Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 8959635, 12 pages
https://doi.org/10.1155/2018/8959635

Research Article

Achieving Incentive, Security, and Scalable Privacy Protection in Mobile Crowdsensing Services

Jinbo Xiong,[1,2] Rong Ma,[1] Lei Chen,[3] Youliang Tian,[2] Li Lin,[1] and Biao Jin[1]

[1]*College of Mathematics and Informatics, Fujian Normal University, Fuzhou, 350117, China*
[2]*Guizhou Provincial Key Laboratory of Public Big Data, Guizhou University, Guiyang, 550025, China*
[3]*College of Engineering and Computing, Georgia Southern University, GA 30458, USA*

Correspondence should be addressed to Youliang Tian; youliangtian@163.com

Received 9 March 2018; Revised 4 June 2018; Accepted 31 July 2018; Published 12 August 2018

Academic Editor: Kim-Kwang Raymond Choo

Mobile crowdsensing as a novel service schema of the Internet of Things (IoT) provides an innovative way to implement ubiquitous social sensing. How to establish an effective mechanism to improve the participation of sensing users and the authenticity of sensing data, protect the users' data privacy, and prevent malicious users from providing false data are among the urgent problems in mobile crowdsensing services in IoT. These issues raise a gargantuan challenge hindering the further development of mobile crowdsensing. In order to tackle the above issues, in this paper, we propose a reliable hybrid incentive mechanism for enhancing crowdsensing participations by encouraging and stimulating sensing users with both reputation and service returns in mobile crowdsensing tasks. Moreover, we propose a privacy preserving data aggregation scheme, where the mediator and/or sensing users may not be fully trusted. In this scheme, differential privacy mechanism is utilized through allowing different sensing users to add noise data, then employing homomorphic encryption for protecting the sensing data, and finally uploading ciphertext to the mediator, who is able to obtain the collection of ciphertext of the sensing data without actual decryption. Even in the case of partial sensing data leakage, differential privacy mechanism can still ensure the security of the sensing user's privacy. Finally, we introduce a novel secure multiparty auction mechanism based on the auction game theory and secure multiparty computation, which effectively solves the problem of prisoners' dilemma incurred in the sensing data transaction between the service provider and mediator. Security analysis and performance evaluation demonstrate that the proposed scheme is secure and efficient.

1. Introduction

Crowdsensing, also known as crowdsourced sensing, mainly originates from the notion of crowdsourcing. In recent years, crowdsourcing was fused with mobile embedded sensors (such as acceleration sensors, digital compasses, GPS, microphones, and cameras) in an ordinary user mobile device into a powerful sensing unit, which consciously or unconsciously collaborates one another through the mobile Internet to form a mobile crowdsensing network [1, 2]. Compared with the traditionally fixed deployment sensing mode, the cloud-based mobile crowdsourcing has proven to be an attractive solution to provide data storage and share services for resource-limited mobile devices in a privacy preserving manner [3]. Mobile crowdsensing has been widely applied to environmental monitoring [4], intelligent traffic systems [5], social behavior analyses [6, 7], urban management [6], public security [8] and other fields with the advantages of low deployment costs, simple maintenance, and excellent scalability.

As a novel service schema of the IoT, mobile crowdsensing provides an innovative way to implement the ubiquitous social sensing [9]. Despite its innovation, the application of mobile crowdsensing is limited by the insufficient number of perceived participants and the low data quality [10], which may seriously affect the development of mobile crowdsensing for the following reasons. First, sensing users expect to receive actual incentives, rather than providing free sensing data. Without appropriate incentives, sensing users may not be interested at all in the task of data sensing due to the facts that mobile sensing devices have to consume resources, such as battery power, computation and storage resources, and data

traffic. Second, in a mobile crowdsensing network [9], the collected sensing data may contain a significant amount of sensitive and private information with the risk of private data leakage. Therefore, users anticipate that effective measures are taken to protect their privacy when sensing data is uploaded to the service providers [11]. Third, there may be malicious activities in the course of the data transactions between the mediator and the service provider, possibly leading to loss of profits.

In order to tackle the aforementioned problems, this paper proposes a hybrid incentive mechanism based on both reputation and service return to motivate users to participate in the sensing tasks. Meanwhile, a privacy preserving data aggregation scheme is proposed to allow sensing users to upload encrypted data to an incompletely trusted mediator, enabling the mediator to acquire the sensing data aggregation for each time interval without decrypting each ciphertext. In this scenario, the mediator cannot derive additional information from its background knowledge or expect statistical data. We further discuss the prisoners' dilemma problem of data transactions between the service provider and the mediator and propose a novel secure multiparty auction mechanism to solve the problem that the service provider lowers the price in data transaction. The main contributions of this paper are as follows:

(i) A reliable hybrid incentive mechanism is proposed based on both reputation and service returns. Additional reputation rewards are given to the sensing users who continue to provide high-quality sensing data, and the quality of the user sensing data reflects the quality of service (QoS) provided by the service provider. For the purpose of obtaining better QoS, sensing users are required to provide more accurate and authentic data. This ensures a sustainable growth of the number of sensing participants and the overall QoS of sensing data.

(ii) A privacy preserving data aggregation scheme is proposed based on differential privacy and homomorphic encryption to solve the problem of private data leakage, where the sensing users can securely contribute their encrypted data. The simulation results indicate that the proposed scheme is effective and efficient, even in the scenario where the mediator has the access to sensing user's auxiliary information while user privacy is still protected.

(iii) The problem of prisoners' dilemma in the transactions between the service provider and mediator is discussed, and a novel secure multiparty auction mechanism is designed based on both auction game theory and secure multiparty computation. In the process of transactions, the parties choose to process the actual value of the transaction data based on the goal of maximizing the profit and implement the privacy preserving for the transaction data.

(iv) The security analysis shows that the encryption algorithm scheme used in this paper is provably secure. The constructed privacy protection scheme meets the security objectives, and the performance analysis and simulation results indicate that the proposed scheme is effective and efficient.

The rest of this paper is organized as follows: Section 2 covers the related work of the incentive mechanism and privacy protection scheme in the mobile crowdsensing systems. The system model, adversary model, and the security requirements of the proposed schemes are described in Section 3. In Section 4, we first describe the hybrid incentive mechanism of mobile crowdsensing system and then introduce a privacy preserving data aggregation scheme by jointly integrating differential privacy and homomorphic encryption between sensing users and the mediator. Finally, the secure multiparty auction model between service providers and the mediator is presented. Sections 5 and 6 analyze the proposed scheme in the aspects of security and performance. Section 7 points out future research directions and summarizes the entire paper.

