InPPTD: A Lightweight Incentive-Based Privacy-Preserving Truth Discovery for Crowdsensing Systems

Kaiping Xue, *Senior Member, IEEE*, Bin Zhu, *Graduate Student Member, IEEE*, Qingyou Yang, *Member, IEEE*, Na Gai, *Student Member, IEEE*, David S. L. Wei, *Senior Member, IEEE*, and Nenghai Yu, *Member, IEEE*

***Abstract*—Recently, truth discovery in crowdsensing systems has received considerable attention with its appealing features for extracting truthful information from multiple unreliable data sources. However, it also poses new challenges to the issues of privacy and security. On the one hand, workers' sensed data can be used to infer their privacy. On the other hand, workers may be selfish and lazy, especially in the Internet-of-Things environment, devices are usually resource constrained, so they may dishonestly execute the costly sensing task so as to reduce resource consumption, or even break the protocol to obtain illegal rewards. Although some privacy-preserving truth discovery schemes have been proposed, they still cannot achieve strong privacy protection while keeping efficiency on the worker side, and still has no efficient incentive mechanism to persuade workers to participate in the system operations. In this article, we propose an incentive-based privacy-preserving truth discovery framework, named InPPTD. By adopting the Paillier homomorphic cryptosystem and two noncolluding servers, InPPTD not only effectively protects workers' sensed data information but also preserves the privacy of these workers' weight information. Meanwhile, a weight-based incentive mechanism is introduced in InPPTD to reduce the number of lazy workers. Security and performance analysis shows that InPPTD can guarantee stronger security features, while also ensure efficiency in terms of computation and communication overhead.**

***Index Terms*—Crowdsensing, incentive, privacy-preserving, truth discovery.**

I. Introduction

With the rapid development of portable and mobile devices (e.g., smartphone, smartwatch, smartglass, etc.) equipped with a series of sensors (such as accelerometer, GPS, camera, and compass), a new sensing paradigm that collects and analyzes sensed data from a crowd of diverse participating users (referred to as workers) has emerged. However, workers' sensed data are usually unreliable due to various factors, including background noise, hardware quality, insufficient skill, as well as lack of effort. To discover truthful information from these unreliable data, truth discovery has recently been widely studied [1]–[3] and applied in many areas [4]–[7]. Rather than treating all participating workers equally, truth discovery algorithms differentiate them by estimating the worker reliability (also referred to as weight) from reported sensed data, while updating truthful facts (referred to as the ground truth) by quality-aware data aggregation [8], [9].

Although truth discovery can significantly improve the accuracy of aggregation results, this process also poses privacy concerns. On the one hand, worker's sensed data may be used to infer private information and utilized by adversaries for criminal purposes. If the system does not equip with the privacy preservation of workers' sensed data, workers may not be willing to provide their data to the system. On the other hand, the worker's reliability (i.e., weight) is also sensitive information that should be well protected. For example, by aggregating opinions regarding challenging social problems may lead to a better solution, but the leakage of weight may disclose worker's education and intellectual level [10], [11]. Additionally, workers may be lazy and selfish as they usually worry about the overuse of the resources of their own devices, especially in the Internet-of-Things environment, devices are often resource-constrained. In order to obtain higher benefits, they may reduce their costly sensing effort, such as spent time, resources, attention, and carefulness in the sensing tasks. Obviously, these misbehaviors will significantly impair the aggregation accuracy [12]. More seriously, workers may maliciously manipulate the weight, which is often used as the reference for rewards [13], [14], and get additional illegal rewards.

In order to address the privacy issues, a few privacy-preserving truth discovery schemes (PPTDs) have been proposed. In general, they can be divided into two categories: 1) single-server scheme and 2) two-server scheme. Single-server schemes, such as [15] and [16], can be easily

Manuscript received February 10, 2020; revised June 6, 2020 and September 3, 2020; accepted September 19, 2020. Date of publication October 7, 2020; date of current version March 5, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61972371, and in part by the Youth Innovation Promotion Association of the Chinese Academy of Sciences under Grant 2016394. *(Corresponding author: Kaiping Xue.)*

Kaiping Xue and Nenghai Yu are with the School of Cyber Security and the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China (e-mail: kpxue@ustc.edu.cn).

Bin Zhu and Na Gai are with the School of Cyber Security, University of Science and Technology of China, Hefei 230027, China.

Qingyou Yang is with the Department of Security, Baidu Inc., Shenzhen 518000, China.

David S. L. Wei is with the Computer and Information Science Department, Fordham University, Bronx, NY 10458 USA.

Digital Object Identifier 10.1109/JIOT.2020.3029294

implemented and can achieve the privacy preservation for worker's sensed data and weight. But each worker has to perform some costly operation for encrypting data and also has to keep online until the algorithm of iteration ends. Thus, any workers' failures would impair the accuracy of the final aggregation result. While two-server schemes are proposed aiming at freeing the workers from the big burden of computation. Among them, Zheng *et al.* [15] proposed a lightweight two-server scheme, in which each worker's data are perturbed by a random number, which is more efficient than performing complex encryption. However, the worker directly participates in the calculation of its weight, which may cause the weight to be modified by a malicious worker, and the problem of workers' failures still exists in their scheme. In [17], workers are offline after they report their sensed data, and the iteration is only executed between the two clouds. However, this scheme cannot guarantee strong privacy protection since each worker's weight information is disclosed to one of the two clouds.

The existing PPTD schemes, whether with a single server or two servers, have not yet considered the issue of incentives, which is also an important factor in determining whether workers would participate in the truth discovery program. Although the incentive mechanisms have been studied [14], [18], which use the auction model and game-theoretic model to encourage workers to honestly carry out the sensing tasks, the privacy issues of the incentive mechanism have not yet been further discussed. Especially, when it is incorporated into the truth discovery, how to simultaneously ensure privacy in truth discovery and incentive mechanism remains a research challenge.

Observing that the above issues have not yet been addressed adequately, in this article, we design an incentive-based PPTD (InPPTD) for crowdsensing systems. Our proposed scheme integrates the truth discovery with an incentive mechanism which offers payments to incentivize high-effort sensing task from workers. It can achieve strong privacy protection by leveraging a high-security level homomorphic encryption algorithm, while it is extremely efficient on the worker side. Moreover, the well-designed protocol is able to avoid workers' selfish and malicious behaviors for gaining the benefits through illegal means. Specifically, this article makes the following contributions.

1) We propose an efficient and privacy-enhanced truth discovery scheme for crowdsensing systems. Both the workers' sensed data and weight information can be well protected, while each worker needs less computation and can be offline after reporting his perturbed data.
2) A weight-aware incentive mechanism is integrated into the truth discovery. The amount of earned rewards for each worker is determined based on his contribution to the final truth. Therefore, lazy workers can be easily identified and removed from the system.
3) Based on the two noncolluding servers, the proposed incentive mechanism can not only preserve privacy for each worker but also prevent malicious workers from manipulating the interactions to get additional rewards illegally.

The remainder of this article is organized as follows. We discuss the related work in Section II. We state our system model, security assumptions, and design goals in Section III. We give the brief preliminaries of truth discovery, incentive mechanism, and the Paillier cryptosystem in Section IV and present the detailed incentive-based PPTD in Section V. Security analysis is shown in Section VI and performance evaluation is presented in Section VII. We conclude this work in Section VIII.

II. Related Works

Recently, truth discovery is recognized as an effective method to extract truthful information from multiple data sources and has attracted more and more attention from researchers and practitioners. Compared to the traditional averaging or voting approaches, recent truth discovery schemes, such as conflict resolution on heterogeneous (CRH) [19], optimized CRH [1], and TruthFinder [20], can provide more reliable results by considering device reliability. However, these schemes execute the truth discovery process in the plaintext domain and do not consider participants' privacy.

Some PPTD schemes have been proposed with an increase in privacy concerns. In general, they can be classified into two categories: 1) single-server scheme and 2) two-server scheme. The single-server scheme considers the framework with only one server (also called cloud server), which usually requires workers (participants) to perform some complex computations and participate in every iteration of truth discovery. For example, Miao *et al.* [16] adopted the threshold Paillier cryptosystem to design a single-server PPTD system, in which both the sensed data and weight are privacy preserved. However, each worker needs to encrypt its data with the Paillier cryptosystem, which is challenging for resource constrained devices. To reduce the computation cost, Zheng *et al.* [15] and Xu *et al.* [21] utilized a lightweight homomorphic encryption technology, respectively, which can achieve the requirements of privacy protection of sensed data and weight. However, these two single-server schemes cannot be successfully executed without the real-time participation of each worker, which means the problem of workers' failures would be an obstacle for the deployment of single-server schemes. Moreover, workers' weights are calculated by themselves, which makes the schemes vulnerable to modification attacks from malicious workers. In addition, Li *et al.* [22], [23] utilized local differential privacy technology and proposed two schemes, which use a two-layer randomized response mechanism and a Gaussian noise mechanism, respectively. PPTD methods based on local differential privacy are generally more efficient and strike a balance between accuracy and the level of privacy protection.

