IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 27, 2019, accepted March 17, 2019, date of publication April 1, 2019, date of current version April 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2908429

LBOA: Location-Based Secure Outsourced Aggregation in IoT

JUNWEI ZHANG, YUE ZONG, CHAO YANG, YINBIN MIAO, AND JINGJING GUO
Department of Cyber Engineering, Xidian University, Xi'an 710071, China

Corresponding author: Yue Zong (zy_zongyue@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61472310, Grant U1536202, Grant U1708262, Grant 61672413, Grant 61672415, Grant 61671360, Grant 61602360, and Grant 61702404, and in part by the China 111 Project under Grant B16037.

ABSTRACT Secure outsourced aggregation in the Internet of Things (IoT) can solve the problem that sensing devices are limited in energy and bandwidth by outsourcing data aggregation task to a third-party service provider. Location-based secure outsourced aggregation (LBOA), aggregating data whose location satisfies user's location strategy, is very important in some location-critical scenarios (e.g., smart homes, intelligent transportation, and smart city). Recent work studied secure data aggregation to reduce transmission overhead and network bandwidth by optimizing topology of networks or adopting the cryptographic approach. However, as far as we know, scarcely any work considers the location information of the data source and the privacy protection of the data at the same time in the studies of secure outsourced aggregation. In this paper, we first propose an LBOA scheme LBOA_{Max}, which can return the maximum value of sensory data whose location satisfies location strategy by applying one-way chain, order-preserving encryption, and some other cryptographic operation. Then, we proposed scheme LBOA_{Top-k} and scheme LBOA_{Sum}, which can return the largest k values of data and the summation value of data based on location, respectively. The security analysis results show that our schemes can satisfy the defined requirements and the experiment results show that our schemes are feasible and efficient for each entity in practice.

INDEX TERMS Location-based, secure aggregation, cloud computation, privacy protection, one-way chain, order-preserving encryption.

I. INTRODUCTION

With the continuous development and improvement of wireless network technology [1], IoT has been more and more widely used in our life [2]. For example, in smart homes [3], smart devices are connected to the external service through the network of IoT, realizing interaction between external service and smart devices. In the field of electronic medical system [4], patients' vital signs such as heart rate and blood pressure can be monitored by wearable devices. IoT is also widely used in intelligent transportation, environmental monitoring, military and many other fields.

However, sensing devices in IoT are usually limited in energy and bandwidth [5], and their computation and storage power are limited. What's more, in general, data requester cannot interact with sensing devices directly in IoT. Outsourcing [6]–[8] data aggregation task to a third service provider such as aggregator in cloud [9], [10] is an effective way to solve the above problems.

Unfortunately, aggregator is usually untrusted, it may have malicious behaviors. On one hand, aggregator may filch sensory data. In some fields such as electronic medical, the confidentiality of sensory data is very important [11]. Sensing devices do not want aggregator to learn any knowledge about the data while outsourcing aggregation task to aggregator. Thus, how to ensure the confidentiality of data is a challenging question in outsourced aggregation. On the other hand, aggregator may tamper the aggregation process and report a wrong aggregated result to data requester. Thus, the verifiability of the aggregated result is also essential in outsourced aggregation.

Most of the existing work related to secure data aggregation has not considered the location of data source. However, location-based secure outsourced aggregation, i.e. securely aggregate data under location strategy, is very important in some location-critical scenarios. In many IoT applications,

The associate editor coordinating the review of this manuscript and approving it for publication was Guoqi Xie.

the data collected by the sensing device is closely related to the location information of the sensing device [12]. For example, in intelligent transportation, the location-critical data sensed by smart devices is one of the most important elements for monitoring, analyzing and forecasting road conditions. In smart homes, the location of data source plays an important role. In applications such as geographic key distribution [13] and geographic routing [14], the smart devices' location also plays a crucial role.

Therefore, a location-based secure outsourced aggregation scheme is necessary in IoT to realize aggregating data based on devices' location. Due to the reason that Max, Top-k and Sum are some of the most basic operations of data. Thus, we try to construct location-based secure outsourced aggregation schemes in which aggregator only aggregates data whose location is at the specified position and return the Max or Top-k or Sum value to data requester. However, this work is very challenging. First, it is a challenging task to achieve location-based in IoT because the traditional methods to determine location is not efficient and not suitable to IoT. Second, it is difficult to protect data confidentiality against outside attacker and untrusted aggregator while aggregator needs to aggregate data and return the aggregated result to data requester. Third, it is hard to ensure that aggregator does not have malicious behaviors and guarantee that the user could verify the validity of aggregated result because user cannot get all the raw sensory data in the whole process of aggregation. What's more, the privacy of devices' location and the confidentiality of user's location strategy should also be protected.

Our contributions. We put emphasis on realization of secure outsourced aggregation based on location in this study. First, in order to determine whether the sensing device's location satisfies location strategy or not, we adopt vector operation, which using the geometric relationship of the spatial data's location. Second, in order to protect the confidentiality of data, we adopt order-preserving encryption scheme [15] to realize comparing directly on ciphertexts. Third, in order to ensure the verifiability of the aggregation process, we adopt one-way chain [16], which consists of a series of one-way functions. At the same time, we utilize public key encryption algorithm to achieve protecting the privacy of data's location and the confidentiality of user's location strategy.

Our contributions in this study are fourfold.

1) We design the system model of location-based secure outsourced aggregation in IoT. Then we propose the threat model. Next we propose our design goal. The system model of LBOA defines the participants including location-sensitive devices, user and cloud service provider. The system model also defines the participants' task. The threat model in this study describes the adversarial behaviors including data tampering, cheating, data deleting and so on. The design goal of LBOA in this study presents the requirements such as providing service based on location, achieving location privacy protection, data confidentiality protection and location strategy confidentiality protection.
2) We propose a novel location-based secure outsourced aggregation scheme $LBOA_{Max}$ which can return the maximum value under user's location strategy. Then we propose scheme $LBOA_{Top-k}$ and scheme $LBOA_{Sum}$ which can return the largest k values and the summation value under user's location strategy respectively. Our schemes could aggregate data whose location is at specified location correctly. They could also protect the confidentiality of data and the privacy of the location strategy.
3) We theoretically analyze the security of LBOA. The analysis results show that our schemes satisfy our design goal. At the same time, our schemes support much more data aggregation query operations and are much more secure than existing schemes.
4) We report experimental evaluations of LBOA. The evaluation results show that our schemes are efficient and feasible in practice.

II. RELATED WORK

A. AGGREGATION

Tan and Körpeoglu [17] proposed a power efficient data gathering and aggregation scheme in wireless sensor networks. Rajagopalan and Varshney [18] introduced data aggregation techniques in sensor networks. Chen et al. [19] presented a data aggregation scheme with distributed randomized algorithms. Hekmat and Van Mieghem [20] constructed the shortest path aggregation tree that maximizes network lifetime. Chang and Yen [21] constructed a spanning tree based aggregate routing algorithm, selecting the node that performs the data aggregation operation through the coding tree. Lee *et al.* [22] presented a construct which use geographical route to balance network traffic, and optimize network lifetime and aggregate data rate through optimization methods. However, these data aggregation schemes are mainly concerned with the issue of energy conservation without focusing on the security of data aggregation.

The security problem of data aggregation [23] began to be studied at home and abroad in recent years, but they mainly focus on safe energy conservation and safe routing at early period. Some work focus on the security of data later. Work in [24], [25] introduced privacy homomorphism technology and proposed data aggregation schemes based on privacy homomorphism. These schemes aggregate encrypted data directly without decryption in order to protect end-to-end privacy of data. Zhu *et al.* [26] proposed a secure data aggregation scheme based on commitment-proof and back testing in order to protect the integrality of data. Chen *et al.* [27] combined homomorphic encryption with bilinear-based signature, and proposed a recoverable data aggregation scheme to ensure data privacy and integrity. Li *et al.* [28] proposed a privacy preserving data aggregation scheme for mobile edge computing assisted IoT application. However, all of these work only focus on protecting data privacy and integrity

but not pay attention to the basic data aggregation query operations such as max, min, count, top-k and so on.

Later, some work focus on the basic query operations in secure aggregation. Chan *et al.* proposed a secure data aggregation scheme (SIA) [29]. However, this scheme only supports one aggregator, and it is not applicable to large amounts of data. Then they extended SIA and proposed a secure hierarchical aggregation scheme (SHIA) [30] that supports multiple aggregators. This work only supports limited sets of aggregation functions but not supports aggregation functions such as max and top-k. Nath *et al.* [16] proposed a secure outsourced aggregation scheme which uses one-way chain and related cryptography operations to ensure the security of aggregation. This work supports several aggregation functions such as max, count and top-k, but it doesn't protect the privacy of data.

In summary, none of the above data aggregation schemes consider the location of data source.

B. LOCATION VERIFICATION

Vora and Nesterenko [31] proposed a location verification scheme that can achieve verifying location in-region of provers. Sastry *et al.* [32] proposed a location verification scheme which can realize verification in a small circular region. Čapkun *et al.* [33] proposed a scheme which can verify the location through mobile base stations. Sciancalepore *et al.* [34] proposed a scheme which realize secure location verification by the help of meteor burst communication.

Chandran *et al.* [35] proposed location-based cryptography in 2009, using user's geographic location information as the user's unique credential. Under BRM model, they proposed secure positioning (SP) protocol which can be proven secure. SP protocol can be used to verify whether the user's location is at the specified location or not. However, SP protocol requires multiple verifiers work together to verify the legitimacy of the location.

Zhang *et al.* [12] proposed a universally composable secure positioning scheme in the bounded retrieval model. It realized secure location verification. Zhang *et al.* [36] investigated a scheme which can achieve secure geographical area verification without pre-shared secret. This work is propitious to massive location-critical devices in IoT.

All of work above need verifiers to realize location verification. They also require precise time synchronization, and are not robust to computation delay.

C. LOCATION-BASED SOLUTION

Kwon *et al.* [13] proposed a scheme of location-based pairwise key distribution for wireless sensor networks which can achieve perfect resilience and higher connectivities with less resources. Li *et al.* [14] proposed an energy efficent cooperative geographic routing scheme in wireless sensor networks which based on sensor nodes' location information. Zhang *et al.* [37] proposed a position based key exchange scheme which can achieve both security and performance perspectives. Ji *et al.* [38] proposed a blockchain-based multi-level privacy-preserving location sharing scheme which can achieve security and flexibility of location privacy protection. Gao *et al.* [39] proposed a logistics information privacy protection scheme with position and attribute-based access control which can achieve privacy protection of both logistics information and personal information.

