	08:12:00	08:42:00	30
4	08:26:00	09:11:00	45
5	08:40:00	09:40:00	60

6.2.1. Performance Analysis under Different Numbers of Participants

The total volume of sensing data collected by the three data collection mechanisms is shown in Figure 5. It can be seen that the total volume of sensing data increased rapidly and then slowly and finally stabilized with the growing number of participants, because the number participants was small at the beginning and the platform actively received sensing data to complete the sensing task efficiently and accurately. When the number of participants increased gradually, there was some redundancy between sensing data and the platform only processed the valid sensing data, so the total volume of sensing data grew slowly. When the number of participants is large, the overall reward expense of the platform was limited and the sensing data quality among participants was different, which means that the platform needed to reject some participants' sensing data, and therefore the total volume of sensing data remained stable. In addition, ProMoT only considers the total volume of sensing data collected by the platform, so its total volume of sensing data is the largest. However, compared to IIM, PSAA analyzes the data transmission delay based on the total volume of sensing data, so its total volume of sensing data is the least.

The overall sensing time of the three data collection mechanisms is shown in Figure 6. Apparently, the overall sensing time decreased rapidly and then stabilized with the growing number of participants. When the number of participants started to increase, the total volume of sensing data collected by the platform also grew to facilitate the completion of sensing tasks and reduce the overall sensing time. When the number of participants was large enough, the total volume of sensing data processed by the platform did not increase, so the overall sensing time stabilized. In addition, the timeliness of participants is considered by PSAA in the service ability analysis and the transmission mode of sensing data is determined according to the transmission delay, so the overall sensing time of PSAA

is the lowest. IIM updates the participant reputation dynamically and selects participants with high reputation, which has lower overall sensing time than ProMoT.

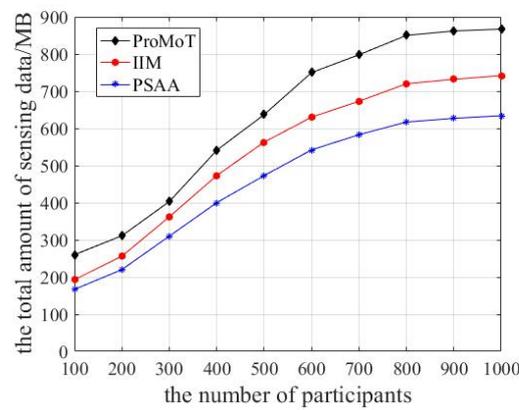


Figure 5. The total volume of sensing data.

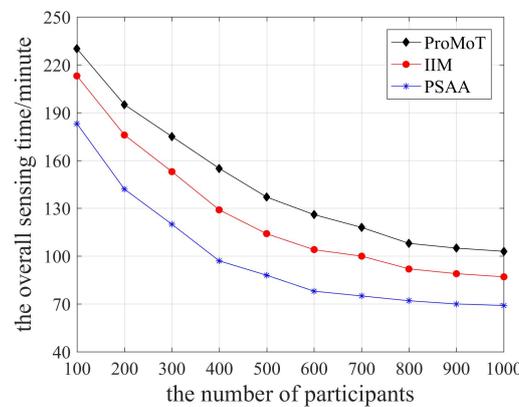


Figure 6. The overall sensing time.

The number of participants in each sensing task is shown in Figure 7. It can be seen that the number of participants in sensing task No. 5 was the largest for all three mechanisms, because the platform allowed the longest sensing time for task No. 5 due to the large and complex scene scale. Therefore, the platform required many participants in No. 5 scene. Moreover, participants select the best sensing task and the platform chooses sensing data according to the participant service ability in PSAA, so its number of participants is the smallest. The reputation of participants is analyzed by the platform in IIM, whereas ProMoT only maximizes the rewards of participants, resulting in the larger number of participants in ProMoT.

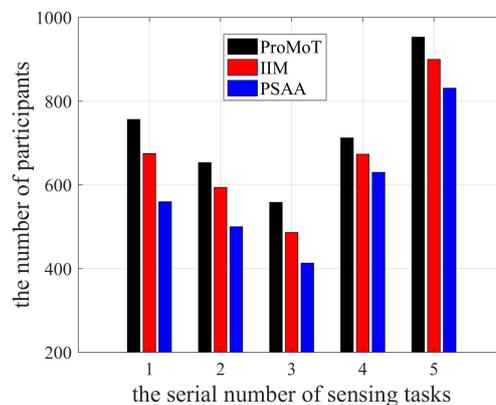


Figure 7. The number of participants.