The two-server truth discovery scheme utilizing two noncolluding servers is more appropriate for the PPTD scenario. Through the division of labor between the two servers, the two-server schemes can protect workers' privacy, while not requiring workers to participate in iterations of truth discovery. Therefore, workers do not need to perform large burden computing, and their failures also do not have any impact on the system running. However, some serious drawbacks have not been tackled in existing works. Zheng *et al.* [15] introduced

noise to perturb the plaintext data so that workers' sensed data and weights can be protected. The drawback of this scheme is that each worker still has to participate in each round iteration of the truth discovery procedure. Miao *et al.* [17] proposed two lightweight PPTD protocols, named L-PPTD and L^2-PPTD, respectively. L-PPTD can protect each worker's sensed data and weight, but workers are still involved in the calculation of their own weight for each iteration. In L^2-PPTD, workers can be offline after they upload their data to the clouds. However, workers' privacy of weights is disclosed to a cloud. Then, Tang *et al.* [24] used the garbled circuit (GC) technology to design a noninteractive PPTD (named NPPTD) protocol on the basis of two noncolluding servers. This design removes the online requirement for workers and guarantees workers' privacy of sensed data and weights. But owing to the inherent disadvantages of GC, NPPTD has a huge computation and communication overhead of GC generation and transmission. Meanwhile, Zheng *et al.* [11] constructed a new system architecture that can protect both of the sensed data and the corresponding weight of each worker, but it cannot prevent the modification of malicious workers and reduce lazy workers.

In fact, it is quite challenging to design a PPTD scheme that can meet all of the following requirements: high efficiency, workers' failure resistance, data, and weight privacy preservation, resistance to the modification attack from malicious workers, and reduction of lazy workers. As far as we know, there is no existing PPTD scheme possessing all of the above important properties at the same time. In addition, the existing PPTD schemes have not yet considered the incentives for workers. Although some works have conducted such research that constructs an incentive mechanism for crowdsensing system [14], [18], [25], it is still a challenging work to combine the incentive mechanism with PPTD protocol while preserving worker's privacy. In this article, we thus propose a novel PPTD scheme, named InPPTD, which leverages a two-server framework and introduces an incentive mechanism. Our scheme does not need workers to be online during the entire process of truth discovery and thus can achieve failure resistance and strong privacy guarantee while providing fairness of rewards according to the workers' contributions.

III. System Model, Security Assumptions, and Design Goals

A. System Model

We consider a two-server crowdsensing architecture that consists of a service provider (SP), a cloud provider (CP), and a group of participating workers, as shown in Fig. 1. The SP is a cloud-based platform that posts sensing tasks that usually require the sensed data on a collection of objects (e.g., litter, pothole, automated external defibrillator objects, etc.) from participating workers, and rewards each participating worker according to the worker's contributions. The CP is responsible for performing some secure computation and assists the SP to execute the secure weight estimation and truth estimation during the truth discovery processes. The participating workers are a set of mobile device users who carry out sensing tasks with their mobile devices for financial incentives.

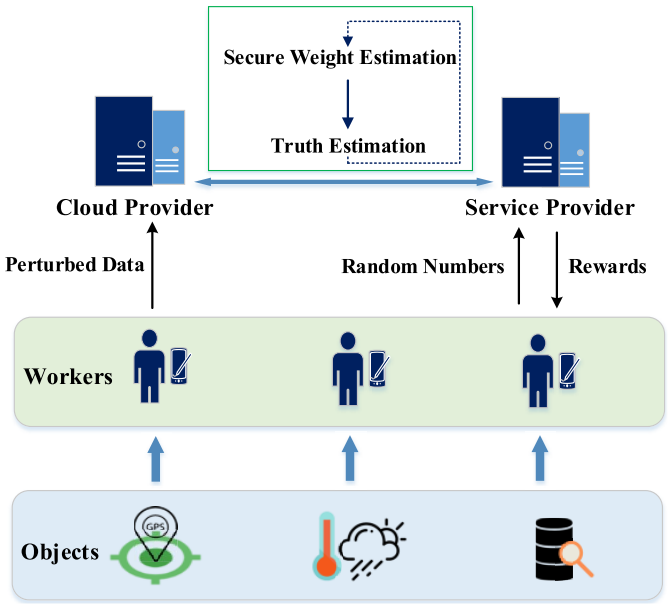


Fig. 1. System model.

In this article, we assume that there are M objects (denoted $O = \{o_1, o_2, \ldots, o_M\}$) in the sensing task, and K participating workers (referred to as $U = \{u_1, u_2, \ldots, u_K\}$), and the weights for workers are presented as $W = \{w_1, w_2, \ldots, w_K\}$. The problem we aim to address is to facilitate the SP to accurately estimate the ground truths $\{t_m\}_{m=1}^{M}$ for objects in a secure and efficient way with the assistance of CP. To make our system model clearer, we give an example of a practical crowdsensing application, i.e., indoor floorplan [26]. In such a crowdsensing system, SP poses the sensing task for measuring the distances between two specified points and constructs the indoor floorplan based on the collected sensed data from participating workers. For the sake of privacy and efficiency for workers, CP is introduced into our truth discovery framework such that the iterations of weight and truth estimation are only performed between CP and SP.

B. Security Assumptions

In truth discovery, the sensed data and weight information of each worker may be disclosed to CP, SP, and other parties. These parties may try to infer the sensitive and private information of each worker for the purpose of their own benefits or crime. On the other hand, workers may be lazy and malicious. Lazy workers may strategically reduce the participation of costly sensing tasks. Meanwhile, malicious workers may try to tamper and forgery the transmitted data to illegally get some financial benefits.

In this article, we follow the same assumptions that the two servers, SP and CP, are semihonest and nonconcluding as defined in [17] and [24]. We consider that the two servers are honest-but-curious. They will honestly follow the protocol, while they hold all the data they have sent and received, and may attempt to infer the private information of other parties, such as individual worker's sensed data and weight information. We also assume that there is no collusion between these two servers in the system, which means they will not

collude with each other outside the protocol. Additionally, we assume that there exist secure channels among workers, SP, and CP. These secure channels can be constructed by the TLS/SSL protocol.

C. Design Goals

In this article, we intend to devise an PPTD scheme, which can provide more accurate ground truth. Specifically, our scheme achieves the following design goals.

1) *Efficiency:* The proposed scheme must achieve high efficiency in terms of computation and communication, especially for workers, since the workers' devices are usually resource-constrained and thus are not given the ability of complex operations.
2) *Strong Privacy Preservation:* Individual worker's sensed data and weight information should be well protected from disclosure to any other parties during the truth discovery and incentive processes.
3) *Failure Resistance:* The proposed scheme should have the capability of failure resistance, so workers' failures will not influence the operation of the system and the accuracy of the final results.
4) *Resistance to the Modification Attack From Malicious Workers:* The proposed scheme should be able to resist the modification attack so that malicious workers cannot gain benefits through their misbehavior.
5) *Reduction of Lazy Workers:* Considering the existence of lazy workers, the proposed scheme should be able to distinguish normal workers from lazy workers, e.g., introducing an incentive mechanism to reward honest and diligent workers according to their contributions (i.e., weight), thereby making lazy workers unprofitable so as to reduce lazy workers.

IV. Preliminaries

A. Truth Discovery

With the explosion of information, data for one object can be collected from multiple sources. However, there usually exist conflicts among multisource noisy information. To extract truthful information from these data, traditional methods usually treat all information sources equally and derive the final result by averaging/voting method. While truth discovery approaches resolve the conflicts and infer truthful information based on the different reliability degrees between different sources. Recently, some truth discovery algorithms have been proposed. Though the algorithmic details are a bit different from each other, the fundamental principle of assigning device weights and estimating ground truth is more or less the same. The CRH data framework [19] has been widely used in many PPTDs as the truth discovery algorithm. Therefore, without loss of generality, we follow the CRH framework to briefly introduce the procedure in truth discovery.

The truth discovery algorithm starts with a random guess of ground truths and then iteratively carries out the weight estimation and truth estimation phase until it satisfies the convergence criterion. The convergence criterion can be a predefined iteration number or a threshold of the change between two estimated truths in consecutive iterations.