2. Related Work

2.1. Incentive Mechanism. Various incentives strategies and methods are available in mobile crowdsensing. Generally, in regard of the form of returns, it can be divided into monetary incentives and nonmonetary incentives [12]. Monetary incentives are mainly through reward payments to encourage the sensing users to participate in sensing tacks. Based on the sensing users' quotation of the sensing data, the system selects a subset of the sensing users with lower payment costs to complete the sensing task. Monetary incentives reward the sensing users' with money, a direct and currently the most common form of incentives. One of the most important incentive mechanisms is based on the auction game theory mechanism, including reverse auction, combined auction [13], multiattribute auction, full auction, two-way auction, and vickrey-clarke-groves (VCG) auction. In addition to the above, there are many other incentives methods, such as those based on the Stackelberg game model [14, 15]. The literature [14] proposed a Stackelberg game-based pricing mechanism to inspire core users to distribute videos to the multicast users via device-to-device (D2D) communication. Nonmonetary incentives include entertainment game incentives [16], social relations incentives [15, 17], and virtual integration incentives [18]. The entertainment game incentive [16] uses game entertainment and attractiveness to encourage the sensing users to complete the sensing task. Social relation incentives refer to a type of social networking relationship where the sensing users already exist or server platform is built, and the sensing users are motivated to maintain a sense of belonging in social relations [15]. The literature [17] exploits the coalition game by employing the user's social preference list to dynamically establish virtual communities. Virtual integral incentive [18] refers to the fact that the sensing users will receive the virtual integral from the sensing task. The virtual currency converts into a real currency or some other types of physical or virtual return, which encourages the sensing users to participate in a sensing task. In order to better motivate the sensing users, more recent research

works tend to integrate two or more types of incentives (hybrid incentive). The literatures [19] use reverse auction to select the winner to receive incentives in exchange of the sensing data, while motivating the noncompliant users to remain active in the system through virtual integration. The literature [15] uses the Stackelberg game model and establishes the endorsement relationship among the sensing users to motivate their participation. The literature [20] uses three types of incentives: reverse auction payment method based on game theory, social relation incentive based on the user's reputation level, and entertainment game incentive based on the user's psychological satisfaction. The hybrid incentive method provides satisfactory incentive strategies for various situations. In order to meet the psychological requirements of the sensing users and promote the participation of sensing task, it is encouraged to integrate various incentives, such as reward incentives, entertainment incentives, spiritual incentives, honor incentives, for an optimized solution.

2.2. Privacy Protection Scheme. The security of private information in mobile crowdsensing includes the privacy of the sensing users and the security of the service provider. The sensing users anticipate that their personal data privacy is preserved during the uploading to service provider [21, 22]. Since the sensing users and the mediator may not be completely trusted, the uploaded false data may potentially cause security problems to the service provider, and therefore countermeasures to malicious users and malicious attacks should be considered. Most privacy protection schemes assume that the mediator is fully trusted, and the sensing user takes privacy protection measures in the sensing task. Each user in the t-sensing task provides real-time sensing data and chooses an anonymity level. Data anonymity [23] can be used to add noise to the real data, such as k-anonymity and others. Each user uploads the private data and indicates data anonymity level to the mediator without knowing the privacy preferences of other users. The mediator collects all anonymously processed data sets and the anonymity levels from all sensing users, followed by trading them with the service provider. Wu [9] combined key distribution with trust management to construct a novel dynamic trust relationships-aware data privacy protection (DTRPP) mechanism for mobile crowdsensing. Zhang et al. [24] proposed a novel technique called match-then-decrypt, in which a matching phase is additionally introduced before the decryption phase. Rastogi and Nath [25] considered aggregating the sum statistical sum in the presence of an untrusted mediator, proving that the mediator cannot calculate a linear combination of user values other than the sum. However, this implicit security definition is not complete in a sensing task, and it requires the aggregator to interact with the participants in order to decrypt ciphertext for each time interval. Rieffel et al. [26] considered a specific application scenario in which the manager attempts to decrypt the statistical sum of a group of users on a regular basis without decrypting a single value. Ni et al. [27] proposed a fog-assisted mobile crowdsensing framework, enabling fog nodes to allocate tasks based on users' mobility for improving the accuracy of task assignment.

Table 1: Notations and descriptions.

Notations	Descriptions
g	a random generator
ε	a measure in differential privacy
Δ	the sensitivity
δ	the probability of adding noise
n	the number of sensing users
x_i^t	the sensing data of user i at t-th task
a_i	the data feature set
b_i	the class tag
v	the final transaction price
$\widehat{x}_i$	added noise sensing data
r_i	noise
$H(t)$	anti-collision hash function
sk_i	the sensing user's key
sk_0	the mediator's key
c_i	the ciphertext
V	the aggregation data
$\{sp_1, sp_2, \ldots, sp_m\}$	a list of service providers
r_0	a random number
$\{price_1, price_2, \ldots, price_m\}$	a list of the expected cost price
Z	the mediator

But their construction falls short in completely resisting against the collusion attack; in other words, users may collude with managers to decrypt the victim's data. Therefore, it is necessary to design a novel privacy protection scheme immune from collusion attacks.

In summary, limitations exist in the existing privacy protection schemes in mobile crowdsensing systems, mainly due to neglecting the consideration of mediator being not completely trusted. There exist security threats to the privacy protection methods used by sensing users (e.g., an anonymous antibackground knowledge attack). They also lack the privacy protection solutions against malicious users colluding with one another [28]. At the same time, the incentive mechanism of the mobile crowdsensing system is too simple to be effective. How to design an effective hybrid incentive mechanism to ensure high participation of sensing users in the sensing task while providing long-term high-quality data is of great importance.

3. Problem Description

This section first gives the notations and descriptions in Table 1 and then introduces our system model, followed by an adversary model, security requirements, and our design goals.

3.1. System Model. The mobile crowdsensing system consists of the following three main entities, as shown in Figure 1.

Crowdsensing Users. They are the participants who collect the sensing data using their personal mobile-aware devices (such as intelligent terminal equipment, wearable equipment, and automotive equipment). Crowdsensing users participate