Wang *et al.* [40] combined data's location with searchable encryption and proposed a secure geometric search scheme on encrypted spatial data. This scheme determined whether the data's location is at specified location or not by executing vector operation which the geometric relationship of the spatial data's location is used. This method of verifying the location of data is very efficient and is suitable to IoT.

III. PRELIMINARIES

A. VECTOR OPERATION BASED ON GEOMETRIC RELATIONSHIP

The essence of vector operation based on geometric relationship is based on circular geometry [40].

Given a random point (x, y) and a circle which (x_c, y_c) is the circle center and R is the radius. If the point (x, y) is on the boundary of the given circle, we have:

$$\begin{aligned}
&(x-x_c)^2+(y-y_c)^2-R^2\\
&=x^2+y^2-2x\cdot x_c-2y\cdot y_c+x_c^2+y_c^2-R^2\\
&=(x^2+y^2)\cdot 1+(-2x)\cdot x_c+(-2y)\cdot y_c+1\cdot(x_c^2+y_c^2-R^2)\\
&=<\vec{u}=(x^2+y^2,-2x,-2y,1),\vec{v}=(1,x_c,y_c,x_c^2+y_c^2-R^2)>\\
&=0.
\end{aligned}$$

B. ORDER-PRESERVING ENCRYPTION

Order-preserving encryption (OPE) [15] is a special symmetric encryption scheme which guarantees the orders of ciphertexts are the same as the orders of plaintexts.

An OPE scheme generally contains *KeyGen*, *Enc* and *Dec*. Specifically,

- $KeyGen(1^\lambda) \rightarrow k$: input a security parameter λ, output a secret key k.
- $Enc(k, m) \rightarrow C_m$: input a secret key k and plaintext m, output ciphertext C_m.
- $Dec(k, C_m) \rightarrow m$: input a secret key k and ciphertext C_m, output plaintext m.

OPE has the property that:

$$m_1 < m_2 \Leftrightarrow Enc_{OPE}(k, m_1) < Enc_{OPE}(k, m_2).$$

C. ONE-WAY FUNCTION AND ONE-WAY CHAIN

A function $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ is a one-way function [41] if $f(x)$ can be calculated by a polynomial time algorithm which takes input x, but the possibility that any adversary attempt to compute a pseudo-inverse for f successfully can be negligible. That is, for every PPT adversary A there is a negligible function v_A such that for a sufficiently large k,

$$\begin{aligned}
Pr[z \leftarrow A(1^k, y) : x \xrightarrow{R} \{0, 1\}^k;\\
y \leftarrow f(x); f(z) = y] \leq v_A(k).
\end{aligned}$$

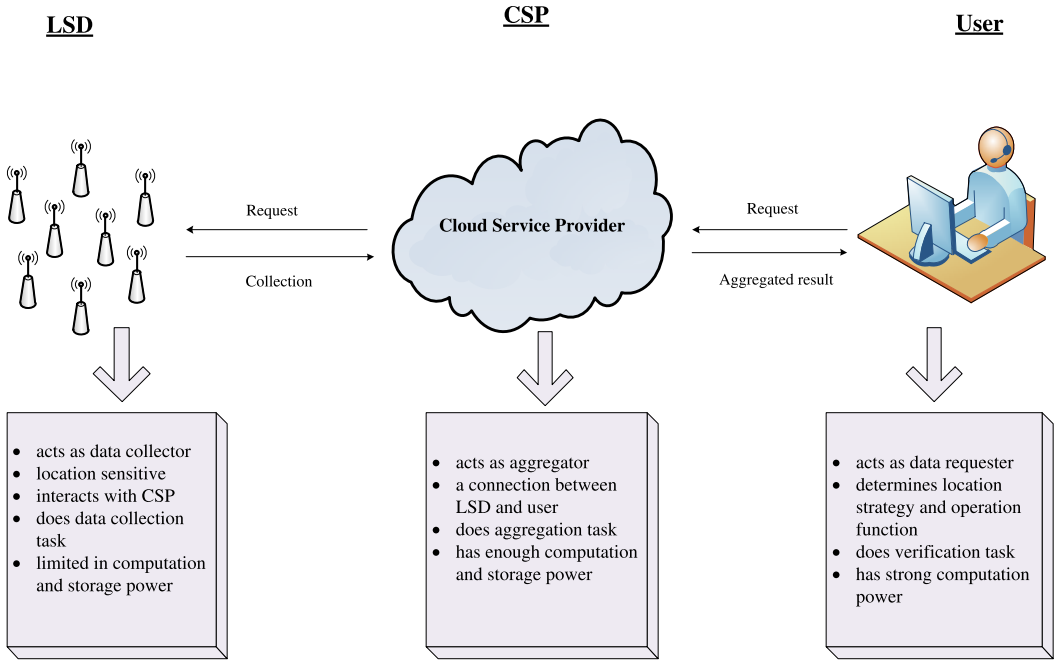

FIGURE 1. System model.

One-way chain [42] is based on secure one-way function and random seed s. One-way chain recurs applying one-way function for multiple times. We use $F^x(s)$ to denote recursively applying one-way function F for x times. The chain has the value $F^1(s), F^2(s), F^3(s), F^4(s), \ldots$ sequentially. Due to the property of one-way function, given the value of $F^x(s)$, we can calculate $F^{x'}(s)$ if $x' > x$, however, if $x' < x$, the probability to get $F^{x'}(s)$ can be negligible. In other words, one-way chain can only be rolled forward, it is infeasible to be rolled backward.

IV. PROBLEM FORMULATIONS

A. SYSTEM MODEL

The system model of location-based secure outsourced aggregation (LBOA) is shown in Figure 1. It contains of three entities: user (User), cloud service provider (CSP) and location-sensitive device (LSD).

If a user, which can also be called a data requester, wants to obtain some values (such as the maximum value, the largest k values, the summation value and so on) of LSD's sensory data, he (or she) will send request to the CSP. CSP asks for LSD's data after receiving the request, then LSD submits data to CSP. CSP does aggregation operation after receiving all the data and then returns aggregated result to the user. User verifies aggregated result after receiving the aggregated result. More specifically, we describe as follows.

User. User acts as data requester. User determines the location strategy and operation function. He (or she) would like to get LSD's data that satisfying the location strategy and the operation function. However, user cannot interact with LSD directly. Thus, user sends requests to CSP, then CSP aggregates LSD's data and returns aggregated result to user. User will verify the correctness and completeness of the aggregated result after receiving the aggregated result from CSP.

CSP. CSP acts as a connection between location sensitive devices and user. There may be one or multiple aggregators in CSP to do aggregation task. CSP has the ability to do calculation operation and aggregation operation. He (or she) receives data from LSDs and aggregates the data to get an aggregated result, and then sends the aggregated result to user.

LSD. LSD acts as data collector. There are multiple location sensitive devices. Each LSD has a location coordinate. LSD can interact with CSP. LSD may collect sensory data and then outsource data aggregation task to CSP.

B. THREAT MODEL

We considered four participants in threat model.

1) **User.** We assume the user is totally trusted.
2) **CSP.** The cloud service provider may have three types of malicious behaviors. First, CSP may tamper the aggregation process and trick user with wrong aggregated result. Second, CSP may ignore or delete some location-sensitive devices' data. Third, CSP may add an illegal LSD into the aggregation process.
3) **LSD.** The location-sensitive devices are honest-but-curious. They will not transmit faked data, they will not misreport their location, but they will try to pry the location of other location-sensitive devices.
4) **Outside Attacker.** Outside attacker also may cause attacks. Outside attacker refers to the adversary who obtains some knowledge about the legitimate

location-sensitive device's data or user's location strategy via public channels. An outside attacker may intercept the data sensed by LSD and may intercept user's location strategy.

C. DESIGN GOAL

Accoring to the requirements and the threat model of location-based secure outsourced aggregation, the proposed schemes should satisfy the following design goal:

1) Achieve secure aggregation based on sensing devices' location (LB). Our schemes should realize aggregation based on location. CSP should aggregate data whose location satisfies user's location strategy.
2) Support some basic query operations in aggregation such as Max, Top-k and Sum. Our schemes should achieve returning the Max (Top-k/Sum) value of data under user's location strategy correctly.
3) Guarantee the verifiability of aggregation (AV). Our schemes should guarantee CSP does not tamper the aggregation process. Our schemes should also ensure that the correctness and completeness of the aggregated result reported by CSP can be verified by user.
4) Guarantee the privacy of the LSD's location (LSDP). Our schemes should ensure any entity except the location-sensitive device itself and the totally trusted user could not learn any location information about the legitimate sensing devices.
5) Guarantee the confidentiality of the data (DC). Our schemes should guarantee the data confidentiality against outside attacker (DCO). As the CSP is untrusted, our schemes should also guarantee the data confidentiality against CSP (DCC). In other words, our schemes should ensure the data sensed by location-sensitive devices will not be intercepted by outside attacker and CSP.
6) Guarantee the confidentiality of location strategy (LSC). Our schemes should ensure the confidentiality of user's location strategy. Only the user itself and location-sensitive devices can learn the location strategy.

V. THE PROPOSED SCHEMES

Starting from this section, we present our LBOA schemes. The notations used in this study are listed in Table 1. Section 5.*A* explains scheme $LBOA_{Max}$ which user requests the maximum value of data whose location satisfies user's location strategy. Section 5.*B* explains scheme $LBOA_{Top-k}$ which user requests the largest k values of data whose location satisfies user's location strategy. Section 5.*C* explains scheme $LBOA_{Sum}$ which user requests the summation value of data whose location satisfies the location strategy.

A. $LBOA_{MAX}$

The overview of $LBOA_{Max}$ is shown in Figure 2. Scheme $LBOA_{Max}$ consists of the following five phases: the initialization phase, the request phase, the collection phase, the aggregation phase and the verification phase. Initialization phase generates parameters which are required later. During the request phase, user formulates location strategy and operation function, then sends request messages to CSP. CSP transmits request messages to LSD and requests LSD to submit the value of sensory data. During the collection phase, location-sensitive devices judge whether the location of data meet user's location strategy or not, and then submit the response to CSP. During the aggregation phase, CSP does aggregation task, and then sends the aggregated result to user. During the verification phase, user verifies the aggregated result reported by the CSP.