1) Weight Estimation: In this phase, weight for each worker (e.g., worker k) is computed. Given the estimated ground truth t_m for object m, the weight w_k is derived as

$$w_k = \log\left(\sum_{k=1}^{K}\sum_{m=1}^{M} d\big(x_{m,k}, t_m\big) \Big/ \sum_{m=1}^{M} d\big(x_{m,k}, t_m\big)\right) \tag{1}$$

in which the distance function is adopted by the squared distance $d(x_{m,k}, t_m) = (x_{m,k} - t_m)^2$. In this article, we focus on continuous sensed data type, but it can be easily extended to support categorical data in the same way as [17].

2) Truth Estimation: After the weight of each worker has been estimated, the ground truth for each object o_m can be computed as follows:

$$t_m = \sum_{k=1}^{K}\big(w_k x_{m,k}\big) \Big/ \sum_{k=1}^{K} w_k \tag{2}$$

where t_m represents the estimated ground-truth value.

B. Incentive Mechanism

Recently, some incentive mechanisms for truth discovery have been studied to avoid lazy workers. The worker's effort (weight) in these mechanisms is an important reference for rewards. Generally, the number of rewards for a worker can be formulated as

$$p_k = f\left(w_k \Big/ \sum_{k=1}^{K} w_k\right) \tag{3}$$

where w_k is worker k's weight representing his/her effort in the sensing task, p_k is rewards for worker k, and function $f(\cdot)$ is a defined incentive algorithm. Without loss of generality, we utilize a simple incentive algorithm that is

$$p_k = P \cdot \left(w_k \Big/ \sum_{k=1}^{K} w_k\right) \tag{4}$$

where P is the total rewards for this sensing task. Other complex incentive algorithms that may be implemented by the game theory or optimization theory can be extended from our scheme, and will be discussed in the future.

C. Paillier Cryptosystem

The Paillier cryptosystem [27] is one of the most popular public-key encryption schemes, which can achieve the homomorphism properties. Specifically, the Paillier cryptosystem includes the following three algorithms.

1) *Key Generation:* Select two large and independent prime numbers p and q randomly, and compute $\lambda = lcm(p-1, q-1)$ and $N = pq$, where λ is the least common multiple of $p-1$ and $q-1$. Then, define a function $L(x) = [(x-1)/N]$, choose a generator $g = (1+N)$, and compute $\mu = (L(g^{\lambda} \bmod N^2))^{-1} \bmod N$. The public key is (N, g), and the private key is (λ, μ).
2) *Encryption:* Given the message $m \in \mathbb{Z}_N^*$, we select a random number $r \in \mathbb{Z}_N^*$. Then, we can compute the ciphertext as

$$C = E(m) = g^m \cdot r^N \bmod N^2. \tag{5}$$

3) *Decryption:* To decrypt a ciphertext c, where $C \in \mathbb{Z}^*_{N^2}$ using the private key (λ, μ), we can compute the plaintext message as

$$m = D(C) = L\left(C^{\lambda} \bmod N^2\right) \cdot \mu \bmod N. \quad (6)$$

The Paillier cryptosystem has an additive homomorphism property, that is, given two ciphertexts C_1 and C_2 for messages m_1 and m_2, respectively, we can obtain the ciphertexts of $m_1 + m_2$ by performing the product $C_1 C_2$.

V. Our Proposed InPPTD

A. Overview

As described in Section IV, truth discovery is an iteration between weight estimation and truth estimation. To protect workers' privacy and make the protocol efficient, two noncolluding servers SP and CP are introduced. As shown in Fig. 1, there are mainly three phases in our proposed InPPTD: *Report Phase, Iteration Phase, and Rewards Phase*.

In the report phase, each worker reports his perturbed sensed data to CP, while the random number is uploaded to SP. After that, workers can stay offline until the end of truth discovery. Then, the iteration phase is carried out, which contains a secure weight estimation step and a truth estimation step. In the secure weight estimation step, CP computes the secure distance function, and sends the ciphertexts of aggregated distance $\prod_{k=1}^{K} C_k$ and perturbed distances $C_k^{2a_k}$ to SP. Then, SP decrypts them and computes each worker's perturbed weight $w_k - a_k$. In the truth estimation step, SP sends encrypted perturbed weights $E(w_k - a_k)$ and $E(\sum_{k=1}^{K}(w_k - a_k) \cdot r_{m,k})$ to CP. Then, CP computes $E(\sum_{k=1}^{K} w_k x_{m,k})$ and $E(\sum_{k=1}^{K} w_k)$, and sends them to SP. SP decrypts them and computes the estimated ground truth as $t_m = \sum_{k=1}^{K} w_k x_{m,k} / \sum_{k=1}^{K} w_k$. These two steps are iteratively executed until the convergence condition is satisfied, and output the final estimated truth. Every time a truth discovery task is completed, each participating worker obtains corresponding rewards based on his contributions (i.e., weight) to the final truth. In the rewards phase, each worker can redeem his rewards from SP without violating his privacy.

B. Initialization Phase

In this phase, SP generates the necessary public and private parameters for the system setup. Assume that there are a totally K participating workers and M objects are required to be sensed for each worker. SP initializes the system as follows.

1) SP first chooses two large and independent random primes p and q, computes $N = pq$, and chooses a generator $g = (1 + N)$. The public key is $PK = (N, g)$.
2) Then, SP computes $\lambda = lcm(p - 1, q - 1)$ and $\mu = (L(g^{\lambda} \bmod N^2))^{-1} \bmod N$, where $L(x) = [(x - 1)/N]$. The private key is $SK = (\lambda, \mu)$.
3) Finally, SP publishes the public key PK to CP to finish this phase.

C. Report Phase

In this phase, each worker reports its sensed data to CP. For the sake of privacy preservation, data will be obfuscated by random noises before being reported to CP. The procedure of this phase is shown in Fig. 2.

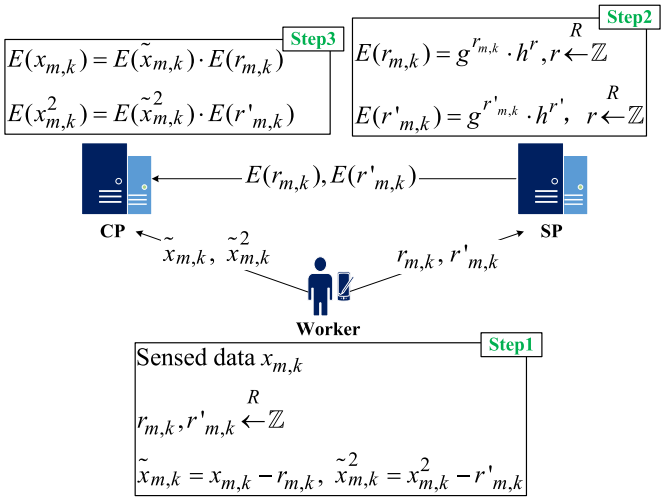


Fig. 2. Report phase of InPPTD.

Step 1: To perturb worker k's sensed data $x_{m,k}$ for object m, the worker first chooses two random numbers $r_{m,k}$ and $r'_{m,k}$, then computes the perturbed data as

$$\tilde{x}_{m,k} = x_{m,k} - r_{m,k}$$
$$\tilde{x}^2_{m,k} = x^2_{m,k} - r'_{m,k}. \quad (7)$$

Finally, worker k reports all perturbed data and random chosen numbers to CP and SP through a secure channel, respectively.

Step 2: When SP receives the reported random number from workers, it first encrypts them by utilizing the Paillier cryptosystem, then sends the ciphertexts to CP through a secure channel. The encryption process for $r_{m,k}$ and $r'_{m,k}$ is given as follows:

$$E(r_{m,k}) = g^{r_{m,k}} \cdot h^r$$
$$E(r'_{m,k}) = g^{r'_{m,k}} \cdot h^{r'} \quad (8)$$

where r and r' are random numbers in $\mathbb{Z}_N$ chosen by SP.

Step 3: After receiving perturbed data (e.g., $\tilde{x}_{m,k}$ and $\tilde{x}^2_{m,k}$) and encrypted random numbers from workers and SP, CP first encrypts the perturbed data $\tilde{x}_{m,k}$ and $\tilde{x}^2_{m,k}$ by SP's public key to get $E(\tilde{x}_{m,k})$ and $E(\tilde{x}^2_{m,k})$, then performs the following computation to obtain the ciphertexts of $x_{m,k}$ and $x^2_{m,k}$:

$$E(x_{m,k}) = E(\tilde{x}_{m,k}) \cdot E(r_{m,k})$$
$$E(x^2_{m,k}) = E(\tilde{x}^2_{m,k}) \cdot E(r'_{m,k}). \quad (9)$$

It should be noted that the data which are encrypted by the Paillier cryptosystem may not be integers. We use a public parameter R to round the factional values. For example, the rounded integer of a can be computed as $\tilde{a} = \lfloor Ra \rfloor$, and the value of a can be recovered by computing $\tilde{a}/R$.