6.2.2. Performance Analysis under Different Total Reward Expenses

The total volume of sensing data collected by the three data collection mechanisms is shown in Figure 8. It can be seen that the total volume of sensing data grew to a certain extent and then stabilizes with the increase of overall reward expense. When the overall reward expense increased gradually, it inevitably attracted participants to perform the sensing task actively, so the total volume of sensing data grew gradually. When the overall reward expense was high, the platform selectively received the sensing data under the three data collection mechanisms, so the total volume of sensing data stabilized. In addition, the platform expects to collect enough sensing data in ProMoT, so its total volume of sensing data is the largest. Compared to IIM, PSAA considers the data transmission delay, so its total volume of sensing data collected by the platform is the least.

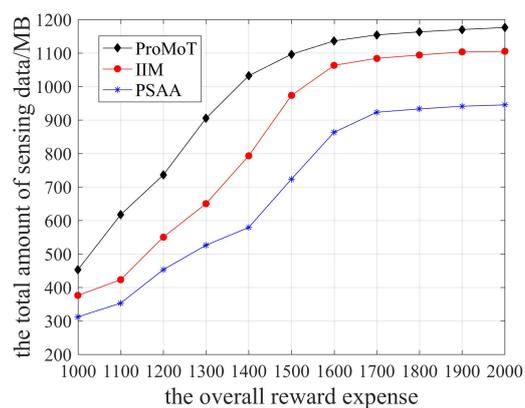


Figure 8. The total volume of sensing data.

The overall sensing time of the three data collection mechanisms is shown in Figure 9. It can be seen that the overall sensing time decreased and then stabilized with the increase of overall reward expense. When the overall reward expense was low and the number of participants was small, the platform could not collect sensing data efficiently, resulting in a high overall sensing time. However, more and more participants chose to perform sensing tasks due to the increase of overall reward expense, so the overall sensing time decreased. When the overall reward expense was relatively large, the total volume of sensing data stabilized, which means the platform collected no more sensing data, so the overall sensing time remained almost unchanged. In addition, the utility of participants and platform in PSAA can reach the maximum and its overall sensing time stabilizes first. While aiming to maximize the utility of participants, ProMoT has the largest overall reward expense.

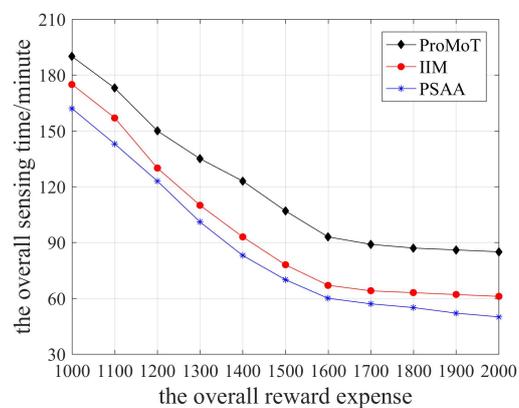


Figure 9. The overall sensing time.

The overall reward expense of each sensing task is shown in Figure 10. Obviously, the overall reward expense of sensing task No. 5 was the largest for all three mechanisms. According to Figure 6,

the number of participants in sensing task No. 5 was the largest, so it had the largest overall reward expense. Combined with the number of participants in Figure 6, Figure 9 shows the overall reward expenses of PSAA and ProMoT are the lowest and the highest, respectively.

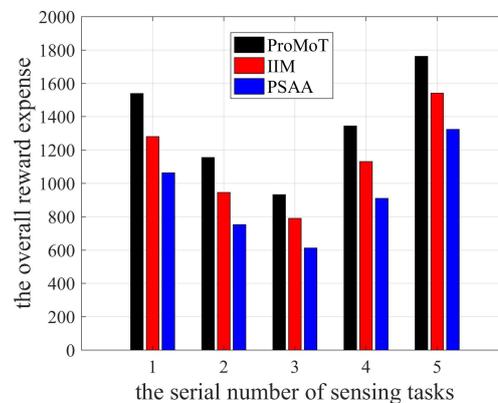


Figure 10. The overall reward expense.

7. Conclusions

To collect data efficiently and reliably in MCS, PSAA is proposed in our paper. First, participants select the best sensing task by measuring the task complexity and desired reward. Second, the Stackelberg game model is established according to the mutual choice between participants and the platform to maximize the utilities of participants and platform. Then, the next service ability of participants is employed along with the initial service ability for the platform to selectively receive sensing data. Finally, participants transmit data directly or indirectly to the platform, ensuring the efficient and accurate completion of sensing tasks the minimum overall reward expense. In our future work, the mobility of participants would be considered for MCS data collection.