TABLE 1. Notations.

Notation	Description
S_i	the $i'th$ location-sensitive device
ID_i	the identity of S_i
(PK, SK)	the public/private key pair
K_i	pre-shared symmetric key consulted between S_i and user
GK	the group key used in broadcast
K_{OPE}	the symmetric key used in OPE
Enc	the asymmetric encryption algorithm
Dec	the asymmetric decryption algorithm
Enc_{OPE}	the order-preserving encryption algorithm
Dec_{OPE}	the order-preserving decryption algorithm
$C_{Request}$	the ciphertext of user's request message
(x_i', y_i')	S_i's location coordinate
(x_i, y_i)	S_i's discretizated integer coordinate
$\vec{u_i}$	the vector related to S_i's location
$\vec{v_k}$	the vector related to location strategy
C_{u_i}	the ciphertext of S_i's location
Sig	the signature algorithm
MAC	message authentication code
a	a random integer
d_i	the raw value of S_i's sensory data
D_i	the processed value of S_i's sensory data
C_{D_i}	the ciphertext of D_i
f	the operation function
$\mathbb{P}$	the location strategy
$epoch\#$	the timestamp of the current time
L	error tolerance in coordinate discretization
S_i^+, S_i^-	the parameters generated by S_i
F	the one-way function
$F^x(s)$	recursively applying one-way function F for x times
$\odot$	the folding operation(modulo multiplication)

1) INITIALIZATION

During the initialization phase, each S_i $(1 \leq i \leq n)$ with an ID_i registers a public/private key pair (PK_i, SK_i). User also has a certified public/private key pair (PK_U, SK_U). Each S_i consults a symmetric key K_i with user. And then user maintains a group key GK with all LSDs.

2) REQUEST

The specific steps of this phase are described below.

(1) The user picks a random $a \in \mathbb{Z}_p$, assuming that the range of d_i is $d_i \in [lowest, largest]$, so the range of a is $a \in [2 - lowest, \infty]$. In other words, the random integer a meets the condition that for any d_i, $d_i + a \geq 2$ is satisfied.

(2) The user decides the operation function f. In this $LBOA_{Max}$ scheme, $f = Max$.

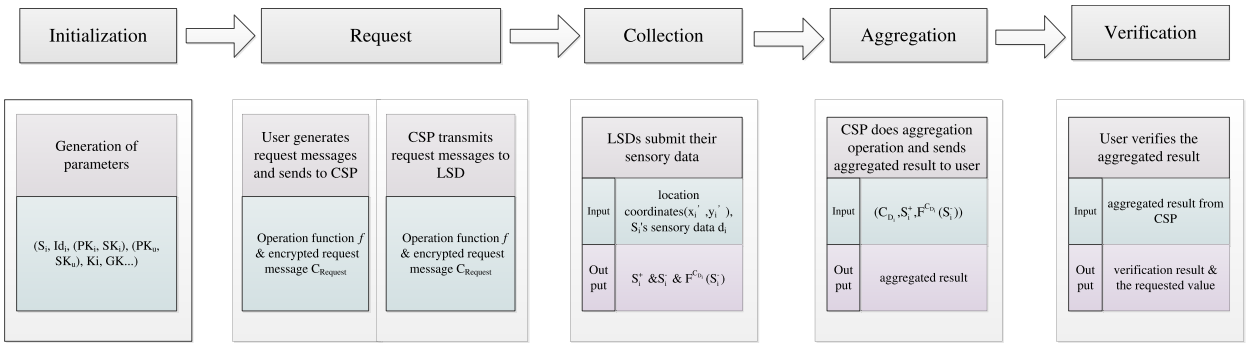

FIGURE 2. Overview of $LBOA_{Max}$.

(3) The user presets the location strategy $\mathbb{P} = \{P_1, P_2, \ldots, P_n\}$, each P_i contains of m vectors $\vec{v}_i$ $(1 \leq m \leq 2)$, where $\vec{v}_i = (1, x_{ic}, y_{ic}, x_{ic}^2 + y_{ic}^2 - r_i^2)$.

(4) The user utilizes group key GK to encrypt a, $\mathbb{P}$ and the symmetric key K_{OPE}, i.e.,

$$Enc(GK, a\|\mathbb{P}\|K_{OPE}\|epoch\#) \rightarrow C_{Request},$$

where $epoch\#$ denotes timestamp of the current time.

(5) User sends $\{f = Max\|C_{Request}\}$ to CSP.

(6) CSP preserves the operation function $f = Max$ after receiving all the messages send by the user, at the same time, CSP transmits $\{f = Max\|C_{Request}\}$ to location-sensitive devices, and requests to collect LSD's sensory data.

3) COLLECTION

After receiving the messages from CSP, location-sensitive devices submit their sensory data. The details of this phase are illustrated as follows.

(1) Each S_i decrypts the broadcast message to get the plaintexts of a, $\mathbb{P}$ and K_{OPE}, i.e.,

$$Dec(K_i, C_{Request}) \rightarrow a\|\mathbb{P}\|K_{OPE}\|epoch\#.$$

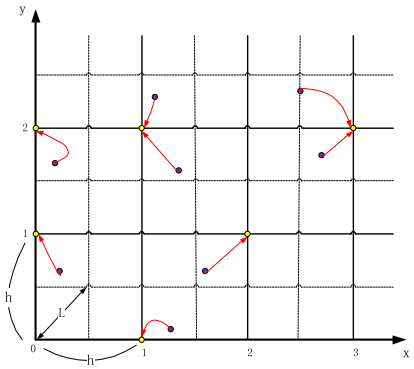

FIGURE 3. Coordinate discretization.

(2) As shown in Figure 3, each LSD discretizes its position coordinate (x_i', y_i') to get integer coordinate (x_i, y_i). If the error tolerance is no more than L. For any noninteger point, we can approximate it as a nearest integer point which the distance between them is no more than L. The detail of coordinate discretization is described in Algorithm 1. Each S_i utilizes its integer location coordinate (x_i, y_i) to generates a vector $\vec{u}_i = (x_i^2 + y_i^2, -2x_i, -2y_i, 1)$. Signing $\vec{u}_i$ with $S_i's$ secret key SK_i, and then encrypt $\vec{u}_i$ with user's public key PK_u, i.e.,

$$Enc(PK_U, Sig(SK_i, \vec{u}_i)) \rightarrow C_{u_i}.$$

Algorithm 1 Coordinate Discretization

Require:
Point location coordinate (x', y'); ouput space S; Error tolerance $L = \frac{\sqrt{2}}{2}h$

Ensure:
Integer coordinate (x, y)

LSD executes:
$(\lfloor x' \rfloor, \lceil x' \rceil) \leftarrow x'$;
if $\lfloor x' \rfloor \leq x' < \lfloor x' \rfloor + \frac{1}{2}$ **then**
  $x = \lfloor x' \rfloor$;
else if $\lfloor x' \rfloor + \frac{1}{2} \leq x' \leq \lceil x' \rceil$ **then**
  $x = \lceil x' \rceil$;
end if
$(\lfloor y' \rfloor, \lceil y' \rceil) \leftarrow y'$;
if $\lfloor y' \rfloor \leq y' < \lfloor y' \rfloor + \frac{1}{2}$ **then**
  $y = \lfloor y' \rfloor$;
else if $\lfloor y' \rfloor + \frac{1}{2} \leq y' \leq \lceil y' \rceil$ **then**
  $y = \lceil y' \rceil$;
end if
return integer coordinate(x, y)

(3) Each S_i calculates the value of $< \vec{u}_i, \vec{v}_k >$ $(1 \leq k \leq m)$ respectively, if $\exists P_i \forall \vec{v}_k ((P_i \in \mathbb{P}) \bigwedge (\vec{v}_k \in P_i) \rightarrow (\Pi_{k=1}^{m} < \vec{u}_i, \vec{v}_k >= 0))$, S_i sets $D_i = d_i + a$ (d_i denotes the raw value of S_i's sensory data), else, sets $D_i = 1$.

(4) Each location-sensitive device S_i encrypts D_i by order-preserving encryption, i.e.,

$$Enc_{OPE}(K_{OPE}, D_i) \rightarrow C_{D_i}.$$

(5) Each location-sensitive device S_i generates:

$$S_i^+ = \{ID_i, C_{u_i}, MAC_{K_i}(C_{D_i}||C_{u_i}||epoch\#)\},$$
$$S_i^- = \{MAC_{K_i}(epoch\#)\},$$

where $epoch\#$ denotes timestamp of the current time.

(6) Each location-sensitive device S_i calculates $F^{C_{D_i}}(S_i^-)$. We use RSA as the one-way function F in this study.

4) AGGREGATION

During the phase of aggregation, the following steps are executed in sequence.

(1) Each S_i sends $(C_{D_i}, S_i^+, F^{C_{D_i}}(S_i^-))$ to CSP.

(2) CSP compares the value of C_{D_i} after receiving all the message from LSDs. Assuming that the value of S_m's data is the maximum value, which is known by the CSP. We use C_{D_m} to represent the encrypted maximum value.

(3) CSP computes the aggregated result

$$(C_{D_m}, S_m^+, \odot_i F^{C_{D_m}}(S_i^-)),$$

where $\odot_i F^{C_{D_m}}(S_i^-) = \prod_{i=1}^{n}(F^{C_{D_m}}(S_i^-)) mod(pq)$, p and q are two large prime number used in one-way function (RSA).

(4) CSP sends the aggregated result

$$(C_{D_m}, S_m^+, \odot_i F^{C_{D_m}}(S_i^-))$$

to the user.

5) VERIFICATION

User verifies the correctness of the aggregated result reported by CSP in this phase. The specific steps of this phase are described below.

(1) User verifies the validity of ID_m and MAC_{K_m} first after receiving

$$S_m^+ = \{ID_m, C_{u_m}, MAC_{K_m}(C_{D_m}||C_{u_m}||epoch\#))\}$$

from CSP.