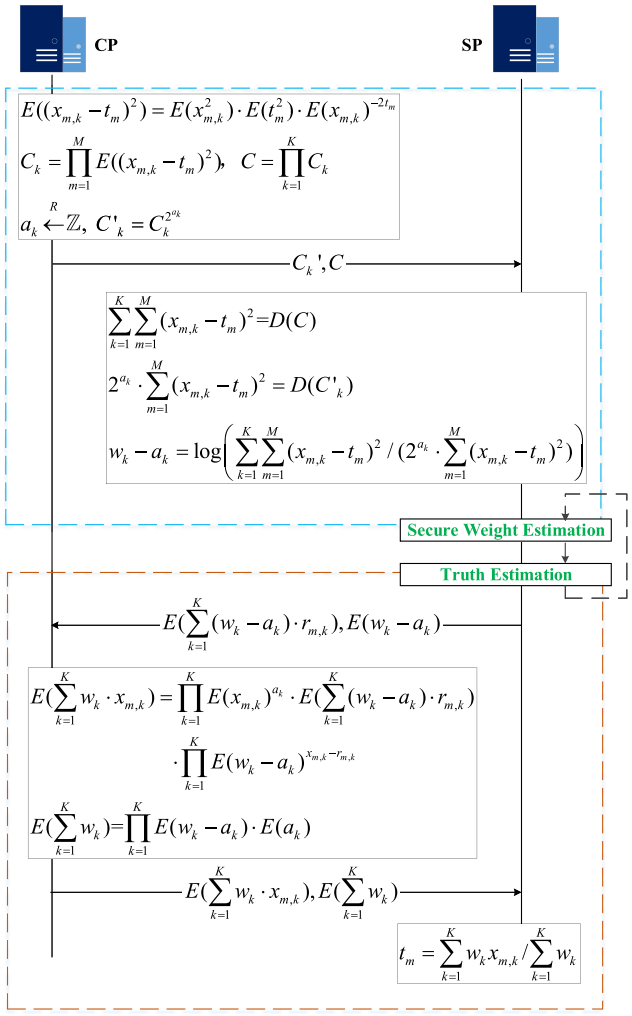


Fig. 3. Iteration phase of InPPTD.

D. Iteration Phase

As shown in Fig. 3, this phase iteratively executes the secure weight estimation step and the truth estimation step until it satisfies the convergence criterion.

1) Secure Weight Estimation: In this step, the weight of each worker is estimated based on the published ground truth and worker's sensed data. It should be noted that the ground truth for the first time weight estimation is randomly chosen. First, CP computes the secure distance function $E((x_{m,k}-t_m)^2)$ as

$$E\left((x_{m,k}-t_m)^2\right)=E\left(x_{m,k}^2\right)\cdot E\left(t_m^2\right)\cdot E\left(x_{m,k}\right)^{-2t_m}. \quad (10)$$

Then, CP aggregates the distances of objects and workers as

$$C_k=\prod_{m=1}^{M}E\left((x_{m,k}-t_m)^2\right) \quad (11)$$

$$C=\prod_{k=1}^{K}C_k. \quad (12)$$

Finally, CP randomly chooses a random number a_k for each worker k, computes $C'_k=(C_k)^{2^{a_k}}$, and sends C'_k and C to SP.

In fact, C'_k is the encryption form of $2^{a_k}\cdot\sum_{m=1}^{M}(x_{m,k}-t_m)^2$, C is the encryption form of $\sum_{k=1}^{K}\sum_{m=1}^{M}(x_{m,k}-t_m)^2$.

After receiving C'_k and C from CP, SP decrypts them to obtain $2^{a_k}\cdot\sum_{m=1}^{M}(x_{m,k}-t_m)^2$ and $\sum_{k=1}^{K}\sum_{m=1}^{M}(x_{m,k}-t_m)^2$, respectively. Then, performs following formula for each worker to derive its perturbed weight:

$$w_k-a_k=\log\times\left(\sum_{k=1}^{K}\sum_{m=1}^{M}\left(x_{m,k}-t_m\right)^2/\left(2^{a_k}\sum_{m=1}^{M}\left(x_{m,k}-t_m\right)^2\right)\right). \quad (13)$$

2) Truth Estimation: In the weight estimation step, SP can obtain a perturbed weight for each worker. To estimate the ground truth, SP first computes $E(\sum_{k=1}^{K}(w_k-a_k)r_{m,k})$ and encrypts the perturbed weight [e.g., $E(w_k-a_k)$], and sends them to CP. Then, the CP, respectively, performs (14) and (15), and sends $E(\sum_{k=1}^{K}w_k x_{m,k})$ and $E(\sum_{k=1}^{K}w_k)$ back to SP

$$E\left(\sum_{k=1}^{K}w_k\right)=\prod_{k=1}^{K}E(w_k-a_k)\cdot E(a_k) \quad (14)$$

$$\begin{aligned}E\left(\sum_{k=1}^{K}w_k x_{m,k}\right)&=E\left(\sum_{k=1}^{K}a_k x_{m,k}+\sum_{k=1}^{K}(w_k-a_k)x_{m,k}\right)\\&=E\left(\sum_{k=1}^{K}a_k x_{m,k}\right)\cdot E\left(\sum_{k=1}^{K}(w_k-a_k)r_{m,k}\right)\\&\quad\times E\left(\sum_{k=1}^{K}(w_k-a_k)\left(x_{m,k}-r_{m,k}\right)\right)\\&=\prod_{k=1}^{K}E\left(x_{m,k}\right)^{a_k}\cdot E\left(\sum_{k=1}^{K}(w_k-a_k)r_{m,k}\right)\\&\quad\times\prod_{k=1}^{K}E(w_k-a_k)^{x_{m,k}-r_{m,k}}.\end{aligned} \quad (15)$$

Upon receiving $E(\sum_{k=1}^{K}w_k x_{m,k})$ and $E(\sum_{k=1}^{K}w_k)$ from CP, SP decrypts them and computes the estimated truth as

$$t_m=\sum_{k=1}^{K}w_k x_{m,k}/\sum_{k=1}^{K}w_k. \quad (16)$$

Subsequently, SP determines whether the convergence condition is satisfied, and if not, SP transmits the ground truth to the CP to continue the next iteration.

E. Reward Phase

At the end of the iteration phase, SP gets the final perturbed weight for each worker, while CP holds the corresponding randomly chosen number. For each task i, the following steps will be executed.

1) First, CP aggregates all random numbers as $S^i=\sum_{k=1}^{K}a_k$ and transmits it to SP.
2) Then, SP computes the total weight $W^i=\sum_{k=1}^{K}(w_k-a_k)+S^i$, and $ps_k^i=[(w_k-a_k)/W^i]\cdot P^i$, where P^i is the total reward for task i. Subsequently, SP sends the total weight W^i to CP.

3) When receiving the total weight from SP, CP computes $pc_k^i = (a_k/W^i) \cdot P^i$.

It should be noted that the real reward of task i for worker k is $ps_k^i + pc_k^i$, and workers can learn the real-time rewards for each task by querying the CP and SP. However, we do not recommend workers redeem their rewards in real time. On the one hand, workers' weights can be easily inferred by SP when redeeming the rewards after complete the task intermediately. On the other hand, the amount of rewards for each task is generally minor in reality, and frequent redemption will bring excessive transaction costs. Therefore, we adopt an aggregation incentive mechanism. In detail, SP and CP will store ps_k^l and pc_k^l for multiple tasks (e.g., $l = i, \ldots, j$). Upon worker k wants to redeem its rewards, CP first sends aggregated result $\sum_{l=i}^{j} pc_k^l$ to SP. Then, SP computes the total rewards during the period from task i to j as follows:

$$p_k = \sum_{l=i}^{j} ps_k^l + \sum_{l=i}^{j} pc_k^l. \tag{17}$$

Finally, SP transfers the corresponding incentives to worker k's account. For clarity, all the above steps are depicted in Fig. 4.

It should be noted that SP is able to recognize the lazy workers from the total rewards. In this weight-aware incentive mechanism, lazy workers are usually assigned fewer weights, and gain fewer rewards. In fact, our experiments in Section VII-E have shown that the gap in rewards between normal workers and lazy workers is obvious. Therefore, SP can remove lazy workers from the system through the incentive mechanism.

F. Generalization

In our proposed scheme, the truth estimation step and the incentive mechanism are both general to be applied to other truth discovery algorithms, but the weight estimation step cannot be directly applied to other weight functions. This is because in our proposed scheme, no server is able to obtain the plaintext value of each worker's distance, which is perturbed by a random number, and thus no server can estimate each worker's weight locally.