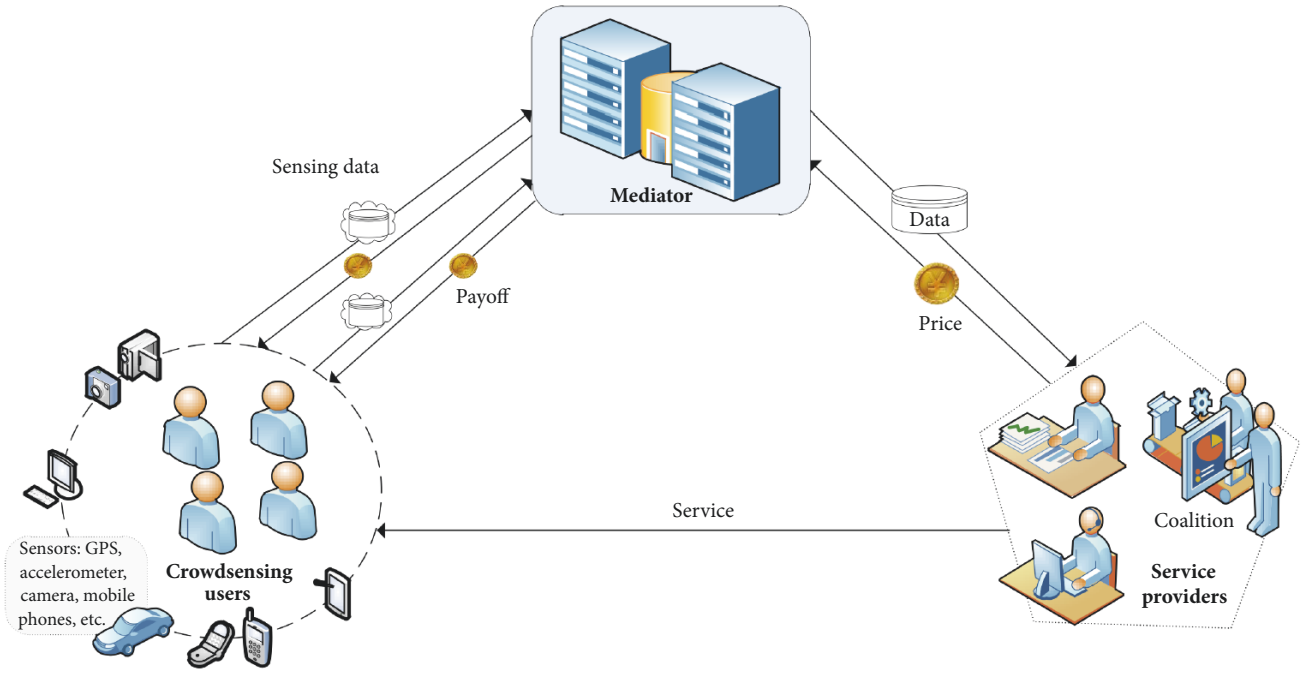


Figure 1: System model of mobile crowdsensing.

in sensing tasks and receive the corresponding maximal revenue via the mobile network by uploading the sensing data to the mediator. Continuous participation will help acquire additional reputation incentives. In order to implement privacy protection for the sensing data, crowdsensing users usually encrypt the data using the homomorphic encryption algorithm [22] with added noise, followed by uploading the ciphertext to the mediator. The quality of the sensing data provided by sensing user is reflected in the final quality of service (QoS) provided by the service provider: the higher the quality of the data provided by the sensing user, the higher the QoS be returned to the customers.

Service Providers. Due to the different requirements of service customers, service providers are responsible for participating in the final aggregated data transactions and providing various services to the customers. The aggregated sensing data received by the service provider is then used for machine learning, data visualization and other studies. A rational service provider aims to acquire higher value data from a mediator at a reasonable price. To reduce the cost of purchase, service provider may choose to share the data with shared other providers to average the total costs.

Mediators. They interact with both the sensing users and the service providers. Under the privacy protection mechanism though combining the differential privacy with homomorphic encryption, the mediator advertises the sensing task to the mobile crowdsensing users and adopts the hybrid incentive mechanism to attract more users to upload their encrypted sensing data. The mediator aggregates and sells the sensing data to the service providers to receive the corresponding rewards.

3.2. Adversary Model. Cryptographic Hash function is used for securing both the sensing user key and the mediation key of the homomorphic encryption algorithm. We assume that the Hash function is cryptologically secure. More specifically, it is assumed that the Hash functions used in our schemes are resistant against the weak collision attacks and the strong collision attacks [29]. Moreover, we assume that the mediator is a semitrusted data aggregator.

In our adversary model, we consider a strong attacker $\mathcal{A}$, who cannot only listen to all the communication data in our system model but also launches the following attacks:

(i) Attacker $\mathcal{A}$ may intercept one single sensing user's private data in the process of uploading data by consuming a substantial amount of cost. However, there are a large number of sensing users in the system, it will be extremely high cost for attacker $\mathcal{A}$ to intercept private data from each individual sensing user. Therefore, the attacker $\mathcal{A}$ may attempt to analyze the privacy of other sensing user's by using the user's private key which has already intercepted [30].

(ii) Attacker $\mathcal{A}$ may obtain private data by colluding with an incompletely trusted mediator. By doing this, attacker $\mathcal{A}$ will be able to access a large amount of sensing data stored at the mediator.

(iii) Attacker $\mathcal{A}$ can break a small number of sensing users, and attempt to obtain the other users' private keys and decrypt the ciphertext of the sensing data.

3.3. Security Requirements. The reliability and efficiency of the mobile crowdsensing system depends on the security of the communication system. Mobile crowdsensing systems are increasingly complex, interactive, and dynamic and therefore require advanced network technologies and complex security protocols to address potential security vulnerabilities. The design of a privacy protection mechanism in mobile crowdsensing systems needs serious and comprehensive consideration of the security of communications. In order to prevent attacker $\mathscr{A}$ from obtaining user's private data, it is desired to achieve the following security requirements:

(i) Even if $\mathscr{A}$ can listen to the communication data flow, he still could not obtain the private data from any sensing user.

(ii) Even if $\mathscr{A}$ can break into individual sensing user device, he still could not acquire other sensing users' private data.

(iii) Even if $\mathscr{A}$ can collude with the mediator to access the aggregation results of the sensing data, $\mathscr{A}$ still could not obtain the sensing user's personal private data.

(iv) Even if $\mathscr{A}$ can break into a small number of sensing users to access their private key, $\mathscr{A}$ still could not get the sensing user's original personal data.

3.4. Design Goals. With the above system model, adversary model, and security requirements, our goals are to propose a reliable hybrid incentive mechanism, an efficient privacy preserving data aggregation scheme, and a secure multiparty auction mechanism in mobile crowdsensing system. Specifically, the following objectives should be achieved:

(i) The proposed incentive mechanism should be reliable and effective. The sensing effect is closely related to the number of participants and the quality of data provided by the sensing users in mobile crowdsensing. Incentive mechanism must ensure that a sufficient number of sensing users have long-term involvement in sensing tasks and provide reliable data.

(ii) The proposed privacy preserving data aggregation scheme should meet the security requirements. As mentioned above, if the security and privacy issues are not considered in the mobile crowdsensing system, the privacy of the individual sensing user will be disclosed, hindering the further development and application of mobile crowdsensing system. Therefore, the proposed scheme must be able to meet the above security requirements.

(iii) The proposed aggregation scheme should be highly efficient in communication. While the sensing user and the mediator communicate via high-bandwidth, low-latency wired/wireless connections, it is essential to support a large number of sensing users simultaneously sending data to the mediator. The proposed scheme should consider the efficiency of communication, so that real-time sensing data can be sent to the intermediary in a timely manner.

(iv) The proposed data transaction mechanism should ensure the security of the service providers and the mediator. Service providers participate in data transaction process with rational choices to solve the problem of the prisoners' dilemma and achieve the goal of maximizing payoff without unnecessary disclosure of the parties' private information.

4. Construction

This section elaborates the details of the proposed reliable hybrid incentive mechanism for mobile crowdsensing, the privacy preserving data aggregation scheme combining differential privacy with homomorphic encryption, and the secure multiparty auction mechanism based on auction game theory.