(2) User computes all individual S_i^-, for each S_i^-, computes $F^{C_{D_m}}(S_i^-)$, and then computes $\odot_{i-user} F^{C_{D_m}}(S_i^-)$, where

$$\odot_{i-user} F^{C_{D_m}}(S_i^-) = \prod_{i=1}^{n}(F^{C_{D_m}}(S_i^-)) mod(pq).$$

(3) If the $\odot_{i-user} F^{C_{D_m}}(S_i^-)$ computed by user is the same as the $\odot_i F^{C_{D_m}}(S_i^-)$ in the aggregated result reported by CSP, user accepts the aggregated result.

(4) User decrypts C_{D_m} to get the plaintext of the maximum value at the specified location. The maximum value of location-sensitive devices' sensory data that meets the location strategy is d_m. At the same time, user decrypts C_{u_m} in S_m^+ to get the location (x_m, y_m) from $\vec{u_m}$, i.e.,

$$\begin{aligned}
&Dec_{OPE}(K_{OPE}, C_{D_m}) \rightarrow D_m,\\
&d_m = D_m - a,\\
&Dec(PK_m, Dec(SK_U, C_{u_m})) \rightarrow \vec{u_m},\\
&\vec{u_m} = (x_m^2 + y_m^2, -2x_m, -2y_m, 1).
\end{aligned}$$

The flow diagram of protocol $LBOA_{Max}$ is shown in Figure 4.

In order to help reader to understand scheme $LBOA_{Max}$ better, we take a simple example. As shown in Figure 5, the user specifies the location strategy $\mathbb{P}$ which is $\vec{v} = (1, 1, 1, 1)$. There are nine LSDs. S_2, S_4, S_6 and S_8 satisfy the location strategy while S_1, S_3, S_5, S_7, S_9 do not satisfy the location strategy. If S_i satisfies user's location strategy, i.e. $< \vec{u_i}, \vec{v} >= 0$, sets $D_i = d_i + a$, else, sets $D_i = 1$. We assume $a = 3$ in this example. During the phase of collection, each S_i encrypts D_i by order-preserving encryption to obtain the encrypted ciphertext C_{D_i}. And then each S_i calculates $F^{C_{D_i}}(S_i^-)$. Due to the property of OPE, because $D_2 = 8$ is the largest value among D_i. Thus C_{D_2} is still the largest value among the ciphertext C_{D_i}. We use C_8 to represent C_{D_2} in this example. Next, at the phase of aggregation, for each $F^{C_{D_i}}(S_i^-)$, CSP computes the value of $F^{C_8}(S_i^-)$ through one-way chain. Then CSP computes the aggregated result $(C_8, S_2^+, \odot_i F^{C_8}(S_i^-))$ and sends the aggregated result to user. User verifies the correctness of the aggregated result reported by CSP and decrypts C_8 to obtain the plaintext of the maximum value at the specified location.

6) DISCUSSION OF $LBOA_{MAX}$

a: STATIC VS. DYNAMIC

In our scheme $LBOA_{Max}$, location-sensitive devices may be static or dynamic. If the location-sensitive devices are static, in other words, if the location of location-sensitive devices will not change after deploying, each S_i will only need to register once. If the location-sensitive devices are dynamic, at the beginning of each different process of $LBOA_{Max}$, the newly added LSDs should be registered and have certified ID. At the same time, the LSDs which are outdated should be logged out.

b: LOCATION STRATEGY

As shown in Figure 6(a), if we want to determine whether a point (x, y) is on the boundary of a circle C, we could split point (x, y) into a vector $\vec{u}$, then we distribute the circle into a vector $\vec{v}$, and then compute the inner product $< \vec{u}, \vec{v} >$. If $< \vec{u}, \vec{v} >= 0 \Rightarrow (x, y) \in C$. As shown in Figure 6(b), in a two-dimensional space, a particular point (x, y) can be uniquely determined by two tangent circles, if $< \vec{u}, \vec{v_1} >= 0$ and $< \vec{u}, \vec{v_2} >= 0$ are met at the same time, the point (x, y) is at the specified location.

c: ONE-WAY CHAIN

One-way chain is constructed by a series of one-way functions. There may be multiple aggregators in CSP executing aggregation process at the same time. Thus, the one-way function we adopted should have the property of homomorphism. Due to the reason that RSA has the property of homomorphism, we use RSA as the one-way function F in this study.

If the one-way function used in one-way chain is RSA, multiple values of one-way chain based on the same encryption key can be folded together by using modulo multiplication operation. We use symbol $\odot$ to denote the folding operation in this article. The folded value can be used to make more efficient communication and do verification efficiently.

LSD	CSP	User
ID_i (PK_i, SK_i)		(PK_u, SK_u) K_i
(x_i', y_i') K_i		GK

$a \in Z_p, f = Max$

$P = \{P_1, P_2, ..., P_n\}$

$Enc(GK, a \| P \| K_{OPE} \| epoch\#) \rightarrow C_{\text{Request}}$

$\{f = Max \| C_{\text{Request}}\}$ ←

$\{f = Max \| C_{\text{Request}}\}$ ←

$Dec(K_i, C_{\text{Request}}) \rightarrow a \| P \| K_{OPE} \| epoch\#$

$Enc(PK_U, Sig(SK_i, \vec{u}_i)) \rightarrow C_{u_i}$

if $\exists P_i \forall \vec{v}_k ((P_i \in P) \wedge (\vec{v}_k \in P_i) \rightarrow (\prod_{k=1}^{m} < \vec{u}_i, \vec{v}_k > = 0))$

sets $D_i = d_i + a$ else, sets $D_i = 1$

$Enc_{OPE}(K_{OPE}, D_i) \rightarrow C_{D_i}$

$S_i^+ = \{ID_i, C_{u_i}, MAC_{K_i}(C_{D_i} \| C_{u_i} \| epoch\#)\}$

$S_i^- = \{MAC_{K_i}(epoch\#)\}$

calculates $F^{C_{D_i}}(S_i^-)$

$\{C_{D_i}, S_i^+, F^{C_{D_i}}(S_i^-)\}$ →

compares $C_{D_i} \rightarrow C_{D_m}$

computes $F^{C_{D_m}}(S_i^-)$

aggregates, computes $\odot_i F^{C_{D_m}}(S_i^-)$

aggregated result $(C_{D_m}, S_i^+, \odot_i F^{C_{D_m}}(S_i^-))$

$\{C_{D_m}, S_m^+, \odot_i F^{C_{D_m}}(S_i^-)\}$ →

computes all individual S_i^-

computes $F^{C_{D_m}}(S_i^-)$

computes $\odot_{i-user} F^{C_{D_m}}(S_i^-)$

verifies $\odot_{i-user} F^{C_{D_m}}(S_i^-) = \odot_i F^{C_{D_m}}(S_i^-)$

$Dec_{OPE}(K_{OPE}, C_{D_m}) \rightarrow D_m$

$Dec(PK_m, Dec(SK_U, C_{U_m})) \rightarrow \vec{u}_m$

$d_m = D_m - a$

FIGURE 4. Scheme ***LBOA$_{Max}$***.

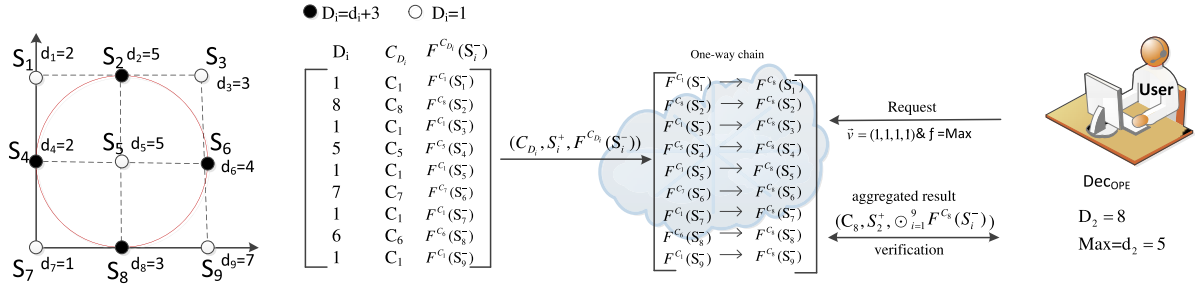

FIGURE 5. An example of ***LBOA$_{Max}$***.

B. LBOA$_{TOP-K}$

In this section, we do some extension to propose scheme $LBOA_{Top-k}$.

User requests to obtain the largest k values of data under location strategy in $LBOA_{Top-k}$ scheme. It is obvious that we can achieve this by executing $LBOA_{Max}$ scheme repeatedly for k times.

The first invocation of $LBOA_{Max}$ scheme will return the aggregated result $(C_{D_m}, S_m^+, \odot_i F^{C_{D_m}}(S_i^-))$ to user. The user could learn the maximum value d_m and the location of S_m

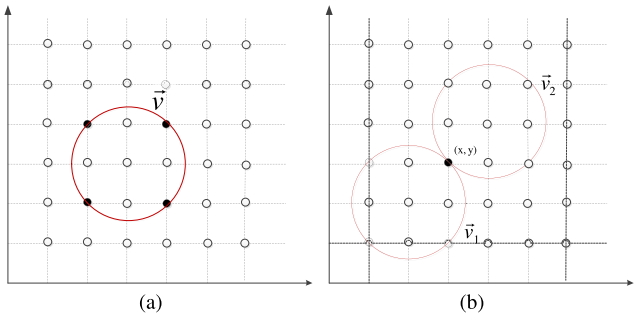

FIGURE 6. (a) Circle area. (b) Specified location.

who reporting the maximum value of sensory data. Then executing the second invocation of $LBOA_{Max}$ scheme again after excluding the location-sensitive device S_m. The second invocation will return the second largest value of data. And then, executing the third invocation of $LBOA_{Max}$ scheme again after excluding the location-sensitive device who reports the second largest value of data. In a similar fashion, the user can get top-k values of data whose location satisfy the location strategy after executing $LBOA_{Max}$ scheme repeatedly for k times.

C. $LBOA_{SUM}$

$LBOA_{Sum}$ scheme requests to obtain the summation value of data under user's location strategy. Similar to $LBOA_{Max}$ and $LBOA_{Top-k}$, scheme $LBOA_{Sum}$ also contains of five phases.