If the weight estimation function is changed, the secure weight estimation step should be modified accordingly. For example, in the weight estimation step, if we use the affine function $w_k = 1 - p \sum_{m=1}^{M} d(x_{m,k,t_m})$ as the weight estimation function, CP computes $E(p \sum_{m=1}^{M} d(x_{m,k,t_m}) + a_k) = C_k^p \cdot E(a_k)$ and sends it to SP, where p is a parameter chosen based on the specific application scenarios, $d(x_{m,k,t_m}) = (x_{m,k} - t_m)^2$, $C_k = \prod_{m=1}^{M} E(d(x_{m,k,t_m}))$, and a_k is randomly chosen by CP. Then, SP decrypts $E(p \sum_{m=1}^{M} d(x_{m,k,t_m}) + a_k)$ and computes $w_k - a_k = 1 - (p \sum_{m=1}^{M} d(x_{m,k,t_m}) + a_k)$. Thus, the truth estimation step and the reward phase can be continued without any modification.

VI. Security Analysis

In this section, we first prove that each worker's privacy of sensed data and weight information will not be disclosed to CP, SP, and other workers. Then, we demonstrate that our scheme can provide failure resistance, and reduce malicious and lazy workers. Finally, the security comparisons of our work with the related works are presented.

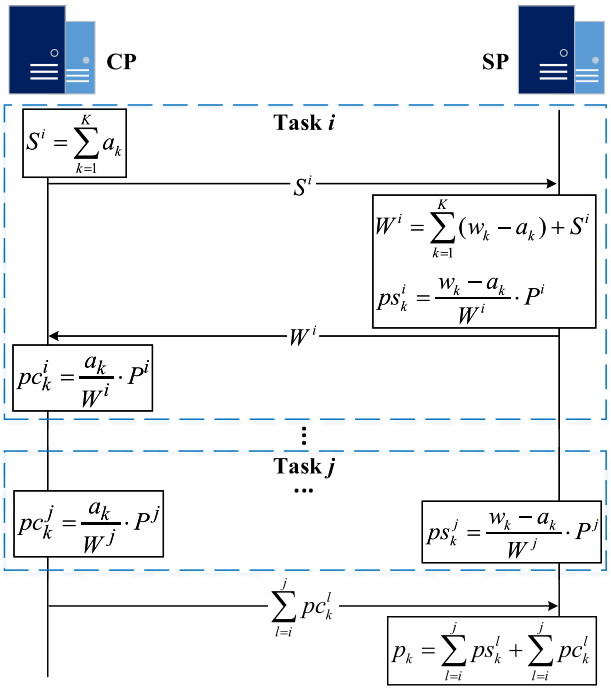


Fig. 4. Reward phase of InPPTD.

A. Preventing Privacy Disclosure to CP

During the InPPTD procedure in Section V, CP can obtain plaintexts of data $\{x_{m,k}+r_{m,k}\}_{m,k=1}^{M,K}$ and $\{x_{m,k}^2+r'_{m,k}\}_{m,k=1}^{M,K}$, total weight W^i, and some ciphertexts. Obviously, without knowing random numbers $\{r\}_{m,k=1}^{M,K}$ and $\{r'\}_{m,k=1}^{M,K}$, CP cannot infer individual worker's sensed data and weight from these plaintexts. On the other hand, the ciphertexts are encrypted by the Paillier cryptosystem which is semantically secure against the chosen plaintext attack. CP cannot decrypt any ciphertext without knowing the private key SK. Therefore, our scheme achieves the protection of privacy disclosure to CP.

B. Preventing Privacy Disclosure to SP

For SP, it holds the private key for the Paillier cryptosystem, so it can decrypt all the ciphertexts it received. However, the following theorem can prove that SP cannot obtain an individual worker's sensed data and weight information.

Theorem 1: Suppose that the number of participating workers' tasks satisfies $K \geq 2$ and $|j - i| \geq 2$, and for each object, there are at least two workers providing different sensed data. If the two servers are semihonest and noncolluding, individual worker's sensed data and weight information will not be leaked to SP under the InPPTD framework.

Proof: SP receives two random numbers, $\{r_{m,k}\}_{m,k=1}^{M,K}$ and $\{r'_{m,k}\}_{m,k=1}^{M,K}$, in the report phase, four aggregated results $\sum_{k=1}^{K} \sum_{m=1}^{M} (x_{m,k} - t_m)^2$, $2^{a_k} \cdot \sum_{m=1}^{M} (x_{m,k} - t_m)^2$, $\sum_{k=1}^{K} w_k \cdot x_{m,k}$, and $\sum_{k=1}^{K} w_k$, and perturbed weights $\{w_k - a_k\}_{k=1}^{K}$ in the iteration phase, and the aggregated results of $\sum_{k=1}^{K} a_k$ and $\{\sum_{l=i}^{j} pc_k^l\}_{k=1}^{K}$ in the reward phase.

First, we only consider one iteration. Clearly, based on these information and the assumption that $K \geq 2$, SP cannot deduce individual worker's sensed data. For weight information, although SP can learn the numerator $\sum_{k=1}^{K}\sum_{m=1}^{M}(x_{m,k}-t_m)^2$ of weight formula, the denominator is perturbed by random number $\{2^{a_k}\}_{k=1}^{K}$. Thus, SP can only deduce $\{w_k - a_k\}_{k=1}^{K}$ by (13). While in the reward phase, SP only receives the aggregation results $\sum_{k=1}^{K} a_k$ and $\{\sum_{l=i}^{j} pc_k^l\}_{k=1}^{K}$. When $|j-i| \geq 2$, it is impossible for SP to deduce the weight information about workers.

Furthermore, we consider multiple iterations. Since a_k is randomly chosen in each iteration, a_k is always a new unknown variable for SP in $2^{a_k} \cdot \sum_{m=1}^{M}(x_{m,k}-t_m)^2$ in different iterations. Also, SP cannot calculate the difference in weights of worker k between any two iterations with $\{w_k - a_k\}_{k=1}^{K}$. As w_k differs in different iterations, $\{w_k\}_{k=1}^{K}$ are always new unknown variables for SP in $\sum_{k=1}^{K} w_k \cdot x_{m,k}$ and $\sum_{k=1}^{K} w_k$. So it is impossible to form an equation set with a unique solution by accumulating $2^{a_k} \cdot \sum_{m=1}^{M}(x_{m,k}-t_m)^2$, $\sum_{k=1}^{K} w_k \cdot x_{m,k}$, and $\sum_{k=1}^{K} w_k$ of multiple iterations.

For the aggregated distance, there are $K \cdot M$ unknown variables for SP in $\sum_{k=1}^{K}\sum_{m=1}^{M}(x_{m,k}-t_m)^2$. Therefore, to enable workers' sensitive and private data to be calculated, at least $K \cdot M$ iterations are needed to accumulate a set of $K \cdot M$ equations. Our experiments in Section VII-B will illuminate that the proposed algorithm can converge in just a couple of iterations. In most practical cases, the number of iterations is much less than $K \cdot M$. Therefore, it is reasonable to believe that the aggregated distance will not expose an individual worker's sensed data. When stronger security is needed, we further provide an extension method as follows.

In the secure weight estimation step, CP randomly chooses another random number b_k for each worker, and calculates $C = (\prod_{k=1}^{K} C_k)^{2^{b_k}}$, $C'_k = C_k^{2^{a_k+b_k}}$. Thus, $D(C) = 2^{b_k} \cdot \sum_{k=1}^{K}\sum_{m=1}^{M}(x_{m,k}-t_m)^2$, $D(C'_k) = 2^{a_k+b_k} \cdot \sum_{m=1}^{M}(x_{m,k}-t_m)^2$. The remaining process is the same as the basic scheme in Section V-D1. Since the aggregated distance is perturbed by random numbers $\{2^{b_k}\}_{k=1}^{K}$, it is impossible for SP to form an equation set with a unique solution by accumulating $D(C) = 2^{b_k} \cdot \sum_{k=1}^{K}\sum_{m=1}^{M}(x_{m,k}-t_m)^2$. ■

C. Preventing Privacy Disclosure to Workers and Providing Failure Resistance

In InPPTD, workers are kept out of the execution of protocol after reporting their sensed data. Therefore, workers will not receive any information from other parties. In addition, workers' sensed data and the interactions between SP and CP are transmitted through the secure channel. So workers cannot obtain sensitive and private information from the communications between other parties. Moreover, the feature that workers can be offline after reporting sensed data allows the implementation of our scheme to be free from the influence of the problem of workers' failures. Overall, our proposed protocol can prevent privacy disclosure to workers and provide effective failure resistance.