Author Contributions: The work presented in this paper is a collaborative development by all of the authors. J.Y. and J.X. proposed the idea, were involved in the theoretical performance analysis, designed and optimized the algorithm, and wrote the paper.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 61771082 and Grant 61871062, and in part by the Program for Innovation Team Building at Institutions of Higher Education in Chongqing under Grant CXTDX201601020.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wu, D.; Liu, Q.; Li, Y.; McCann, J.A.; Regan, A.C.; Venkatasubramanian, N. Adaptive Lookup of Open WiFi Using Crowdsensing. *IEEE/ACM Trans. Netw.* **2016**, *24*, 3634–3647. [[CrossRef](#)]
2. Miao, Y.; Ma, J.; Liu, X.; Li, X.; Liu, Z.; Li, H. Practical Attribute-Based Multi-Keyword Search Scheme in Mobile Crowdsourcing. *IEEE Internet Things J.* **2018**, *5*, 3008–3018. [[CrossRef](#)]
3. Xiang, C.; Yang, P.; Tian, C.; Zhang, L.; Lin, H.; Xiao, F.; Zhang, M.; Liu, Y. CARM: Crowd-Sensing Accurate Outdoor RSS Maps with Error-Prone Smartphone Measurements. *IEEE Trans. Mob. Comput.* **2016**, *15*, 2669–2681. [[CrossRef](#)]
4. Luo, T.; Kanhere, S.S.; Das, S.K.; Tan, H. Incentive Mechanism Design for Heterogeneous Crowdsourcing Using All-Pay Contests. *IEEE Trans. Mob. Comput.* **2016**, *15*, 2234–2246. [[CrossRef](#)]
5. Zhang, Y.; Yu, R.; Xie, S.; Yao, W.; Xiao, W.; Guizani, M. Home M2M networks: Architectures, standards, and QoS improvement. *IEEE Commun. Mag.* **2011**, *49*, 44–52. [[CrossRef](#)]
6. Han, K.; Zhang, C.; Luo, J. Taming the Uncertainty: Budget Limited Robust Crowdsensing Through Online Learning. *IEEE/ACM Trans. Netw.* **2016**, *24*, 1462–1475. [[CrossRef](#)]