The initialization phase is completely identical to 5.A.1 with an additional initialization step: CSP has a certified public/private key pair that we represent as (PK_{CSP}, SK_{CSP}).

The request phase is identical to 5.A.2 except the operation function f in step (2) is $f = Sum$, and the step (4) is user utilizes group key GK to encryption a and the location strategy $\mathbb{P}$, i.e.,

$$Enc(GK, a\|\mathbb{P}\|epoch\#) \rightarrow C_{Request}.$$

In the phase of collection, Each S_i decrypts the broadcast message from the CSP to get the plaintexts of a and $\mathbb{P}$, i.e.,

$$Dec(K_i, C_{Request}) \rightarrow a\|\mathbb{P}\|epoch\#.$$

The second and the third steps are identical to 5.A.3(2) and 5.A.3(3). Then each S_i utlizes CSP's public key to encrypt D_i, i.e.,

$$Enc(PK_{CSP}, D_i) \rightarrow C_{D_i}.$$

Next each location-sensitive device S_i generates:

$$S_i^+ = \{ID_i, C_{u_i}, MAC_{K_i}(C_{D_i}||C_{u_i}||epoch\#)\},$$
$$S_i^- = \{MAC_{K_i}(epoch\#)\}.$$

Finally, Each S_i calculates $F^{D_i}(S_i^-)$.

During the phase of aggregation, the following steps are executed in sequence.

(1) Each S_i sends $(C_{D_i}, S_i^+, F^{D_i}(S_i^-))$ to CSP.

(2) CSP decrypts C_{D_i} to get the plaintext D_i, i.e.,

$$Dec(SK_{CSP}, C_{D_i}) \rightarrow D_i.$$

(3) CSP counts the number of $D_i = 1$, and use h to denote the total quantity.

(4) CSP compares the value of D_i to get the maximum value of D_i. Assuming the maximum value is D_m.

(5) CSP computes the summation of data whose location satisfies the location strategy, i.e.,

$$Sum_{CSP} = \Sigma_{i=1}^{n} D_i - h.$$

(6) CSP computes the value of $Top - (n - h)$ and $Top - (n - h + 1)$.

(7) CSP computes $F^{D_m}(S_i^-)$ for each S_i, and then CSP computes $\odot_i F^{D_m}(S_i^-)$, where

$$\odot_i F^{D_m}(S_i^-) = \prod_{i=1}^{n} (F^{D_m}(S_i^-)) mod(pq).$$

(8) CSP sends the aggregated result $(h, D_m, Top - (n - h), Top - (n - h + 1), Sum_{CSP}, \odot_i F^{D_m}(S_i^-))$ to the user.

After receiving the aggregated result sent by CSP, the user verifies the correctness of aggregated result. The specific steps of this phase are described as follows.

(1) User computes all individual S_i^-, computes $F^{D_m}(S_i^-)$ for each S_i, and then computes $\odot_{i-user} F^{D_m}(S_i^-)$, where

$$\odot_{i-user} F^{D_m}(S_i^-) = \prod_{i=1}^{n} (F^{D_m}(S_i^-)) mod(pq).$$

(2) If $\odot_{i-user} F^{D_m}(S_i^-)$ computed by the user is the same as $\odot_i F^{D_m}(S_i^-)$ in the aggregated result reported by CSP, then user verifies $Top - (n - h)$ and $Top - (n - h + 1)$ reported in the aggregated result.

(3) If the following three conditions are met at the same time, user believes that the aggregated result reported by CSP is correct. And then user computes $Sum = Sum_{CSP} - a(n - h)$. Sum denotes the summation value of data who meets user's location strategy.

- $\forall D_i (D_i \in Top(n - h) \rightarrow D_i \geq 2)$.
- The values in $Top - (n - h)$ is the same as the first $(n - h)$ values in $Top - (n - h + 1)$.
- The last value in $Top - (n - h + 1)$ is 1, in other words, the minimum value in $Top - (n - h + 1)$ is 1.

VI. SECURITY ANALYSIS

In this section, we analyze that our schemes satisfy all the security goal. We take $LBOA_{Max}$ as an example. The analysis of $LBOA_{Top-k}$ and $LBOA_{Sum}$ are similar to that of $LBOA_{Max}$.

A. AGGREGATION VERIFIABILITY

Our scheme $LBOA_{Max}$ can guarantee the verifiability of aggregation. In other words, our scheme can guarantee the aggregation process not to be tampered by CSP. We analyze it from the following two aspects.

1. In this part, we don't discuss the case that CSP maliciously delete or increase the nodes participating in the aggregation, we only consider that CSP false reports the aggregated result.

Our scheme can guarantees CSP reports the true Max, put another way, our scheme can prevent CSP reporting a value smaller or larger than the true Max. Let D_m denotes the true Max of reported values and D'_m denotes the result reported by the CSP. We use MAC_{Ki} to generate S_i^+ and S_i^- in the phase of collection. Due to the properties of MAC, S_i^+ and S_i^- only can be generated by S_i. Thus user can verify the validity of S_i^+ for a given epoch. Next, we analyze from the following two aspects.

① $D'_m > D_m$. If $D'_m > D_m$, the parameter S_m^+, which $S_m^+ = \{ID'_m, C'_{u_m}, MAC_{K_i}(C'_{D_m}||C'_{u_m}||epoch\#)\}$ must be generated by a location-sensitive device S'_m. So the adversary must forge the MAC successfully, which is impossible.

② $D'_m < D_m$. We construct one-way chain whose seed is S_i^- in the phase of collection. Each location-sensitive device S_i reports the value at position C_{D_i}, which we express by $F^{C_{D_i}}(S_i^-)$. The CSP rolls $F^{C_{D_i}}(S_i^-)$ forward for several times to obtain the value $F^{C_{D_m}}(S_i^-)$ at position C_{D_m} for each S_i during the phase of aggregation. Then CSP computes $\odot_i F^{D_m}(S_i^-)$. By doing these, if $D'_m < D_m$, it is impossible to utilize $F^{C_{D_m}}(S_i^-)$ to obtain $F^{C_{D'_m}}(S_i^-)$ since F is an one-way function and $F^{D_m}(S_i^-)$ is an one-way chain which can only roll forward but cannot roll backward. So the probability to calculate all the S_i's value of $F^{C_{D_m}}(S_i^-)$ can be negligible.

Therefore, our scheme is secure since the CSP can't abduct the user to accept an incorrect aggregated result by tampering with the aggregation process.

2. In this part, we mainly consider the case that CSP discards the LSD that reads the largest value among all the LSDs who satisfy the location strategy, or CSP adds an illegal LSD that reads a value larger than the true largest reading of legitimate LSDs in our scheme.

① Our scheme can guarantee CSP cannot delete or ignore the true largest value under location strategy. In the initilization phase, each S_i negotiates a symmetric key K_i with user. After LSD sending aggregated result $(C_{D_m}, S_m^+, \odot_i F^{C_{D_m}}(S_i^-))$ to the user, user computes all individaual $S_i^- = MAC_{K_i}(epoch\#)$ and folded all the $F^{C_{D_m}}(S_i^-)$ together to get the $\odot_{i-user} F^{C_{D_m}}(S_i^-)$. And then user compares $\odot_{i-user} F^{C_{D_m}}(S_i^-)$ with $\odot_i F^{C_{D_m}}(S_i^-)$ reported by CSP. In this way, if CSP ignores some legitimate value of LSD's data, $\odot_{i-user} F^{C_{D_m}}(S_i^-)$ computed by user will not equal to the $\odot_i F^{C_{D_m}}(S_i^-)$ reported by CSP. What's more, similar to our analysis above, due to the characteristic of one-way chain, if CSP discards the true largest value and regards a value smaller than it as the largest one, the probability to calculate all the S_i's value of $F^{C_{D_m}}(S_i^-)$ can be negligible.

② Our scheme can guarantee CSP will not add an illegal LSD which reads a value D'_m larger than the true largest value D_m. Assuming that the vector of the illegal LSD's location is $\vec{u'_m}$. After user receiving the aggregated result reported by the CSP, user will verify the validity of S_m^+ firstly. Similar to our analysis above, S_m^+ must generated by a legitimate LSD. But there are no legitimate LSD has the value of D'_m, so the adversary must forge the MAC. However the probability that adversary could forge the MAC successfully can be negligible. What's more, after user verifying the correctness of aggregated result, user will get the location (x'_m, y'_m) from $\vec{u'_m}$. (x'_m, y'_m) is illegal, so it is impossible to pass the verification of the user.

B. LOCATION PRIVACY

Location privacy means the location of LSD should only be known by the LSD itself and the totally trusted user. Thus, any other entity including other LSDs, the CSP and outside attacker cannot learn any knowledge about the location of LSD.

In our scheme $LBOA_{Max}$, we adopt a secure asymmetric encryption algorithm and signature algorithm. At the phase of collection, LSD utilizes S_i's secret key SK_i to sign $\vec{u_i}$, then LSD utilizes user's public key PK_u to encrypt $\vec{u_i}$, i.e. $Enc(PK_U, Sig(SK_i, \vec{u_i})) \rightarrow C_{u_i}$. Since the asymmetric encryption algorithm is semantically secure, it can against chosen ciphertext attack (CCA). The ciphertexts generated by public key encryption algorithm are indistinguishable. In other words, for any two given location plaintext $\vec{u_1}$ and $\vec{u_2}$, and their location ciphertext is C_{u_1} and C_{u_2} respectively, the probability that the adversary can distinguish C_{u_2} with C_{u_1} is $Pr(C_{u_1} \leftarrow Enc(PK_u, \vec{u_1}) : Attacker(PK_u, Enc(PK_u, \vec{u_2})) = C_{u_1}) \leq p(\lambda)$, where $p(\lambda)$ is negligible. This means that only the user who has secret key SK_u could decrypt the ciphertext C_{u_i}, any other third party could not learn any knowledge about the location plaintext $\vec{u_i}$. Since the signature algorithm is semantically secure which can against adaptive chosen message attack (CMA), signature is unforgeable. The probability for adversary who does not have the secret key SK_i to forge a signature of S_i correctly is negligible.

Therefore, the privacy of location-sensitive devices' location can be achieved in our scheme.