D. Resistance to Modification Attack From Malicious Workers

Since our scheme integrates the incentive mechanism and workers are assumed to be selfish, there may be some potential malicious workers trying to gain illegal rewards through system vulnerability. The most direct approach is to manipulate their corresponding weights.

In the proposed schemes in which workers participate in each round iteration of truth discovery, such as PPTD, EPPTD, and L-PPTD, weights are calculated by workers themselves. Although workers' privacy of weights is protected, workers can easily manipulate their respective weights. However, in InPPTD, after reporting their sensed data, workers are kept out of the execution of the protocol. Workers' weights are calculated in the interactions between SP and CP. Thus, workers cannot manipulate weights directly in plaintext. Considering that the Paillier cryptosystem has an additive homomorphism property, the interactions between SP and CP are through the secure channel. Thus, workers cannot tamper with any ciphertexts. Therefore, our scheme can resist modification attacks from malicious workers.

E. Reduce the Number of Lazy Workers

In InPPTD, we adopt a weight-based incentive mechanism, that is, the reward for a task of each worker is proportional to the corresponding weight. This mechanism is simple and effective. According to formula (1), for each task, when the sensed data of other workers are constant, the closer the sensed data of worker k is to the ground truth, the smaller the distance $d(x_{m,k}, t_m) = (x_{m,k}-t_m)^2$ is, and the greater the weight of worker k is. In general, data fabricated or inaccurately sensed by lazy workers are more random, and more likely to differ greatly from the ground truths when the estimated ground truths finally converge. Thus, lazy workers are usually assigned smaller weights and receive fewer rewards than normal workers. If the total rewards for each task are sufficient, lazy workers will not get more benefits by reducing costly sensing effort. They cannot even get enough rewards from our devised incentive mechanism to make up for the power and bandwidth overhead of sending data.

It should be noted that SP is able to recognize the lazy workers from the total rewards and remove them from the system. Our experiments in Section VII-E have shown that the gap in rewards between normal workers and lazy workers is obvious. Therefore, the only way to increase rewards is to be hardworking and provide higher quality/reliable sensed data.

F. Security Comparisons

We compare the proposed InPPTD with several other representative PPTDs in terms of the aspects of data privacy, weight privacy, modification resistance, failure resistance, and incentive integrated, as shown in Table I. Schemes of PPTD, EPPTD, and L-PPTD have no resistance to modification attacks and workers' failures since workers are involved in each iteration phase. In L^2-PPTD and NPPTD, workers can be offline after reporting data, thus modification attacks and workers' failures are prevented. However, L^2-PPTD cannot protect

TABLE I
SECURITY COMPARISON WITH OTHER PPTD SCHEMES

Scheme	Data Privacy	Weight Privacy	Modification Resistance	Failure Resistance	Incentive Integrated
PPTD [16]	Yes	Yes	No	No	No
EPPTD [15]	Yes	Yes	No	No	No
L-PPTD [17]	Yes	Yes	No	No	No
L^2-PPTD [17]	Yes	No	Yes	Yes	No
NPPTD [24]	Yes	Yes	Yes	Yes	No
InPPTD	Yes	Yes	Yes	Yes	Yes

weight privacy. Due to the use of GCs, as the performance analysis in [24], NPPTD requires a massive computation and communication overhead for carrying out the system. Specifically, our proposed InPPTD possesses all of the features while incorporating a secure incentive mechanism.

VII. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed InPPTD in terms of accuracy, convergence, computation overhead, communication overhead, and average rewards. Especially, lazy workers will be taken into consideration to analyze the impact of the performance of some parts. The experiments are conducted on an Intel Xeon 2.13-GHz server running Ubuntu 14.04 with 8-GB RAM. Since the sensed data for normal workers are most likely to be normally distributed [16], in the experiments, data of normal workers are generated from the normal distribution. While lazy workers have no prior knowledge of the given task, they generate random fake sensed data from the uniform distribution (the average of normal workers and lazy workers for each different object is randomly chosen from 10 to 10 000). In what follows, we present the experimental results of each part.

A. Accuracy

First, we evaluate the accuracy of the final estimated ground truths in terms of the number of workers. Many existing methods (e.g., PPTD [16], L^2-PPTD [17], and NPPTD [24]) and our proposed InPPTD all adopt the CRH framework as the basic truth discovery algorithm. If the initial ground truth, weights, and workers' data are the same, the final estimated ground truth will be also the same, thus, the accuracy will also be the same. Therefore, we only evaluate the accuracy of our proposed InPPTD and focus on analyzing the accuracy of the CRH framework when there are different numbers of lazy workers in the system.

We fix the number of objects as 20 and use the root of the mean-squared error (RMSE) to evaluate the deviation between the estimated results and the real truths. Here, RMSE $= (\sum_{m=1}^{M}(t'_m - t)^2/M)^{(1/2)}$, in which t'_m is the final estimated ground truth and t is its corresponding real value. In this experiment, different numbers of lazy workers are, respectively, put into normal workers' data to study their impact on the accuracy of results. The experimental results are shown in Fig. 5. On the one hand, with the increase of the number of workers, the truth discovery algorithm can obtain a more accurate estimated result. On the other hand, it can be seen that only a few lazy workers can damage the system and significantly reduce the accuracy of the estimation results.

Fig. 5. Accuracy for different number of lazy workers.

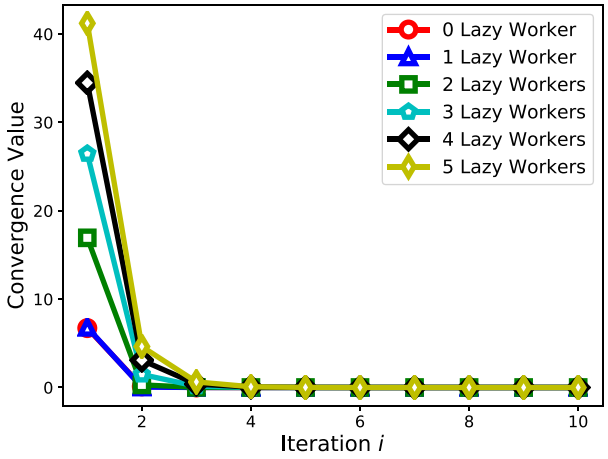


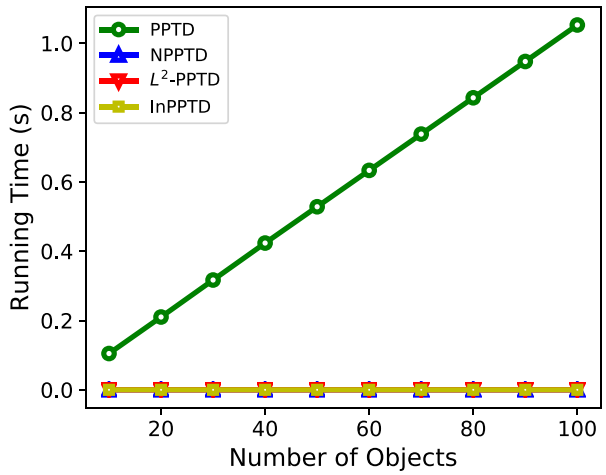
Fig. 6. Convergence for different number of lazy workers.

B. Convergence

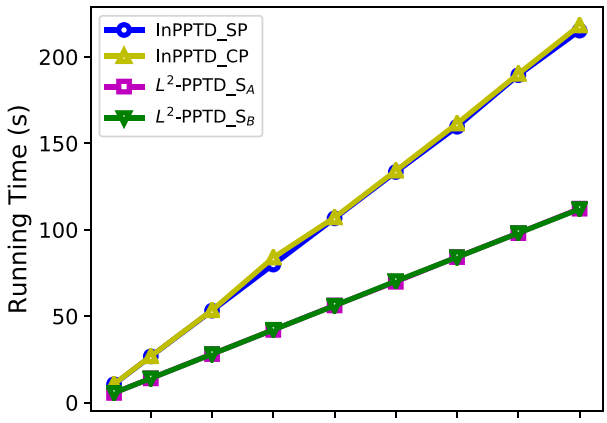
For evaluating convergence, we utilize the Squared Euclidean Distance $||t^i - t^{i-1}||_2$ as convergence value to measure the distance between the estimated truths in two consecutive iterations, in which t^i represents the vector of estimated truths in iteration i (t^0 are randomly generated). The experiments presented with five different number of lazy workers are shown in Fig. 6, where the number of objects and normal workers are fixed as 20 and 100, respectively. As we can see, although the number of lazy workers influences several convergence values at the beginning, the proposed algorithm converges quickly in just a couple of iterations.