4.1. Reliable Hybrid Incentive Mechanism. The data $X_i^t = \{(a_i, b_i)\}_{i=1}^L$ perceived by the sensing user i in the tth perceptual task is a tuples containing the data feature set $a_i \in \mathbb{R}^M$ and the class tag $b_i \in \mathbb{R}$, where L is the number of data tuples and M is the number of data attributes [31]. Feature set a_i consists of sensing data, such as GPS data in smart trip services and personalized recommended social network behavior data. Class tag b_i contains human inputs and is only available in supervised data analytics. After collecting sufficient sensing data, the service provider analyzes the sensing data and builds data-based services. For example, in intelligent transportation services, the use of mobile sensing equipment and urban traffic information collection analyses can provide consumers with more efficient and convenient travel route planning and auxiliary driving information support. We define the expression $U(X_i^t) = u(X_i^t) + R^t + Q$ to represent all of the final proceeds obtained by the sensing data X_i^t in the tth sensing task, which consists of the following three parts:

(i) *Participation income* $u(X_i^t)$. Sensing users choose to participate in a sensing task to obtain their participation income, which depends on the price paid by the service provider in the final transaction with the mediator. Assuming that N sensing users participate in the same sensing task, the transaction price of the final transaction service provider is v; then each participant's revenue in the tth sensing task is $u(X_i^t) = v/N$.

(ii) *Reputation points* R^t. Sensing user participating in the tth sensing task will earn their reputation points. The user's reputation points with continued participation in the sensing task can be accumulated to encourage long-term participation in the sensing task. However, the reputation points will be cleared when the sensing task is interrupted or revoked.

(iii) *Feedback service quality* Q. The quality of the sensing data will reflect the QoS provided by the final service provider to the consumers. Therefore, the system encourages the sensing users to provide data with

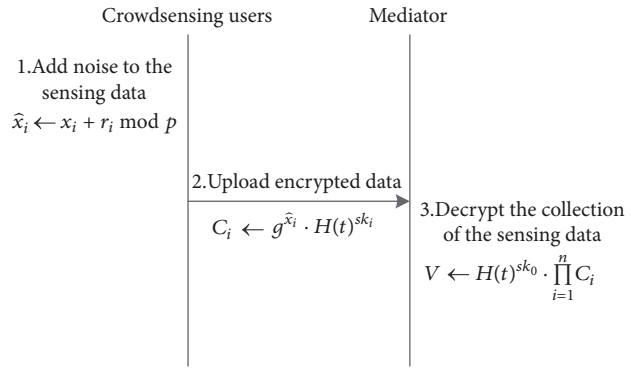


Figure 2: The interactive process of privacy protection.

high level of authenticity to obtain better quality service.

The sensing users can obtain three kinds of incentive income in the hybrid incentive mechanism: reward payment, reputation integral incentive, and service quality incentive. In order to meet the sensing users' psychological expectation and material requirements, the system promotes long-term access to the perceptual task with highly accurate sensing data.

4.2. Privacy Preserving Data Aggregation Scheme. The privacy preserving data aggregation scheme allows the sensing users to upload the encrypted data to an incompletely trusted mediator. The mediator obtains the aggregation of sensing data for each time interval without decrypting each ciphertext. The interaction between the sensing user and mediator is illustrated in Figure 2. The marriage of the differential privacy and homomorphic encryption in privacy preserving data aggregation scheme includes the following three phases: adding noise data, data encryption, and data decryption.

4.2.1. Adding Noise Data. In a mobile crowdsensing system, it is assumed that the sensing user is capable of arbitrarily adding noise to sensing data based on privacy preferences before uploading. This may lead to a serious deviation between the final data aggregation statistics and the actual results, thereby significantly reducing the availability of data collection. For this reason, we introduce a distributed differential privacy protection mechanism [32], where each sensing user only needs to add a small amount of noise with randomness, which still ensures that data privacy is well preserved. The specific description of this phase is shown in Algorithm 1.

4.2.2. Data Encryption. In the proposed model, it is assumed that the mediator is not fully trusted. The sensing user uploads the ciphertext encrypted though the homomorphic encryption [22] data to the mediator, who can only obtain the data aggregation result without knowing the individual private information of any sensing user. In order to counter against background knowledge attack, data encryption is discussed using data collection as an example. In the sensing

input: $\alpha = \exp(\varepsilon/\Delta)$, $\beta = (1/r^n)\log(1/\delta)$, $X = \{x_1, \ldots, x_n\}$
output: $\widehat{X} = \{\widehat{x}_1, \ldots, \widehat{x}_n\}$
(1) **for** each $i \in [n]$ **do**;
(2) Sample noise r_i according to the following;
(3)

$$r_i \longleftarrow \begin{cases} Geom(\alpha) & \text{with probability } \beta \\ 0 & \text{with probability } 1-\beta \end{cases}$$

(4) Randomize data by computing $\widehat{x}_i \longleftarrow x_i + r_i$ **mod** p;
return $\widehat{X}$

Algorithm 1: Adding noise data.

input: $\widehat{X} = \{\widehat{x}_1, \ldots, \widehat{x}_n\}$, $SK = \{sk_1, \ldots, sk_n\}$
output: $C_i = \{c_1, \ldots, c_n\}$
(1) **for** each $i \in [n]$ **do**;
(2) $M_{i,t} = H(t)^{sk_i}$;
(3) Encrypted message by computing $c_i \longleftarrow g^{\widehat{x}_i} \cdot M_{i,t}$;
return C_i

Algorithm 2: Data encryption.

task, the random sequence generator assigns the encryption keys $sk_1, \ldots, sk_n$ to the different sensing users and the decryption key sk_0 to the mediator, which satisfies $\sum_{i=0}^{n} sk_i = 0$. Each sensing user calculates the function $M_{i,t} = H(t)^{sk_i}$, $i \in [n]$ according to its encryption key and the hash function $H(x)$. Then the mediator calculates function $M_{0,t} = H(t)^{sk_0}$. Since $\sum_{i=0}^{n} sk_i = 0$, therefore $\prod_{i=0}^{n} M_{i,t} = 1$. Using this property, the homomorphic addition encryption of the perceptual data can be implemented. The users encrypt the sensing data using the respective encryption keys to obtain the ciphertext c_i:

$$c_i \longleftarrow g^{\widehat{x}_i} \cdot H(t)^{sk_i}. \tag{1}$$

The specific description of the data encryption phase is shown in Algorithm 2.

4.2.3. Data Decryption. The mediator receives the user's uploaded ciphertext and decrypts it with the corresponding decryption key sk_0 by calculating

$$V \longleftarrow H(t)^{sk_0} \cdot \prod_{i=1}^{n} C_i. \tag{2}$$

The specific description of the data decryption phase is shown in Algorithm 3.