C. DATA CONFIDENTIALITY

Data confidentiality refers to the confidentiality of data d_i, which is the raw value of LSD's sensory data. Our schemes can ensure the data sensed by LSD will not be intercepted by external attacker. In $LBOA_{Max}$, the scheme also can ensure the data sensed by LSD will not be intercepted by CSP. In our scheme $LBOA_{Max}$, at the phase of collection, LSD judges whether its location satisfies user's location strategy $\mathbb{P} = \{P_1, P_2, \ldots, P_n\}$ or not. If $\exists P_i \forall \vec{v_k}((P_i \in \mathbb{P}) \bigwedge (\vec{v_k} \in P_i) \rightarrow (\Pi_{k=1}^m < \vec{u_i}, \vec{v_k} >= 0))$, in other words, if the location of S_i satisfies user's location strategy $\mathbb{P}$, S_i sets $D_i = d_i + a$, else, S_i sets $D_i = 1$. Thus, the confidentiality of d_i can be reduced to the confidentiality of D_i.

In the phase of collection, each S_i adopts order-preserving encryption to encrypt D_i, i.e. $Enc_{OPE}(K_{OPE}, D_i) \rightarrow C_{D_i}$. Since order-preserving encryption is a secure symmetric algorithm whose symmetric secret key K_{OPE} is shared

between user and LSD, only the totally trusted user and the LSD itself know the key K_{OPE} which is used in order-preserving encryption, any other entity including outside attackers and CSP know nothing about the key K_{OPE}. Only the LSD itself has D_i and the user could decrypt C_{D_i} with the symmetric key K_{OPE} to obtain plaintext D_i, i.e. $Dec_{OPE}(K_{OPE}, C_{D_i}) \rightarrow D_i$. As for other entities, the probability of obtaining D_i is negligible even if they have the ciphertext C_{D_i}.

What's more, in our scheme, at the phase of request, user picks a random $a \in \mathbb{Z}_p$, which meets the condition that for any value of S_i's sensory data d_i, $d_i + a \geq 2$ is satisfied. a is used as an offset. For each different session, the selection of a is random, that means for the same S_i, D_i is different in different sessions. Due to the properties of order-preserving encryption and offset, for the same plaintext d_i, different sessions generate different ciphertexts C_{D_i}. Adversary cannot distinguish the ciphertexts which are generated by the same plaintext in different sessions. For example, there is a plaintext value of data d_i, in one session, user picks $a_1 \in \mathbb{Z}_p$, thus $D_{i1} = d_i + a_1$, $Enc_{OPE}(K_{OPE}, D_{i1}) \rightarrow C_{D_{i1}}$, the ciphertext of data d_i in this session is $C_{D_{i1}}$. While in another session, user picks $a_2 \in \mathbb{Z}_p$, a_2 may not be equal to a_1. $D_{i2} = d_i + a_2$, $Enc_{OPE}(K_{OPE}, D_{i2}) \rightarrow C_{D_{i2}}$, thus the ciphertext of data d_i in this session is $C_{D_{i2}}$. Adversary cannot guess the plaintext d_i even he (or she) has the ciphertext $C_{D_{i1}}$ and $C_{D_{i2}}$ in different sessions.

Consequently, the confidentiality of the data can be guaranteed in our scheme.

D. LOCATION STRATEGY CONFIDENTIALITY

Location strategy confidentiality refers to that the user's location strategy $\mathbb{P} = \{P_1, P_2, \ldots, P_n\}$ can only be learned by legitimate LSDs and the user itself, any other entity including the untrusted CSP and outside attackers could learn nothing about user's location privacy.

In our scheme, at the phase of initialization, each legitimate S_i consults a symmetric key K_i with the user, user maintains a group key GK. At the phase of request, user adopts a secure asymmetric encryption algorithm, and utilizes group key GK to encrypt the location strategy $\mathbb{P}$ of the user, i.e., $Enc(GK, a \| \mathbb{P} \| K_{OPE} \| epoch\#) \rightarrow C_{Request}$. What's more, the location strategy $\mathbb{P}$ is encrypted by the group key GK. That means for any legitimate S_i who has consulted a symmetric key K_i with the user, it can decrypt the ciphertext $C_{Request}$ to get location strategy $\mathbb{P}$. However, for any other participants, including CSP, the probability to decrypt $C_{Request}$ which is encrypted by user's group key GK to get the plaintext of user's location strategy $\mathbb{P}$ can be negligible.

Therefore, the privacy of location strategy could be protected in our scheme.

E. DISCUSSION

***LBOA*$_{Top-k}$**: $LBOA_{Top-k}$ is achieved by executing $LBOA_{Max}$ repeatedly for k times. Thus, the security analysis of $LBOA_{Top-k}$ is similar to that of $LBOA_{Max}$.

***LBOA*$_{Sum}$**: Due to the purpose of $LBOA_{Sum}$ is to get the summation value of data, CSP will add all legal sensory data together to get the summation. Thus, the scheme $LBOA_{Sum}$ only prevents outside attacker but not CSP to get D_i. In the phase of collection, each S_i adopts a secure asymmetric encryption algorithm, S_i utilizes CSP's public key PK_{CSP} to encrypt D_i. i.e. $Enc(PK_{CSP}, D_i) \rightarrow C_{D_i}$. Only the CSP who has secret key SK_{CSP} could decrypt ciphertext to get D_i. Outside attacker cannot learn any knowledge about the D_i. Therefore, our scheme $LBOA_{Sum}$ can protect data confidentiality against outside attacker.

VII. PERFORMANCE ANALYSIS

A. COMPARISON WITH RELATED WORK

In this section, we compare our schemes with some related work including RCDA [27], SHIA [30], OA-WC [16] and Geo-SE [40]. The comparison results are shown in Table 2. "$\surd$" means satisfied, "$\times$" means dissatisfied and "-" means uninvolved. LB means location-based, DCO means data confidentiality against outside attacker, DCC means data confidentiality against CSP, LSC means location strategy privacy, LSDP means location-sensitive device privacy, AV means aggregation verifiability.

TABLE 2. Comparison with related work.

schemes	RCDA	SHIA	OA-WC	Geo-SE	LBOA		
					Max	Top-k	Sum
Max	$\times$	$\times$	$\surd$	$\times$	$\surd$	-	-
Top-k	$\times$	$\times$	$\surd$	$\times$	-	$\surd$	-
Sum	$\times$	$\surd$	$\times$	$\times$	-	-	$\surd$
LB	$\times$	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\surd$
DCO	$\surd$	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\surd$
DCC	$\surd$	$\times$	$\times$	$\surd$	$\surd$	$\surd$	$\times$
LSC	-	-	-	$\surd$	$\surd$	$\surd$	$\surd$
LSDP	-	-	-	$\surd$	$\surd$	$\surd$	$\surd$
AV	$\surd$	$\surd$	$\surd$	$\times$	$\surd$	$\surd$	$\surd$

From Table 2, it is obvious that RCDA [27] cannot support data aggregation functions. SHIA [30] cannot support Max and Top-k. It also cannot protect the confidentiality of data. OA-WC [16] cannot support Sum and cannot protect the confidentiality of data. Geo-SE [40] is location-based. It can protect the privacy of LSD, the confidentiality of data and location strategy. But this work does not focus on data aggregation. Our scheme $LBOA_{Max}$ and $LBOA_{Top-k}$ satisfy all the properties in Table 2, $LBOA_{Sum}$ satisfies all the properties except DCC.

B. COMPUTATION EVALUATION

In this section, we mainly focus on evaluation of computation overhead of our schemes. The experiments are implemented on a PC (CPU: Intel(R) Core(TM) i5-3470 CPU @ 3.20GHz 3.20GHz, RAM:8.00GB, OS: Windows 7), using jdk 1.8.0.

We take scheme $LBOA_{Max}$ as an example. The evaluation of $LBOA_{Top-k}$ is similar to $LBOA_{Max}$ because $LBOA_{Top-k}$ is achieved by executing $LBOA_{Max}$ repeatedly for k times. We do not repeat the evaluation of $LBOA_{Top-k}$ in this study.

We evaluate the computation overhead of LSD, CSP and user respectively. We adopt RSA-1024 and Hmac-MD5 in our implementation. We adopt the algorithm in [15] to realize order-preserving encryption. We assume there are 100 location-sensitive devices and the range of the value of LSD's sensory data is [0,100].

1) LSD

As shown in Figure 7, we evaluate the computation overhead of LSD in the phase of collection. For $LBOA_{Max}$ and $LBOA_{Top-k}$, we adopt two kinds of different ciphertext-space of OPE respectively. In one case, the ciphertext-space of OPE is $0 - 2^9$. And in the other case, the ciphertext-space of OPE is $0 - 2^{10}$. It is obvious that the bigger value of LSD's sensory data will lead to slightly higher computation cost. The ciphertext-space of OPE also has effect on computation overhead. When the ciphertext-space is larger, the computation cost of LSD is higher. The computation overhead of LSD in $LBOA_{Top-k}$ is similar to that of $LBOA_{Max}$. In brief, the processing time is between 0.2-0.3 second. It is acceptable for LSD whose calculation and storage power is limited. As for $LBOA_{Sum}$, the processing time is much less than 0.05 second. The processing time in $LBOA_{Sum}$ is very short because we does not adopt order-preserving encryption in it. In general, the computation overhead of LSD is practical for LSD who does not have very strong power.

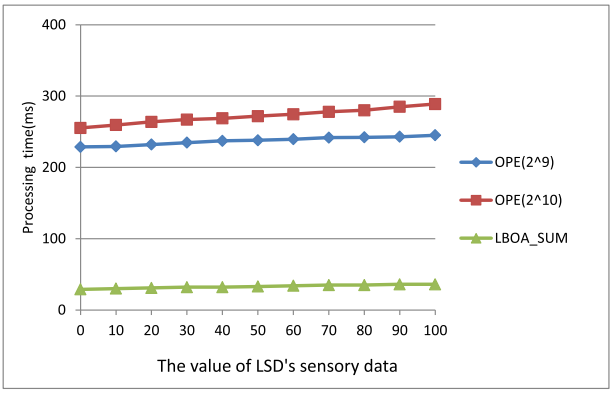

FIGURE 7. Computation overhead of LSD.