C. Computation Overhead

In this part, we evaluate the computation performance of each phase in our proposed InPPTD. The experiments are implemented with the library of GMP [28], and the key size is set as a 2048-b strong security level. The results are compared



(a)



(b)

Fig. 7. Computation overhead in the report phase. (a) Running time for each worker. (b) Running time for servers.

with the baseline method PPTD [16], the state-of-art scheme L^2-PPTD [17] and NPPTD[24].

The experimental results of the report phase are shown in Fig. 7. The number of objects is set as 50, and the number of workers ranges from 10 to 200. We evaluate the time cost for the worker side and the server side, respectively. In our scheme, each worker does not perform any cryptographic operations and just perturbs sensed data with random numbers. The time cost for our scheme is almost negligible for each worker when compared to PPTD [see Fig. 7(a)]. Although, the computation overhead of servers (i.e., SP and CP) is larger than L^2-PPTD, NPPTD, and PPTD [as shown in Fig. 7(b)], it is acceptable since the report phase is executed only once. The servers are usually implemented in the cloud platform with rich computing resources. It can be executed quickly on the server side.

In the iteration phase, workers in L^2-PPTD and our InPPTD are not involved in the procedures. Therefore, only the time cost for servers is evaluated, while the time costs for workers and servers are measured in the PPTD scheme. The experimental results are presented in Fig. 8. The time cost for our InPPTD is much lower than L^2-PPTD and PPTD. It means that when the number of iterations is high, our InPPTD has

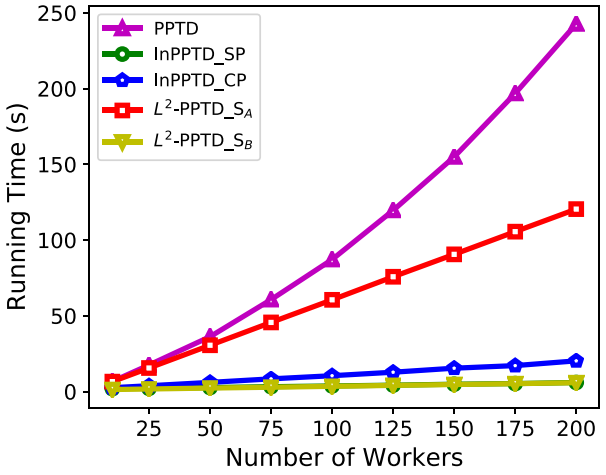


Fig. 8. Computation overhead for each iteration in the iteration phase.

a significant performance advantage. By the way, NPPTD is implemented with GC and it is an NPPTD scheme. So we cannot evaluate the time cost for each iteration of NPPTD. In fact, the major part of computing time is spent on GC generation, much longer than the computing time of InPPTD's server side.

Although, in our InPPTD, a few computing operations are conducted in plaintexts between two servers for the reward phase. They can be regarded as negligible compared to cryptographic operations. In summary, the computational overhead of our proposed InPPTD is very efficient on the worker side when compared to PPTD, and is more efficient than L^2-PPTD in the iteration phase.

The computation complexity analysis is as follows. In our proposed InPPTD, workers are only involved in the report phase. Therefore, the computational cost is $O(M)$ for each worker. SP and CP are both involved in the report phase, the iteration phase, and the reward phase. In the report phase, they both have to take $O(KM)$ encryptions. SP has also to take $O(KM)$ ciphertext multiplications. In the iteration phase, in each iteration, SP has to conduct $O(K+M)$ encryptions and $O(K+M)$ decryptions. CP has to conduct $O(K+M)$ encryptions, $O(KM)$ ciphertext multiplications and $O(KM)$ ciphertext exponentiations.

D. Communication Overhead

In this part, we evaluate the communication overhead with the related works in terms of different numbers of objects first. The size of keys and plaintexts are set as 2048 and 64 b, respectively, and the number of workers and the number of iterations are fixed as 100 and 5, respectively. Table II shows the experimental results. From the table, we can see that the communication overhead between workers and servers in our InPPTD is more efficient than PPTD, and very close to L^2-PPTD and NPPTD. Since PPTD adopts the single-server model, it does not have any cost of server–server communication overhead. While our InPPTD has a similar communication overhead with L^2-PPTD between the two servers.

Communication overhead between two servers of NPPTD is huge. Since it is difficult for NPPTD to support both

TABLE II
COMPARISON OF COMMUNICATION OVERHEAD

Overhead(MB)	Workers-Servers				Server-Server		
	PPTD	L^2-PPTD	NPPTD	InPPTD	PPTD	L^2-PPTD	InPPTD
10 Objects	3.029	0.015	0.015	0.031	-	0.623	0.770
100 Objects	35.34	0.152	0.152	0.305	-	5.128	5.494
500 Objects	124.5	0.763	0.763	1.526	-	25.15	26.49
1000 Objects	248.5	1.526	1.526	3.052	-	50.17	52.73

TABLE III
COMPARISON OF SERVER–SERVER COMMUNICATION OVERHEAD FOR NPPTD AND INPPTD

Scheme	Number of Workers							
	3	4	5	6	7	8	9	10
NPPTD(MB)	1246.9	1536.2	1825.5	2114.9	2404.2	2693.6	2982.9	3272.2
InPPTD(MB)	0.111	0.123	0.135	0.148	0.160	0.172	0.184	0.196

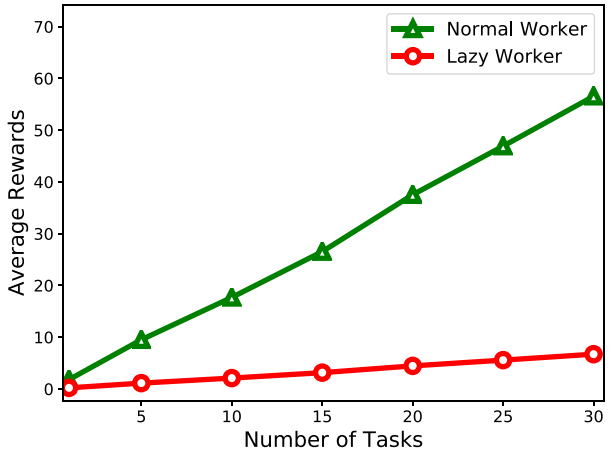


Fig. 9. Average rewards for normal worker and lazy worker.

numerous objects and large worker bases, we compare major communication overhead on servers of InPPTD and NPPTD with fewer workers and objects. The size of plaintexts is set as 16 b. The numbers of objects and iterations are set as 20 and 5, respectively. Table III shows the experimental results. InPPTD can save a lot of bandwidth compared to NPPTD.

E. Average Rewards

To evaluate how our designed incentive mechanism motivates workers to carry out sensing tasks normally rather than lazily, we accumulate the average rewards for each normal worker and lazy worker after multiple tasks. We set the average price for each task to be 75, and Fig. 9 shows the experimental results. From the figure, it can be seen that a normal worker can obtain about six times more rewards than a lazy worker. This huge gap in rewards between normal workers and lazy workers proves that our scheme can greatly lessen the lazy workers since they cannot get enough rewards from our devised incentive mechanism to make up for the power and bandwidth overhead of sending data.

VIII. CONCLUSION

In this article, we proposed an InPPTD, which is based on two noncolluding servers and the Paillier homomorphic cryptosystem. Considering that potential lazy workers exist in the system, an incentive mechanism is integrated into the truth discovery framework to motivate workers to behave normally. Security analysis demonstrates that our scheme can achieve strong privacy guarantees—workers' sensed data and weight information are both well protected. Performance evaluation shows that our InPPTD can significantly reduce both the computation and communication overheads. This enables our designed framework to be more feasible to be implemented in the crowdsensing system.

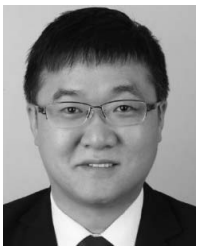
ACKNOWLEDGMENT

The authors sincerely thank all the anonymous referees for their invaluable suggestions that have led to the present improved version from the original manuscript.