In this privacy preserving data aggregation scheme, the mediator cannot decrypt the private data of a particular sensing user even with potential assistance information. In addition to knowing whether a particular user participates in a perceptual task, the mediator only obtains the statistical results of the data aggregation.

input: $sk_0, C_i = \{c_1, \ldots, c_n\}$
output: V
(1) Aggregation by computing $\prod_{i=1}^{n} C_i = c_1 \times c_2 \times, ..., \times c_n$;
(2) $M_{0,t} = H(t)^{sk_0}$;
(3) Decrypted message by computing $V \longleftarrow H(t)^{sk_0} \cdot \prod_{i=1}^{n} C_i$;
(4) $V = g^{\sum_{i=1}^{n} \tilde{x}_i}$;
return V

Algorithm 3: Data decryption.

4.3. Secure Multiparty Auction Mechanism. In addition to relying on sensing users to actively provide high-quality data, mobile crowdsensing system also depends on the mutual trust transactions between the service providers and mediator. However, in the actual transaction process, due to the unequal information accessibility, the data transaction process can be seen as two incomplete information static games [33]. In order to maximize their own profits, service providers tend to purchase data at a lower price, regardless of the quality of the data provided by the mediator. For the mediator, regardless of the price provided by the service provider, providing low-quality data may lead to more revenue. Based on this, the Nash equilibrium solution of the game for the service provider and mediator is <low quality data, low price>, which falls into the prisoner's dilemma. Obviously, this is the worst result, which will lead to lower yields on the mediator, resulting in less revenue for the sensing users. In the proposed privacy preserving data aggregation scheme, since the sensing data is uploaded to the mediator in ciphertext, data accuracy cannot be evaluated. In other words, the real price/value of the data is unknown. This is even more detrimental to data transactions between service providers and mediator.

4.3.1. Secure Multiparty Computation. Using the secure multiparty computation, each service provider can enter the expected price of the data provided by the mediator, respectively. The computational functions $f(x_1), f(x_2), \ldots, f(x_n)$ negotiate the true price of the obtained data. At the end of the protocol, each service provider cannot receive any other information except the value of $f(x_i)$.

Suppose there are more than three service providers $sp_1, sp_2, \ldots, sp_m$ and a mediator Z to provide data for the expected cost price $price_1, price_2, \ldots, price_m$, and $price_0$, respectively. Since the service providers are mutual competitors, it is safe to assume that there is no collusion among them.

To calculate the starting price $price = (\sum_{i=0}^{m} price_i)/(m + 1)$, let the mediator Z generate a random number r_0 and passes $r_0 + price_0$ to service provider sp_1. Then sp_1 calculates $r_0 + \sum_{i=0}^{1} price_i$ and passes it to sp_2, and this process continues. The mediator Z finally obtains $r_0 + \sum_{i=0}^{m} price_i$ and subtracts r_0, where the mediator and its service providers obtain the expected cost of the sum of the price. The mediator Z divides the total price with the number of participants in the calculation to obtain the starting price and publishes it to each service provider.

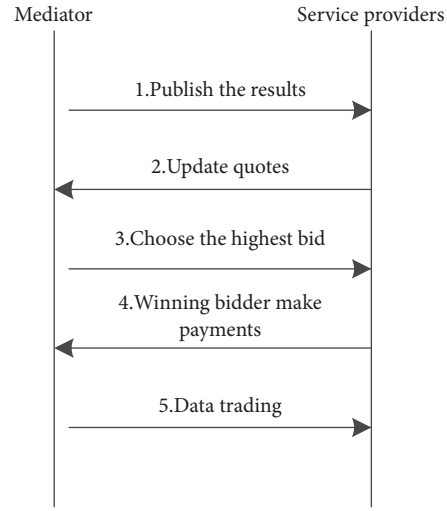


Figure 3: The interactive process of auction mechanism.

4.3.2. Auction Mechanism. Assuming the rationality that the service providers participate in the auction, they choose the calculation strategy for maximized benefit. The interaction between the mediator and service providers is shown in Figure 3.

The mediator first announces the auction starting price from the last step of the secure multiparty computation and, then, the service providers participating in the auction update their quotations according to the starting price adjustment strategy. To close the auction, the mediator selects the service provider with the highest offer as the winning bidder from multiple quotes, followed by data trading to complete this process.

5. Security Analysis

As addressed earlier, we consider scenario with the existence of a strong attacker in the system, who cannot only monitor the communication channel in the system but also collude with the mediator and gain access to the information stored in the mediator. Additionally, attacker $\mathcal{A}$ can break into a small number of sensing users and obtain these users' private keys. Since this paper primarily considers protecting data privacy for the sensing users, attackers tampering with the messages are beyond the scope this particular research, although this

can be prevented by enforcing strong authentication. In this section, we analyze in detail how the proposed schemes are resistant against a variety of attacks by the attacker.

(i) Our scheme can guarantee that the ciphertext submitted by the sensing users will not disclose their personal or private data. As mentioned in the security model, the attacker $\mathcal{A}$ may monitor the data transmitted to the mediator during the user upload phase. Our privacy preserving data aggregation scheme ensures that even if the attacker $\mathcal{A}$ listens to all ciphertext for all sensing users in each time interval, he still cannot derive any plaintext information from these ciphertext, and therefore data privacy is guaranteed. The encryption scheme used in this research is a one-way trap function, based on the Diffie-Hellman hypothesis that the advantage of a polynomial time attacker $\mathcal{A}$ is negligible [30]. At this point, our scheme is to specify what is IND-CPA security.

(ii) In our system, we consider that the attacker $\mathcal{A}$ may break into individual user mobile crowdsensing devices and obtain their private keys. In this case, a user's uploaded ciphertext data will undoubtedly be completely exposed. However, due to the large number of users in the system, the attacker $\mathcal{A}$ may not want to use this costly approach to break into multiple individual user devices. Alternatively, attacker $\mathcal{A}$ may expect to be able to use the private key that he has obtained to analyze the private keys of other users. However, this attack would not be successful, because each user's private key is randomly generated. As a result, data privacy for noncompromised users can still be guaranteed.

(iii) The user's private data will not be exposed at the mediator. At any time interval, the mediator cannot decrypt any ciphertext to obtain the corresponding plaintext information x_i even after collecting the ciphertext from all the sensing users. Therefore, even if the attacker $\mathcal{A}$ colludes with the mediator, $\mathcal{A}$ can only have access to the users' ciphertext. On the other hand, the mediator does not have the corresponding decryption key of any ciphertext and, as a result, $\mathcal{A}$ cannot either obtain the corresponding decryption key for decryption. Moreover, the aggregated ciphertext can only be decrypted with the mediator's random key sk_0, and $\mathcal{A}$ neither knows any users' keys nor knows the mediator's random key sk_0. Therefore, $\mathcal{A}$ cannot break the aggregated ciphertext. Even if the attacker $\mathcal{A}$ obtains the mediator's random key sk_0, the result of the decryption $X + R$ is the aggregated result of both ciphertext and added noise. Based on the above analysis, the user's private data and data aggregation results will not be exposed to the strong attacker $\mathcal{A}$.