2) CSP

As shown in Figure 8, we evaluate the computation overhead of CSP in the phase of aggregation in protocol $LBOA_{Max}$. The CSP computes the folded value $\odot_i F^{C_{D_m}}(S_i^-)$ after receiving $F^{C_{D_i}}(S_i^-)$ from all the location-sensitive devices in the phase of aggregation. The ciphertext-space of OPE is $0 - 2^9$ in this part. We take four kinds of different data distribution strategy in our evaluation. Uniform denotes the values of sensory data are distributed uniformly at the range of [0,100]. High denotes the values of sensory data are concentrated on the range of [80,100]. Medium denotes the values of sensory data are concentrated on the range of [40,60]. Low denotes the values of sensory data are concentrated on the range of [0,20].

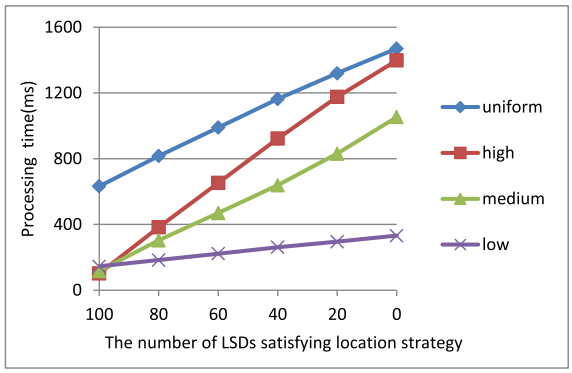

FIGURE 8. Computation overhead of CSP.

The algorithmic we adopted and the data distribution strategy we set are listed detailedly in Table 3.

From Figure 8, it is obvious that the number of LSDs satisfying location strategy influence the processing time. The fewer number of LSDs satisfying the location strategy, the higher computation cost of the CSP. The processing time is only about 0.1s when all the LSDs' location satisfy user's location strategy, while the processing time is 1.6s when nearly no LSD's location satisfies user's location strategy. The different distribution strategy also affect the computation overhead. It is obvious that the computation overhead is the highest when the values of LSD's sensory data are distributed uniformly, while the computation overhead is the lowest when the values of LSD's sensory data concentrated on low values. The processing time of high distribution is a little longer than that of medium distribution. In general, the processing time is within the scope of 0.1s and 1.5s, which is acceptable for CSP in practice.

3) USER

As shown in Figure 9, we evaluate the computation overhead of user in the phase of verification in protocol $LBOA_{Max}$. In the phase of verification, user computes each S_i^- and $F^{C_{D_m}}(S_i^-)$. Then, user folds all the $F^{C_{D_m}}(S_i^-)$ together to get the folded value $\odot_{i-user} F^{C_{D_m}}(S_i^-)$. If the aggregated result can be verified successfully, user decrypts C_{D_m} and C_{u_m}.

From Figure 9, the number of LSDs satisfying location strategy has little effect on the computation overhead of user. The distribution strategy of sensory data has much effect on the computation overhead of user. The processing time is about 1.5s, 2.2s, 2.5s and 2.5s when the data distribution strategy is low, medium, high and uniform respectively. The computation overhead is the highest when the data values are uniform distribution or high distribution. The computation overhead is the lowest when the data values are low distribution. When the data values are medium distribution, the computation overhead is higher than that of low distribution.

4) COMPUTATION OVERHEAD UNDER DIFFERENT OPE

As shown in Figure 10, the ciphertext-space of OPE has effect on the computation overhead. We adopt two kinds of

TABLE 3. Description of evaluation.

Function	Parameter
MAC	Hmac-MD5
OPE	Enc_{OPE} / Dec_{OPE} where ciphertext-space is $(0-2^9)$
one-way function F	RSA-1024
folded operation $\odot$	modulo multiplication
the range of sensory data	[0-100]
uniform	sensory data distributed in [0,100]
high	sensory data distributed in [80,100]
medium	sensory data distributed in [40,60]
low	sensory data distributed in [0,20]

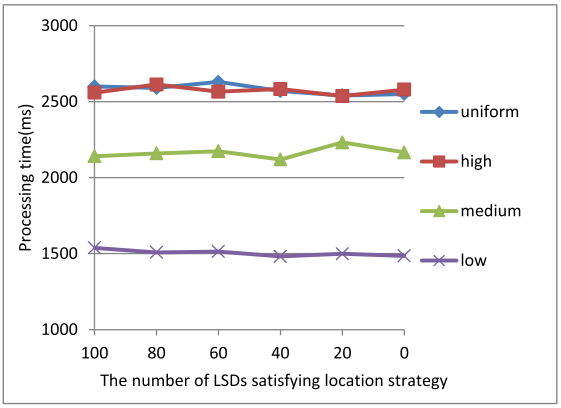

FIGURE 9. Computation overhead of user.

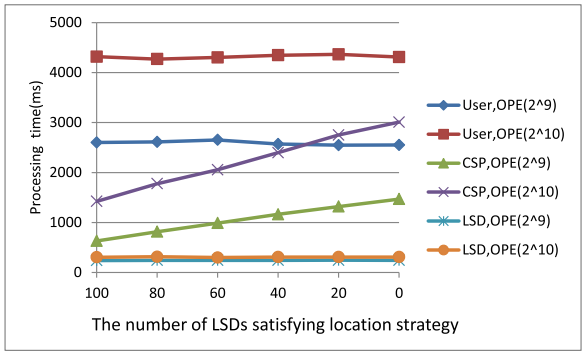

FIGURE 10. Computation overhead under different OPE.

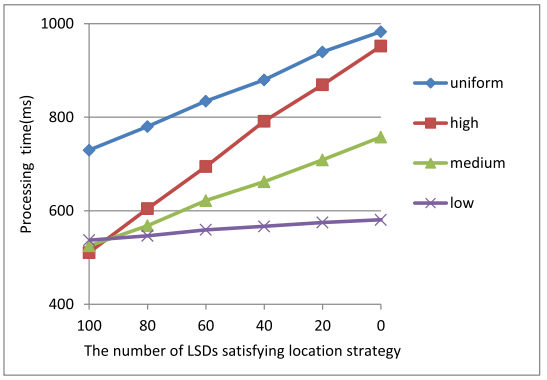

FIGURE 11. Computation overhead of aggregation in $LBOA_{Sum}$.

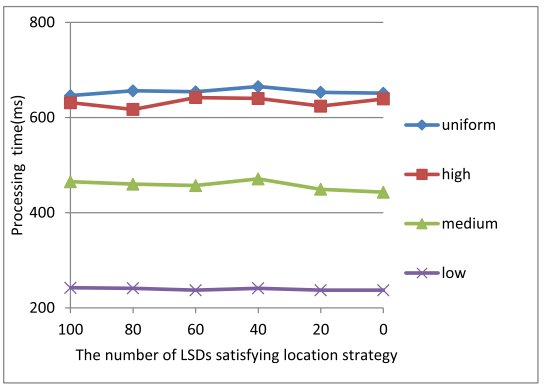

FIGURE 12. Computation overhead of verification in $LBOA_{Sum}$.

different ciphertext-space of OPE respectively. In one case, the ciphertext-space of OPE is $0-2^9$. And in the other case, the ciphertext-space of OPE is $0-2^{10}$. For user and CSP, if the ciphertext-space of OPE is larger, the computation overhead is higer. For user, the processing time is about 2.6s when the ciphertext-space of OPE is $0-2^9$, while the processing time is around 4.2s when the ciphertext-space of OPE is $0-2^{10}$. For CSP, the processing time is within the scope of 0.6s - 1.5s when the ciphertext-space of OPE is $0-2^9$, while the processing time is within the scope of 1.3s - 3s when the ciphertext-space of OPE is $0-2^{10}$. For LSD, the computation overhead under OPE (2^{10}) is slightly higher than OPE (2^9). What's more, the computation overhead of LSD is far less than the computation overhead of CSP and user, which is suitable in practice considering that LSD does not have strong capacity for computation.

5) COMPUTATION OVERHEAD IN $LBOA_{SUM}$

The computation overhead of aggregation in scheme $LBOA_{Sum}$ is shown in Figure 11. The computation overhead of verification in scheme $LBOA_{Sum}$ is shown in Figure 12. The computation overhead of $LBOA_{Sum}$ is much lower than that of $LBOA_{Max}$ and $LBOA_{Top-k}$ because S_i utilizes public key cryptography rather than order-preserving encryption to encrypt D_i in $LBOA_{Sum}$. From Figure 11, it is obvious that the processing time of CSP aggregation is less than 1s. From Figure 12, the processing time of user verification is less than 0.7s, which is acceptable in practice.

VIII. CONCLUSIONS

In this study, we proposed three novel schemes that can achieve secure outsourced aggregation based on data's location. We proposed $LBOA_{Max}$ to obtain the Max aggregated data first, and then we proposed $LBOA_{Top-k}$ and $LBOA_{Sum}$ to obtain the Top-k and Sum aggregated data respectively. Different from existing schemes, our schemes could realize secure aggregation based on location and could achieve location privacy protection, data confidentiality protection and location strategy confidentiality protection. Next we analyze the security of our schemes and the analysis results show that our schemes satisfy all the defined requirements. Finally, the experiment results show that our schemes are practical and feasible in IoT.

REFERENCES

[1] X. Li, J. Niu, M. Z. A. Bhuiyan, F. Wu, M. Karuppiah, and S. Kumari, "A robust ECC-based provable secure authentication protocol with privacy preserving for industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3599–3609, Aug. 2018.

[2] D. Kwon, M. R. Hodkiewicz, J. Fan, T. Shibutani, and M. G. Pecht, "IoT-based prognostics and systems health management for industrial applications," *IEEE Access*, vol. 4, pp. 3659–3670, 2016.

[3] E. Fernandes, J. Jung, and A. Prakash, "Security analysis of emerging smart home applications," in *Proc. IEEE Symp. Secur. Privacy (SP)*, San Jose, CA, USA, May 2016, pp. 636–654.

[4] D. Lu and T. Liu, "The application of IoT in medical system," in *Proc. Int. Symp. IT Med. Edu.*, vol. 1, 2011, pp. 272–275.

[5] T. Dimitriou, "Secure and scalable aggregation in the smart grid," in *Proc. IEEE Int. Conf. New Technol., Mpbility Secur. (NTMS)*, Dubai, United Arab Emirates, Mar./Apr. 2014, pp. 1–5.