REFERENCES

[1] Y. Li *et al.*, "Conflicts to harmony: A framework for resolving conflicts in heterogeneous data by truth discovery," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 8, pp. 1986–1999, Aug. 2016.
[2] S. Yao *et al.*, "Recursive ground truth estimator for social data streams," in *Proc. IEEE 15th Int. Conf. Inf. Process. Sensor Netw. (IPSN)*, 2016, p. 14.
[3] Y. Li *et al.*, "A survey on truth discovery," *ACM SIGKDD Explor. Newslett.*, vol. 17, no. 2, pp. 1–16, 2016.
[4] Y. Zheng, H. Duan, X. Tang, C. Wang, and J. Zhou, "Denoising in the dark: Privacy-preserving deep neural network based image denoising," *IEEE Trans. Depend. Secure Comput.*, early access, Mar. 25, 2019, doi: 10.1109/TDSC.2019.2907081.
[5] J. Xiong *et al.*, "A personalized privacy protection framework for mobile crowdsensing in IIoT," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 4231–4241, Jun. 2020.
[6] C. Cai, Y. Zheng, and C. Wang, "Leveraging crowdsensed data streams to discover and sell knowledge: A secure and efficient realization," in *Proc. IEEE/ACM 38th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, 2018, pp. 589–599.
[7] C. Xu, J. Wang, L. Zhu, C. Zhang, and K. Sharif, "PPMR: A privacy-preserving online medical service recommendation scheme in eHealthcare system," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5665–5673, Jun. 2019.
[8] L. Su *et al.*, "Generalized decision aggregation in distributed sensing systems," in *Proc. IEEE Real-Time Syst. Symp. (RTSS)*, 2014, pp. 1–10.
[9] C. Meng *et al.*, "Truth discovery on crowd sensing of correlated entities," in *Proc. 13th ACM Conf. Embedded Netw. Sensor Syst. (SenSys)*, 2015, pp. 169–182.
[10] C. Zhang, L. Zhu, C. Xu, K. Sharif, X. Du, and M. Guizani, "LPTD: Achieving lightweight and privacy-preserving truth discovery in CIoT," *Future Gener. Comput. Syst.*, vol. 90, pp. 175–184, Jan. 2019.
[11] Y. Zheng, H. Duan, and C. Wang, "Learning the truth privately and confidently: Encrypted confidence-aware truth discovery in mobile crowdsensing," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 10, pp. 2475–2489, Mar. 2018.
[12] Z. Zhang, S. He, J. Chen, and J. Zhang, "REAP: An efficient incentive mechanism for reconciling aggregation accuracy and individual privacy in crowdsensing," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 12, pp. 2995–3007, May 2018.
[13] D. Peng, F. Wu, and G. Chen, "Pay as how well you do: A quality based incentive mechanism for crowdsensing," in *Proc. 16th ACM Int. Symp. Mobile Ad Hoc Netw. Comput. (MobiHoc)*, 2015, pp. 177–186.
[14] H. Jin, L. Su, and K. Nahrstedt, "Theseus: Incentivizing truth discovery in mobile crowd sensing systems," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput. (Mobihoc)*, 2017, p. 1.
[15] Y. Zheng, H. Duan, X. Yuan, and C. Wang, "Privacy-aware and efficient mobile crowdsensing with truth discovery," *IEEE Trans. Depend. Secure Comput.*, vol. 17, no. 1, pp. 121–133, Jan./Feb. 2020.
[16] C. Miao *et al.*, "Cloud-enabled privacy-preserving truth discovery in crowd sensing systems," in *Proc. 13th ACM Conf. Embedded Netw. Sensor Syst. (SenSys)*, 2015, pp. 183–196.

[17] C. Miao, L. Su, W. Jiang, Y. Li, and M. Tian, "A lightweight privacy-preserving truth discovery framework for mobile crowd sensing systems," in *Proc. IEEE Conf. Comput. Commun. (Infocom)*, 2017, pp. 1–9.
[18] H. Jin, L. Su, H. Xiao, and K. Nahrstedt, "Inception: Incentivizing privacy-preserving data aggregation for mobile crowd sensing systems," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput. (Mobihoc)*, 2016, pp. 341–350.
[19] Q. Li, Y. Li, J. Gao, B. Zhao, W. Fan, and J. Han, "Resolving conflicts in heterogeneous data by truth discovery and source reliability estimation," in *Proc. ACM SIGMOD Int. Conf. Manag. Data (SIGMOD/PODS)*, 2014, pp. 1187–1198.
[20] X. Yin, J. Han, and S. Y. Philip, "Truth discovery with multiple conflicting information providers on the Web," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 6, pp. 796–808, Apr. 2008.
[21] G. Xu, H. Li, C. Tan, D. Liu, Y. Dai, and K. Yang, "Achieving efficient and privacy-preserving truth discovery in crowd sensing systems," *Comput. Security*, vol. 69, pp. 114–126, Aug. 2017.
[22] Y. Li *et al.*, "An efficient two-layer mechanism for privacy-preserving truth discovery," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Disc. Data Min.*, 2018, pp. 1705–1714.
[23] Y. Li *et al.*, "Towards differentially private truth discovery for crowd sensing systems," 2018. [Online]. Available: arXiv:1810.04760.
[24] X. Tang, C. Wang, X. Yuan, and Q. Wang, "Non-interactive privacy-preserving truth discovery in crowd sensing application," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2018, pp. 1–9.
[25] J. Lin, M. Li, D. Yang, G. Xue, and J. Tang, "Sybil-proof incentive mechanisms for crowdsensing," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, 2017, pp. 1–9.
[26] R. Gao *et al.*, "JigSaw: Indoor floor plan reconstruction via mobile crowdsensing," in *Proc. ACM 20th Annu. Int. Conf. Mobile Comput. Netw. (Mobicom)*, 2014, pp. 249–260.
[27] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proc. Int. Conf. Theory Appl. Cryptograph. Techn.*, 1999, pp. 223–238.
[28] (2018). *GMP Library*. [Online]. Available: https://gmplib.org/



Kaiping Xue (Senior Member, IEEE) received the bachelor's degree from the Department of Information Security, University of Science and Technology of China (USTC), Hefei, China, in 2003, and the Ph.D. degree from the Department of Electronic Engineering and Information Science (EEIS), USTC in 2007.

From May 2012 to May 2013, he was a Postdoctoral Researcher with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. He is currently an Associate Professor with the School of Cyber Security and the Department of EEIS, USTC. He has authored and coauthored more than 80 technical papers in the areas of communication networks and network security. His research interests include next-generation Internet, distributed networks, and network security.

Dr. Xue work won best paper awards in IEEE MSN 2017, and IEEE HotICN 2019, and best paper runner-up award in IEEE MASS 2018. He serves on the Editorial Board of several journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, and *Ad Hoc Networks*. He was the Guest Editor or the Lead Guest Editor for special issues in several journals and magazines, including the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, *IEEE Communications Magazine*, and IEEE NETWORK. He is a Fellow of IET.

Bin Zhu (Graduate Student Member, IEEE) received the B.S. degree in information security from the School of Cyber Security, University of Science and Technology of China, Hefei, China, in 2019, where he is currently pursuing the graduation degree in information security.

His research interests include network security and applied cryptography.

Qingyou Yang (Member, IEEE) received the B.S. degree in information security from the School of the Gifted Young, University of Science and Technology of China (USTC), Hefei, China, in 2016, and the M.S. degree from the Department of Electronic Engineering and Information Science, USTC in 2019.

He is currently a Software Engineer with Baidu, Shenzhen, China. His research interests include network security and cryptography.

Na Gai (Student Member, IEEE) received the B.S. degree from the Department of Information Security, University of Science and Technology of China, Hefei, China, in 2018, where she is currently pursuing the graduation degree in information security with the Department of Electronic Engineering and Information Science.

Her research interests include network security protocol design and analysis.

David S. L. Wei (Senior Member, IEEE) received the Ph.D. degree in computer and information science from the University of Pennsylvania, Philadelphia, PA, USA, in 1991.

From May 1993 to August 1997, he was the Faculty of Computer Science and Engineering, University of Aizu, Aizuwakamatsu, Japan, (as an Associate Professor and then a Professor). He is currently a Professor with the Computer and Information Science Department, Fordham University, Bronx, NY, USA. He has authored and coauthored more than 120 technical papers in various archival journals and conference proceedings. His research interests include cloud computing, big data, IoT, and cognitive radio networks.

Prof. Wei was the Guest Editor or the Lead Guest Editor for several special issues in the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE TRANSACTIONS ON CLOUD COMPUTING, and the IEEE TRANSACTIONS ON BIG DATA. He also served as an Associate Editor for IEEE TRANSACTIONS ON CLOUD COMPUTING from 2014 to 2018, and *Journal of Circuits, Systems and Computers* from 2013 to 2018.

Nenghai Yu (Member, IEEE) received the B.S. degree from Nanjing University of Posts and Telecommunications, Nanjing, China, in 1987, the M.E. degree from Tsinghua University, Beijing, China, in 1992, and the Ph.D. degree from the Department of Electronic Engineering and Information Science (EEIS), University of Science and Technology of China (USTC), Hefei, China, in 2004.

Since 1992, he has been a Faculty with the Department of Electronic Engineering and Information Science, USTC, where he is currently a Professor. He is the Executive Director of the Department of EEIS, USTC, and the Director of the Information Processing Center, USTC. He has authored or coauthored more than 130 papers in journals and international conferences. His research interests include multimedia security, multimedia information retrieval, video processing, and information hiding.