(iv) The original data from the sensing user attacked by $\mathcal{A}$ will not be compromised. In the proposed system, we consider that the attacker $\mathcal{A}$ has the ability to obtain the sensing user's private key, and subsequently decrypts the user's uploaded ciphertext. As a countermeasure, we introduce the geometric distribution differential privacy for adding noise to the original sensing data. In this case, even if the ciphertext uploaded by the compromised user is completely leaked, the decrypted data is still different from the original data due to the added noise for protecting original data.

6. Performance Analysis and Evaluation

6.1. Complexity Analysis. This paper mainly focuses on the complexity analysis of the algorithm based on the computation costs at the sensing users, mediator, and service providers. In each time interval, the sensing user's computational cost is mainly due to the addition of differential noise and the use of random keys to encrypt the sensing data. Let T_m be the time of modulo multiplication operation and T_e be the time of modulo exponentiation operation. In our scheme, the sensing user generates a ciphertext using formula $c_i \longleftarrow g^{\tilde{x}_i} \cdot H(t)^{sk_i}$. Therefore, the sensing user needs to calculate the modulo exponential operation twice and the modulo multiplication operation once, i.e., $2T_e + T_m$. Also in each time interval, each sensing user needs to use the distributed geometric distribution to add noise. The time to add noise is however negligible compared to the modulo operation. The computational cost of the mediator is mainly reflected in the collection of the decryption of aggregated data, with decryption formula $V \longleftarrow H(t)^{sk_0} \cdot \prod_{i=1}^{n} C_i$. Therefore, the mediator needs to calculate one modulo multiplication operation and one modulo exponential operation, i.e., $T_e + T_m$. The computational cost of the service provider incurs due to the participation in the secure multiparty computation with the time complexity $O(1)$. In our scheme, the total communication overhead is $O(n)$, and its individual user's communication overhead is $O(1)$.

6.2. Performance Evaluation. In this section, we evaluate the performance of the privacy protection scheme proposed in this paper, mainly focusing on the two aspects of calculation time cost and statistical error. In terms of calculating the time cost, we consider the effect of different variables on the calculation time in the three phases of encryption, aggregation, and decryption. In terms of statistical error, we compare the geometric distribution differential privacy with the addition of the random noise. We conducted a set of simulation experiments using C++ on a Linux computer with the following hardware and configurations: Intel Xeon E5-2650 v3 @ 2.30 GHz CPU, 8 GB RAM, Ubuntu 16.04LTS Operating System, and GCC 5.4.0 Compiler.

The experiment tests the calculation running time in data encryption phase, data aggregation phase, and data decryption phase, respectively. Experiment runs 1000 times to receive the average as the experimental results.

(1) Data encryption phase. The experiment tests the running time using symmetric encryption with user key for different sizes of sensing data, which selects 7 randomly generated different sizes of sensing data from 32 bits to 2048

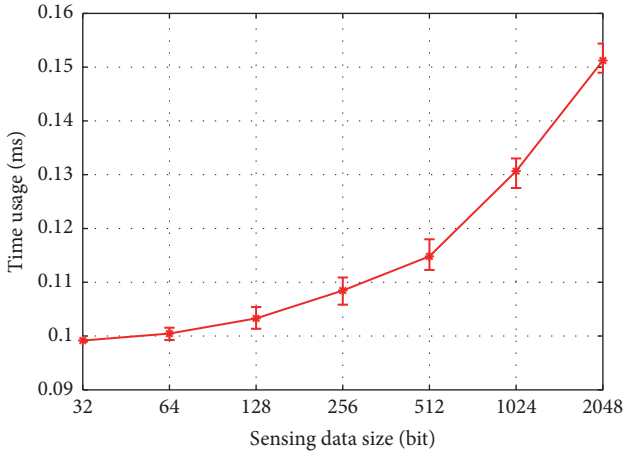


Figure 4: Running time of data encryption.

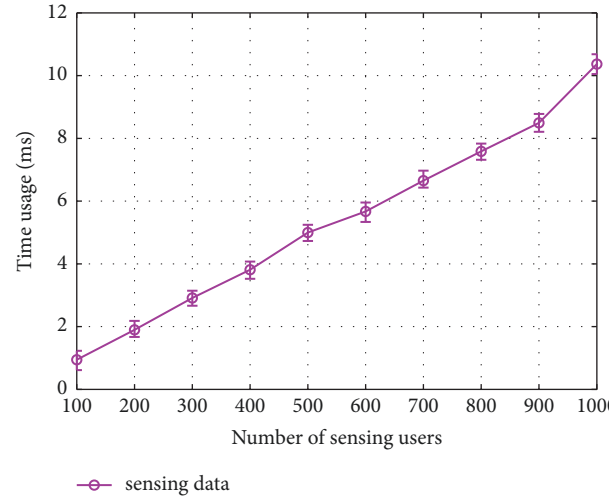


Figure 5: Running time of data aggregation.

bits: 32 bits, 64 bits, 128 bits, 256 bits, 512 bits, 1024 bits, and 2048 bits, as shown in Figure 4.

The data encryption phase mainly considers the relationship between the size of the plaintext and the time consumed for encryption. The plaintext uses a binary string encoding, which is expressed as the length of the string in bits. The plaintext content selection uses randomly generated values that are uniformly distributed. As the size of the plaintext increases, the time required for encryption increases significantly, as shown in Figure 4. As the data length increases, the encryption running time also increases. When the data length reaches 512 bits, the data encryption time is clocked at only 0.112 milliseconds. It is also worth noting that both the ciphertext and the parameter selection within the encryption process affect the output of the encryption. We use multiple sets of generators to randomly generate ciphertexts by repeating experiments to observe their effects in time. Small localized fluctuations, not affecting the overall trend, can be found in Figure 4. Therefore, it can be considered that, in a certain range of errors, the time required for the encryption process is exponential to the size of the plaintext as expected.

(2) Data aggregation phase. The experiment tests the running time of different numbers of sensing data aggregations. Experiments were performed in turn from 100 to 1000 different numbers of the sensing users, i.e., 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1000, as shown in Figure 5. Single data block with a size of 1024 bits is used for 5 different sets of sensing data for polymerization time tests.