[6] L. Hu and D. Evans, "Secure aggregation for wireless networks," in *Proc. Appl. Internet Workshops*, Orlando, FL, USA, 2003, pp. 384–391.

[7] H. Xiong, K.-K. R. Choo, and A. V. Vasilakos, "Revocable identity-based access control for big data with verifiable outsourced computing," *IEEE Trans. Big Data*, to be published.

[8] F. Deng, H. Xiong, Y. Wang, L. Peng, J. Geng, and Z. Qin, "Ciphertext-policy attribute-based signcryption with verifiable outsourced designcryption for sharing personal health records," *IEEE Access*, vol. 6, pp. 39473–39486, 2018.

[9] G. Xie, G. Zeng, R. Li, and K. Li, "Quantitative fault-tolerance for reliable workflows on heterogeneous IaaS clouds," *IEEE Trans. Cloud Comput.*, to be published. doi: 10.1109/TCC.2017.2780098.

[10] Y. Chen, G. Xie, and R. Li, "Reducing energy consumption with cost budget using available budget preassignment in heterogeneous cloud computing systems," *IEEE Access*, vol. 6, pp. 20572–20583, 2018.

[11] H. Xiong, H. Zhang, and J. Sun, "Attribute-based privacy-preserving data sharing for dynamic groups in cloud computing," *IEEE Syst. J.*, to be published.

[12] J. Zhang, J. Ma, C. Yang, and L. Yang, "Universally composable secure positioning in the bounded retrieval model," *Sci. China Inf. Sci.*, vol. 58, no. 11, pp. 1–15, 2015.

[13] T. Kwon, J. Lee, and J. Song, "Location-based pairwise key predistribution for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 11, pp. 5436–5442, Nov. 2009.

[14] B. Li, W. Wang, Q. Yin, H. Li, and R. Yang, "An energy-efficient geographic routing based on cooperative transmission in wireless sensor networks," *Sci. China Inf. Sci.*, vol. 56, no. 7, pp. 1–10, 2013.

[15] A. Boldyreva, N. Chenette, Y. Lee, and A. O'Neill, "Order-preserving symmetric encryption," in *Proc. Annu. Int. Conf. Theory Appl. Cryptographic. Techn.*, Cologne, Germany: Springer, 2009, pp. 224–241.

[16] S. Nath, H. Yu, and H. Chan, "Secure outsourced aggregation via one-way chains," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, Providence, RI, USA, 2009, pp. 31–44.

[17] H. Ö. Tan and I. Körpeoglu, "Power efficient data gathering and aggregation in wireless sensor networks," *ACM SIGMOD Rec.*, vol. 32, no. 4, pp. 66–71, 2003.

[18] R. Rajagopalan and P. K. Varshney, "Data-aggregation techniques in sensor networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 8, no. 4, pp. 48–63, 4th Quart., 2006.

[19] J.-Y. Chen, G. Pandurangan, and D. Xu, "Robust computation of aggregates in wireless sensor networks: Distributed randomized algorithms and analysis," *IEEE Trans. Parallel Distrib. Syst.*, vol. 17, no. 9, pp. 987–1000, Sep. 2006.

[20] R. Hekmat and P. Van Mieghem, "Connectivity in wireless ad-hoc networks with a log-normal radio model," *Mobile Netw. Appl.*, vol. 11, no. 3, pp. 351–360, 2006.

[21] C. W. Yu and L. H. Yen, "Computing subgraph probability of random geometric graphs: Quantitative analyses of wireless ad hoc networks," in *Proc. Int. Conf. Formal Techn. Networked Distrib. Syst.* Berlin, Germany: Springer, 2005, pp. 458–472.

[22] H. J. Lee, A. Cerpa, and P. Levis, "Improving wireless simulation through noise modeling," in *Proc. ACM 6th Int. Conf. Inf. Process. Sensor Netw.*, Cambridge, MA, USA, 2007, pp. 21–30.

[23] Y. Yu, V. K. Prasanna, and B. Krishnamachari, "Energy minimization for real-time data gathering in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3087–3096, Nov. 2006.

[24] T. Jung, X. Mao, X.-Y. Li, S.-J. Tang, W. Gong, and L. Zhang, "Privacy-preserving data aggregation without secure channel: Multivariate polynomial evaluation," in *Proc. IEEE INFOCOM*, Turin, Italy, Apr. 2013, pp. 2634–2642.

[25] Y.-H. Lin, S.-Y. Chang, and H.-M. Sun, "CDAMA: Concealed data aggregation scheme for multiple applications in wireless sensor networks," *IEEE Trans. Knowl. Data Eng.*, vol. 3, no. 3, pp. 1471–1483, Jul. 2014.

[26] L. Zhu, Z. Yang, M. Li, and D. Lium, "An efficient data aggregation protocol concentrated on data integrity in wireless sensor networks," *Int. J. Distrib. Sensor Netw.*, vol. 9, no. 5, 2013, Art. no. 256852.

[27] C.-M. Chen, Y.-H. Lin, Y.-C. Lin, and H.-M. Sun, "RCDA: Recoverable concealed data aggregation for data integrity in wireless sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 4, pp. 727–734, Apr. 2012.

[28] X. Li, S. Liu, F. Wu, S. Kumari, and J. J. P. C. Rodrigues, "Privacy preserving data aggregation scheme for mobile edge computing assisted IoT applications," *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018.22018.2874473.

[29] B. Przydatek, D. Song, and A. Perrig, "SIA: Secure information aggregation in sensor networks," in *Proc. ACM 1st Int. Conf. Embedded Netw. Sensor Syst.*, Los Angeles, CA, USA, 2003, pp. 255–265.

[30] H. Chan, A. Perrig, and D. Song, "Secure hierarchical in-network aggregation in sensor networks," in *Proc. 13th ACM Conf. Comput. Commun. Secur.*, Alexandria, VA, USA, 2006, pp. 278–287.

[31] A. Vora and M. Nesterenko, "Secure location verification using radio broadcast," *IEEE Trans. Dependable Secure Comput.*, vol. 3, no. 4, pp. 377–385, Oct./Dec. 2006.

[32] N. Sastry, U. Shankar, and D. Wagner, "Secure verification of location claims," in *Proc. 2nd ACM Workshop Wireless Secur.*, New York, NY, USA, 2003, pp. 1–10.

[33] S. Čapkun, M. Čagalj, and M. Srivastava, "Secure localization with hidden and mobile base stations," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006, pp. 1–12.

[34] S. Sciancalepore, G. Oligeri, and R. Di Pietro, "Shooting to the stars: Secure location verification via meteor burst communications," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Beijing, China, May/Jun. 2018, pp. 1–9.

[35] N. Chandran, V. Goyal, R. Moriarty, and R. Ostrovsky, "Position based cryptography," in *Proc. Int. Cryptol. Conf. Adv. Cryptol. (CRYPTO)*, Cologne, Germany, 2009, pp. 391–407.

[36] J. Zhang, N. Lu, J. Ma, and C. Yang, "Universally composable secure geographic area verification without pre-shared secret," *Sci. China Inf. Sci.*, vol. 62, Mar. 2019, Art. no. 32113.

[37] J. Zhang, F. Du, J. Ma, and C. Yang, "Position based key exchange: Definitions and implementations," *J. Commun. Inf. Netw.*, vol. 1, no. 4, pp. 33–43, 2016.

[38] Y. Ji, J. Zhang, J. Ma, C. Yang, and X. Yao, "BMPLS: Blockchain-based multi-level privacy-preserving location sharing scheme for telecare medical information systems," *J. Med. Syst.*, vol. 42, no. 8, p. 147, Jun. 2018.

[39] Q. Gao, J. Zhang, J. Ma, C. Yang, J. Guo, and Y. Miao, "LIP-PA: A logistics information privacy protection scheme with position and attribute-based access control on mobile devices," *Wireless Commun. Mobile Comput.*, vol. 2018, Jul. 2018, Art. no. 9436120. doi: 10.1155/2018/9436120.

[40] B. Wang, M. Li, and H. Wang, "Geometric range search on encrypted spatial data," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 4, pp. 704–719, Apr. 2016.
[41] J. Zhang, J. Ma, and S. Moon, "Universally composable one-time signature and broadcast authentication," *Sci. China Inf. Sci.*, vol. 53, no. 3, pp. 567–580, 2010.
[42] N.-S. Jho, J. Y. Hwang, J. H. Cheon, M.-H. Kim, D. H. Lee, and E. S. Yoo, "One-way chain based broadcast encryption schemes," in *Proc. Annu. Int. Conf. Theory Appl. Cryptograph. Techn.*, Berlin, Germany: Springer, 2005, pp. 559–574.

JUNWEI ZHANG received the B.S., M.S., and Ph.D. degrees in computer science and technology from Xidian University, China, in 2004, 2006, and 2010, respectively, where he is currently an Associate Professor with the Department of Cyber Engineering. His research interests include cryptography and information security, especially in provable security, and data location security in cloud computing.

YUE ZONG received the B.S. degree in computer science and technology from Xidian University, China, in 2017, where she is currently pursuing the master's degree with the Department of Cyber Engineering. Her research interests include cryptography and information security, especially data location security in cloud computing.

CHAO YANG received the B.S., M.S., and Ph.D. degrees in cryptograpy from Xidian University, China, in 2002, 2004, and 2008, respectively. He is currently a Professor with the Department of Cyber Engineering, Xidian University, China. His research interests include mobile intelligent computing security, AI, and Big data-based cyberspace security.

YINBIN MIAO received the B.E. degree from the Department of Telecommunication Engineering, Jilin University, Changchun, China, in 2011, and Ph.D. degree from the Department of Telecommunication Engineering, Xidian University, Xi'an, China, in 2016, where he is currently a Lecturer with the Department of Cyber Engineering. He holds a postdoctoral position with Nanyang Technological University, from 2018 to 2019. His research interests include information security and applied cryptography.

JINGJING GUO received the M.Sc. and Ph.D. degrees in computer science from Xidian University, Xi'an, China, in 2012 and 2015, respectively, where she is currently a Lecturer with the Department of Cyber Engineering. Her research interests include trust management, access control, social networks, and wireless network security.