7. Amaxilatis, D.; Mylonas, G.; Diez, L.; Theodoridis, E.; Gutiérrez, V.; Muñoz, L. Managing Pervasive Sensing Campaigns via an Experimentation-as-a-Service Platform for Smart Cities. *Sensors*. **2018**, *18*, 2125. [[CrossRef](#)] [[PubMed](#)]
8. Mylonas, G.; Theodoridis, E.; Muñoz, L. Integrating Smartphones into the SmartSantander Infrastructure. *IEEE Internet Comput.* **2015**, *19*, 48–56. [[CrossRef](#)]
9. Linnap, M.; Rice, A. Managed Participatory Sensing with YouSense. *J. Urban Technol.* **2014**, *21*, 9–26. [[CrossRef](#)]
10. Gisdakis, S.; Giannetsos, T.; Papadimitratos, P. Security, Privacy, and Incentive Provision for Mobile Crowd Sensing Systems. *IEEE Internet Things J.* **2016**, *3*, 839–853. [[CrossRef](#)]
11. Yang, D.; Xue, G.; Fang, X.; Tang, J. Incentive Mechanisms for Crowdsensing: Crowdsourcing With Smartphones. *IEEE/ACM Trans. Netw.* **2016**, *24*, 1732–1744. [[CrossRef](#)]
12. Fiandrino, C.; Anjomshoa, F.; Kantarci, B.; Kliazovich, D.; Bouvry, P.; Matthews, J.N. Sociability-Driven Framework for Data Acquisition in Mobile Crowdsensing Over Fog Computing Platforms for Smart Cities. *IEEE Trans. Sustain. Comput.* **2017**, *2*, 345–358. [[CrossRef](#)]
13. Zhao, Z.; Yu, F.R.; Leung, V.C.M. Opportunistic communications in interference alignment networks with wireless power transfer. *IEEE Trans. Wirel. Commun.* **2015**, *22*, 88–95. [[CrossRef](#)]
14. Luo, C.; Min, G.; Yu, F.R.; Chen, M.; Yang, L.T.; Leung, V.C.M. Energy-Efficient Distributed Relay and Power Control in Cognitive Radio Cooperative Communications. *IEEE J. Select. Areas Commun.* **2013**, *31*, 2442–2452. [[CrossRef](#)]
15. Zhang, Y.; Yu, R.; Nekovee, M.; Liu, Y.; Xie, S.; Gjessing, S. Cognitive machine-to-machine communications: Visions and potentials for the smart grid. *IEEE Netw.* **2012**, *26*, 6–13. [[CrossRef](#)]
16. Zhan, Y.; Xia, Y.; Liu, Y.; Li, F.; Wang, Y. Incentive-Aware Time-Sensitive Data Collection in Mobile Opportunistic Crowdsensing. *IEEE Trans. Veh. Technol.* **2017**, *66*, 7849–7861. [[CrossRef](#)]
17. Shah-Mansouri, H.; Wong, V.W.S. Profit maximization in mobile crowdsourcing: A truthful auction mechanism. In Proceedings of the 2015 IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 3216–3221.
18. Sun, J.; Ma, H. Heterogeneous-belief based incentive schemes for crowd sensing in mobile social networks. *J. Netw. Comput. Appl.* **2014**, *42*, 189–196. [[CrossRef](#)]
19. Gao, H.; Liu, C.H.; Tian, Y.; Xi, T.; Wang, W. Ensuring High-Quality Data Collection for Mobile Crowd Sensing. In Proceedings of the 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, 19–22 March 2017; pp. 1–6.
20. Dai, W.; Wang, Y.; Jin, Q.; Ma, J. An integrated incentive framework for mobile crowdsourced sensing. *Tsinghua Sci. Technol.* **2016**, *21*, 146–156. [[CrossRef](#)]
21. Wen, Y.; Shi, J.; Zhang, Q.; Tian, X.; Huang, Z.; Yu, H.; Cheng, Y.; Shen, X. Quality-Driven Auction-Based Incentive Mechanism for Mobile Crowd Sensing. *IEEE Trans. Veh. Technol.* **2015**, *64*, 4203–4214. [[CrossRef](#)]
22. Krontiris, I.; Albers, A. Monetary Incentives in Participatory Sensing Using Multi-attributive Auctions. *Int. J. Parallel Emerg. Distrib. Syst.* **2012**, *27*, 317–336. [[CrossRef](#)]
23. Yang, S.; Wu, F.; Tang, S.; Gao, X.; Yang, B.; Chen, G. On Designing Data Quality-Aware Truth Estimation and Surplus Sharing Method for Mobile Crowdsensing. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 832–847. [[CrossRef](#)]
24. Liu, F.; Zhao, S.; Weng, M.; Liu, Y. Fire Risk Assessment for Large-Scale Commercial Buildings Based on Structure Entropy Weight Method. *Saf. Sci.* **2017**, *94*, 26–40. [[CrossRef](#)]
25. Wu, D.; Yan, J.; Wang, H.; Wu, D.; Wang, R. Social Attribute Aware Incentive Mechanism for Device-to-Device Video Distribution. *IEEE Trans. Multimed.* **2017**, *19*, 1908–1920. [[CrossRef](#)]
26. Maharjan, S.; Zhu, Q.; Zhang, Y.; Gjessing, S.; Basar, T. Dependable Demand Response Management in the Smart Grid: A Stackelberg Game Approach. *IEEE Trans. Smart Grid* **2013**, *4*, 120–132. [[CrossRef](#)]
27. Liu, C.; Li, K.; Xu, C.; Li, K. Strategy Configurations of Multiple Users Competition for Cloud Service Reservation. *IEEE Trans. Parallel Distrib. Syst.* **2016**, *27*, 508–520. [[CrossRef](#)]
28. Wu, D.; Zhang, F.; Wang, H.; Wang, R. Strategy Configurations of Multiple Users Competition for Cloud Service Reservation. *Future Gener. Comput. Syst.* **2018**, *87*, 803–815. [[CrossRef](#)]
29. Wu, D.; Liu, Q.; Wang, H.; Wu, D.; Wang, R. Socially Aware Energy-Efficient Mobile Edge Collaboration for Video Distribution. *IEEE Trans. Multimed.* **2017**, *19*, 2197–2209. [[CrossRef](#)]
30. Wang, R.; Yan, J.; Wu, D.; Wang, H.; Yang, Q. Knowledge-Centric Edge Computing Based on Virtualized D2D Communication Systems. *IEEE Commun. Mag.* **2018**, *56*, 32–38. [[CrossRef](#)]

31. Zhang, Z.; Zhang, P.; Liu, D.; Sun, S. SRSM-Based Adaptive Relay Selection for D2D Communications. *IEEE Internet Things J.* **2018**, *5*, 2323–2332. [[CrossRef](#)]
32. Yuan, J.; Zheng, Y.; Xie, X.; Sun, G. Driving with Knowledge from the Physical World. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011; pp. 316–324.
33. Yuan, J.; Zheng, Y.; Zhang, C.; Xie, W.; Xie, X.; Sun, G.; Huang, Y. T-Drive: Driving Directions Based on Taxi Trajectories. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, San Jose, CA, USA, 2–5 November 2010; pp. 99–108.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).