We expect that a large influx of sensing users under the current system architecture will not cause a sudden change in system stability, where the number of participating users has a linear growth with the privacy preserving data aggregation time cost. We observe the time cost of the system under a single privacy preserving data aggregation with different users. As shown in Figure 5, the amount of time required for privacy preserving data aggregation increases almost linearly with the number of sensing users. When this number reaches 1000, the time of aggregating data is clocked at only about 10 milliseconds. Considering that the content submitted by

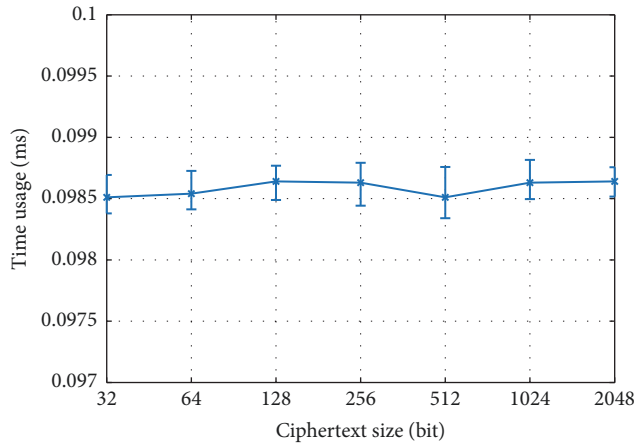
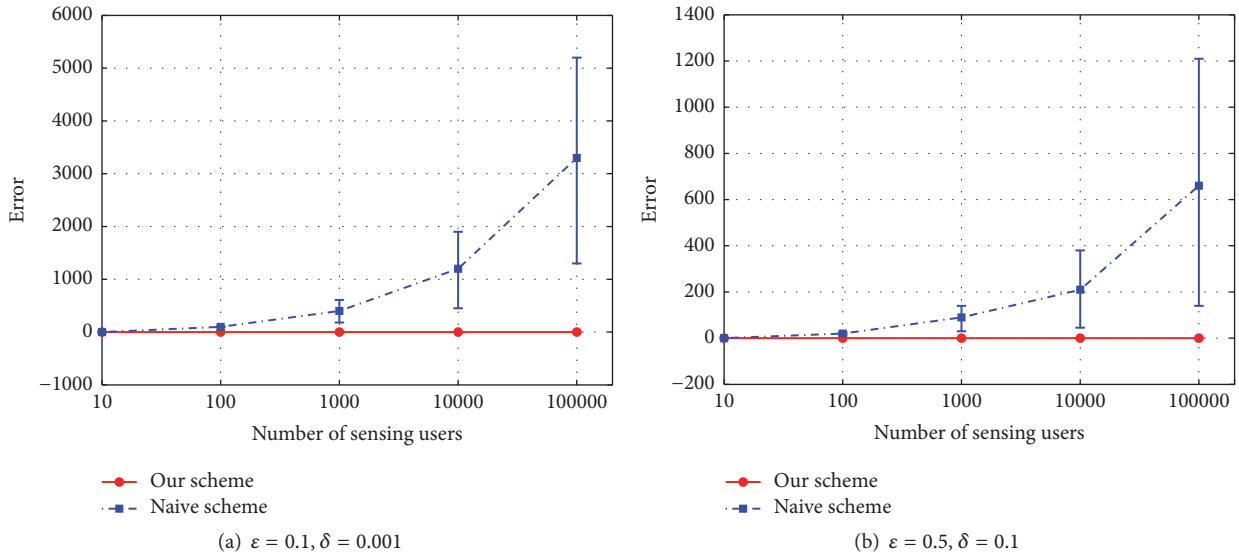


Figure 6: Running time of data decryption.

the sensing users and the aggregation of the internal operations will affect the efficiency of aggregation, our scheme is designed to use multiple sets of the same number of values. By repeating the experiment, it is discovered that the time cost required for the aggregation aligns to the overall trend of the user's participation. Therefore, in a reasonable range of error, the privacy preserving data aggregation is considered stable without causing any sudden change in system overhead as the number of sensing users increases rapidly.

(3) Data decryption phase. This experiment tests the running time of data decryption of the encrypted sensing data of different content. Experiments randomly select 5 groups of encrypted sensing data selected from 32 bits to 2048 bits: 32 bits, 64 bits, 128 bits, 256 bits, 512 bits, 1024 bits, and 2048 bits, as shown in Figure 6.

It is shown that the impact of decryption time variations can be ignored. Specifically, the decryption is only concerned with the results of the previous aggregation. Most of the time spent in the decryption phase is on decrypting aggregated data using decryption key. As shown in Figure 6, there is no



(a) $\varepsilon = 0.1, \delta = 0.001$ (b) $\varepsilon = 0.5, \delta = 0.1$

Figure 7: Error comparison.

significant change in the time cost of decryption as the data size increases. In the vicinity of 0.0985 milliseconds, the trend of the curve is consistent with the computation complexity of the scheme. In other words, the sensing user's overall change will not cause any significant time delay. At the same time, we consider decrypting the results of the data aggregation, and decryption of the internal parameters will jointly affect the decryption efficiency. It is designed to use the overall performance of multiple groups of users and the aggregation of multiple user data to observe the overall trend of the fluctuations. It can be seen from Figure 6 that its fluctuation is confined within a small and permittable range and therefore the data decryption process can be considered stable.

(4) Error comparison of geometric distribution with noise reduction and random noise reduction.

In Figure 7, the x-axis represents the number of sensing users, and the y-axis represents the mean value of the error (absolute). The naive scheme is that each sensing user adds independent geometric noise to his sensing data and uploads the perturbed data to the mediator. It is clear that this approach will cause a significant discrepancy between the aggregated data and the original values. Our scheme uses the distributed geometric plus noise to simulate the discrete Laplacian de-noising, which solves this problem effectively.

7. Conclusion

With the continuous service expansion and extension of mobile crowdsensing systems, privacy protection schemes and incentive mechanisms are highly demanded for their adoptions to the dynamic heterogeneous sensing systems. This paper proposed a reliable hybrid incentive mechanism based on both reputation and service return to inspire more high-quality sensing users to participate in mobile crowdsensing tasks. We constructed a privacy preserving data aggregation scheme based on the assumption that the mediator and/or sensing users as incompletely credible. Homomorphic encryption is used to allow the mediator to decrypt the collection of the sensing data from multiple ciphertext that are encrypted using different sensing user keys. The proposed privacy preserving data aggregation scheme also utilizes the differential privacy mechanism by adding noise to the sensing data, ensuring the security of the remaining data of the sensing users even with the exposure of partial data. We discussed the security problem of data transaction between service provider and mediator and put forwarded a novel secure multiparty auction mechanism based on both the auction game theory and secure multiparty computation to solve the problem of prisoner's dilemma at the service providers. Finally, we proved the security and efficiency of the proposed schemes through security analyses and simulation experiments. Future study will focus on how to improve calculation efficiency and the accuracy-privacy tradeoff for mobile crowdsensing.

Data Availability

Our data is obtained through simulation experiments, and we can provide it if needed.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported in part by the Natural Science Foundation of China (61772008, 61402109, 61502489, 61502248, and 61502102) and Guizhou Provincial Key Laboratory of Public Big Data Research Fund (2017BDKFJJ028).

References

[1] Z. Xu, L. Mei, K. R. Choo et al., "Mobile crowd sensing of human-like intelligence using social sensors: a survey," *Neurocomputing*, vol. 279, pp. 3–10, 2018.

[2] H. Ma, D. Zhao, and P. Yuan, "Opportunities in mobile crowd sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 29–35, 2014.

[3] Y. Miao, J. Ma, X. Liu, X. Li, Z. Liu, and H. Li, "Practical attribute-based multi-keyword search scheme in mobile crowdsourcing," *IEEE Internet of Things Journal*, 2017.

[4] S. Kim, C. Robson, T. Zimmerman, J. Pierce, and E. M. Haber, "Creek watch: pairing usefulness and usability for successful citizen science," in *Proceedings of the 29th Annual CHI Conference on Human Factors in Computing Systems (CHI '11)*, pp. 2125–2134, ACM, Vancouver, Canada, May 2011.

[5] J. Ni, A. Zhang, X. Lin, and X. S. Shen, "Security, privacy, and fairness in fog-based vehicular crowdsensing," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 146–152, 2017.

[6] Z. Xu, H. Zhang, C. Hu et al., "Building knowledge base of urban emergency events based on crowdsourcing of social media," *Concurrency and Computation: Practice and Experience*, vol. 28, no. 15, pp. 4038–4052, 2016.

[7] Y. Zhang, D. Zheng, and R. H. Deng, "Security and privacy in smart health: efficient policy-hiding attribute-based access control," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2130–2145, 2018.

[8] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *Proceedings of the 11th IEEE International Conference on Pervasive Computing and Communications (PerCom '13)*, pp. 193–200, San Diego, Calif, USA, March 2013.

[9] D. Wu, S. Si, S. Wu, and R. Wang, "Dynamic trust relationships aware data privacy protection in mobile crowd-sensing," *IEEE Internet of Things Journal*, 2017.

[10] Y. Wen, J. Shi, Q. Zhang et al., "Quality-driven auction-based incentive mechanism for mobile crowd sensing," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 9, pp. 4203–4214, 2015.

[11] J. Xiong, J. Ren, L. Chen et al., "Enhancing privacy and availability for data clustering in intelligent electrical service of IoT," *IEEE Internet of Things Journal*, 2018.

[12] T. Luo, S. S. Kanhere, J. Huang, S. K. Das, and F. Wu, "Sustainable incentives for mobile crowdsensing: Auctions, lotteries, and trust and reputation systems," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 68–74, 2017.

[13] Z. Feng, Y. Zhu, Q. Zhang, L. M. Ni, and A. V. Vasilakos, "TRAC: truthful auction for location-aware collaborative sensing in mobile crowdsourcing," in *Proceedings of the 33rd IEEE Conference on Computer Communications (IEEE INFOCOM '14)*, pp. 1231–1239, May 2014.

[14] D. Wu, J. Yan, H. Wang, D. Wu, and R. Wang, "Social Attribute Aware Incentive Mechanism for Device-to-Device Video Distribution," *IEEE Transactions on Multimedia*, vol. 19, no. 8, pp. 1908–1920, 2017.

[15] T. Luo, S. S. Kanhere, and H. Tan, "SEW-ing a Simple Endorsement Web to incentivize trustworthy participatory sensing," in *Proceedings of the Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON '14)*, pp. 636–644, Singapore, Singapore, June 2014.

[16] Y. Ueyama, M. Tamai, Y. Arakawa, and K. Yasumoto, "Gamification-based incentive mechanism for participatory sensing," in *Proceedings of the IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS '14)*, pp. 98–103, Budapest, Hungary, March 2014.

[17] D. Wu, Q. Liu, H. Wang, D. Wu, and R. Wang, "Socially aware energy-efficient mobile edge collaboration for video distribution," *IEEE Transactions on Multimedia*, vol. 19, no. 10, pp. 2197–2209, 2017.

[18] T. Yu, Z. Zhou, D. Zhang, X. Wang, Y. Liu, and S. Lu, "INDAPSON: an incentive data plan sharing system based on self-organizing network," in *Proceedings of the IEEE INFOCOM, IEEE Conference on Computer Communications*, pp. 1545–1553, Toronto, ON, Canada, April 2014.

[19] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining 'gamification'," in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek '11)*, pp. 9–15, ACM, Tampere, Finland, September 2011.

[20] Y. Wang, X. Jia, Q. Jin, and J. Ma, "QuaCentive: a quality-aware incentive mechanism in mobile crowdsourced sensing (MCS)," *The Journal of Supercomputing*, vol. 72, no. 8, pp. 2924–2941, 2016.

[21] R. Ma, J. Xiong, M. Lin, Z. Yao, H. Lin, and A. Ye, "Privacy protection-oriented mobile crowdsensing analysis based on game theory," in *Proceedings of the IEEE Trustcom/BigDataSE/ICESS*, pp. 990–995, Sydney, Australia, August 2017.

[22] X. Liu, R. Deng, K.-K. R. Choo, Y. Yang, and H. Pang, "Privacy-preserving outsourced calculation toolkit in the cloud," *IEEE Transactions on Dependable and Secure Computing*, 2018.

[23] H. Wang, D. He, Y. Sun, N. Kumar, and K.-K. R. Choo, "PAT: a precise reward scheme achieving anonymity and traceability for crowdcomputing in public clouds," *Future Generation Computer Systems*, vol. 79, pp. 262–270, 2018.

[24] Y. Zhang, X. Chen, J. Li, D. S. Wong, H. Li, and I. You, "Ensuring attribute privacy protection and fast decryption for outsourced data security in mobile cloud computing," *Information Sciences*, vol. 379, pp. 42–61, 2017.

[25] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proceedings of the SIGMOD International Conference on Management of Data (SIGMOD '10)*, pp. 735–746, ACM, Indianapolis, IN, USA, June 2010.

[26] E. Rieffel, J. Biehl, W. van Melle et al., "Secured histories: computing group statistics on encrypted data while preserving individual privacy," 2010.

[27] J. Ni, K. Zhang, Y. Yu, X. Lin, and X. S. Shen, "Providing task allocation and secure deduplication for mobile crowdsensing via fog computing," *IEEE Transactions on Dependable and Secure Computing*, 2018.

[28] Y. Yang, X. Liu, R. H. Deng, and Y. Li, "Lightweight sharable and traceable secure mobile health system," *IEEE Transactions on Dependable and Secure Computing*, 2017.

[29] J. Xiong, Y. Zhang, X. Li et al., "Rse-pow: a role symmetric encryption pow scheme with authorized deduplication for multimedia data," *Mobile Networks and Applications*, vol. 23, no. 3, pp. 650–663, 2018.

[30] L. Chen, R. Lu, and Z. Cao, "PDAFT: a privacy-preserving data aggregation scheme with fault tolerance for smart grid communications," *Peer-to-Peer Networking and Applications*, vol. 8, no. 6, pp. 1122–1132, 2014.

[31] M. A. Alsheikh, Y. Jiao, D. Niyato, P. Wang, D. Leong, and Z. Han, "The accuracy-privacy trade-off of mobile crowdsensing," *IEEE Communications Magazine*, vol. 55, no. 6, pp. 132–139, 2017.

[32] E. Shi, T. H. H. Chan, and E. Rieffel, "Privacy-preserving aggregation of time-series data," in *Proceedings of the Annual Network And Distributed System Security Symposium*, 2011.

[33] S. Shen, G. Yue, Q. Cao, and F. Yu, "A survey of game theory in wireless sensor networks security," *Journal of Networks*, vol. 6, no. 3, pp. 521–532, 2011.