Hindawi
Security and Communication Networks
Volume 2021, Article ID 6654539, 11 pages
https://doi.org/10.1155/2021/6654539

Research Article

Verifiable Location-Encrypted Spatial Aggregation Computing for Mobile Crowd Sensing

Kun Niu,[1,2] **Changgen Peng**,[1,2] **Weijie Tan**,[1] **Zhou Zhou**,[1,2] **and Yi Xu**[1]

[1]*College of Computer Science and Technology, State Key Laboratory of Public Big Data, Guizhou University, Guiyang 550025, China*
[2]*Institute of Cryptography and Data Security, Guizhou University, Guiyang 550025, China*

Correspondence should be addressed to Changgen Peng; cgpeng@gzu.edu.cn

Received 30 December 2020; Revised 11 March 2021; Accepted 11 April 2021; Published 28 April 2021

Academic Editor: Ximeng Liu

Benefiting from the development of smart urban computing, the mobile crowd sensing (MCS) network has emerged as momentous communication technology to sense and collect data. The users upload data for specific sensing tasks, and the server completes the aggregation analysis and submits to the sensing platform. However, users' privacy may be disclosed, and aggregate results may be unreliable. Those are challenges in the trust computation and privacy protection, especially for sensitive data aggregation with spatial information. To address these problems, a verifiable location-encrypted spatial aggregation computing (LeSAC) scheme is proposed for MCS privacy protection. In order to solve the spatial domain distributed user ciphertext computing, firstly, we propose an enhanced-distance-based interpolation calculation scheme, which participates in delegate evaluator based on Paillier homomorphic encryption. Then, we use aggregation signature of the sensing data to ensure the integrity and security of the data. In addition, security analysis indicates that the LeSAC can achieve the IND-CPA indistinguishability semantic security. The efficiency analysis and simulation results demonstrate the communication and computation overhead of the LeSAC. Meanwhile, we use the real environment sensing data sets to verify availability of proposed scheme, and the loss of accuracy (global RMSE) is only less than 5%, which can meet the application requirements.

1. Introduction

The mobile crowd sensing (MCS) can carry out large-scale social sensing tasks with spatiotemporal attributes by distributed, multiuser collaborative sensing network mode. It is widely used in various fields of urban computing [1]. The truth discovery of MCS has recently been received wide attention. It refers to reliability collecting users' sensing data and aggregating and estimating the unknown data points (regions) accurately [2]. So, we can find reliable information among uneven quality of data collected from MCS network. Further, it has been extensively studied in the field of plaintext, but truth discovery remains largely underexplored in privacy protection MCS.

There are two types of collaborative methods to collect data for sensing users, namely, participant-sensing and opportunity-sensing. They provide strong data support for urban computing, such as environmental monitoring, traffic flow monitoring, and other sensing fields [3, 4]. The Noise Pollution Maps [5] system captures noise information by the participants' smartphones and generates Noise Maps. The GasMobile system [6] has also built a participatory mobile sensing system for air quality monitoring, which gathers data through the terminal equipment and analyzes and produces the pollutant concentration map. In these applications, the computing tasks require data providers to upload multidimensional sensing data and also need to collect the user's location information to participate in the calculation; the control of privacy information is transferred to the Internet service provider, which cannot guarantee the security of the user's privacy data, especially the security of user data in cloud aggregate computing. According to the survey of Crowd Research Partners in recent years, the security problems of cloud computing mainly focus on data

loss and leakage, data privacy, and data confidentiality [7], for example, publishing location privacy of data providers in sensing tasks, sensitive semantic information of data uploading, and malicious privacy association inference. Therefore, in data aggregation analysis, ensuring the privacy security of users, especially for many spatial-relation-based applications with high real-time requirements in sensing computing, is still a hot issue for scholars [8, 9].

Aggregate computation of spatial data based on MCS refers to the following: mobile terminal data providers upload the discrete insensitive data, and the platform collects and processes complex aggregation analysis, such as average calculation and cluster analysis. The aggregate computation can also combine with geography interpolation analysis, get the unknown value of sample point, and realize the overall fitting of regional, such as air quality evaluation. It is worth noting that users do not care about the uploading of feature data in spatial sensing, and there is more concern about the disclosure of personal spatiotemporal attribute privacy.

The traditional client-based location privacy protection method, such as anonymity, generalization, perturbation, and difference privacy, can meet the basic privacy protection needs of users. At the same time, these methods will reduce data accuracy and data availability, leading to deviation of the location-participation aggregated calculation results and lower data service quality of MCS. At present, most researches about spatial sensing privacy protection mainly focus on decentralized location privacy protection schemes, which fail to guarantee the application requirements of sensing data aggregation analysis [10, 11]. Therefore, we focus on two main problems: one is how to carry out ciphertext computing in spatial sensing computing service more efficiently, and the other is how to realize MCS architecture of homomorphic encryption and decryption operation without loss of spatial positioning accuracy. The mainstream method is to use cryptosystem to encrypt user data uploaded to the cloud nowadays [12, 13].

Realizing efficiently the safety of data processing is also the problem that attracted the biggest attention of many cloud service providers [14]. On the premise of protecting the privacy of user data, with the help of cloud computing service to aggregate encrypted data, the platform gets the correct result by decrypting and does not reveal any clear information in the process. However, simple encryption can greatly increase computational costs. Due to the loss of ciphertext data structure stored in the cloud, we need to construct a security solution to perform model calculation, analysis, and other processing of encrypted data without restoring plaintext, so that the cloud service platform can provide efficient services for crowd-sensing platform on the premise of protecting user data privacy [15].

As discussed above, secure data aggregation is the key problem in sensor networks application [16]. However, many ciphertext outsourcing computing schemes for aggregate data mostly satisfy simple algebraic operations and are not suitable for discrete data aggregation based on spatial relations. Moreover, the existing secure communication schemes are still facing security attacks and cannot guarantee all security requirements. In particular, most of the existing schemes are based on an ideal security model.

In order to guarantee the confidentiality of spatial information exchange under MCS architecture, this paper proposes a location-encrypted spatial aggregation computing (LeSAC) scheme. Assuming that the server is untrusted and the platform and the user are semitrusted, we study the public-key-based user information encryption, which is converted into ciphertext for homomorphic computation. In addition, on the basis of satisfying the location privacy protection, the MCS ciphertext computing framework uses the verifiable aggregation signature technology to realize the efficient aggregation of distributed network users' signatures, which ensures the anonymity of computation in the data communication process and prevents illegal modification of data. The main contributions of this paper are summarized as follows:

(i) For secure sharing and ciphertext aggregation of spatial information in MCS, we propose enhanced distance-based linear interpolation method. Based on Paillier homomorphism cryptography mechanism, we design a secure square Euclidean distance calculation protocol. Further, we design a location-encrypted spatial aggregation computing (LeSAC) protocol, which realizes ciphertext computing of crowd sensing platform without disclosing the location privacy.

(ii) In order to ensure computational anonymity during data flow, we implement a verifiable aggregation signature algorithm in the data transmission, and aggregating efficiently digital signatures of distributed network users and realizing multiparty data verify ensuring data security.

(iii) We discuss the indiscriminability of the protocol based on IND-CPA security analysis and analyze the communication complexity and the communication efficiency of the proposed protocol with increasing number of participants. We use the real environment pollution terminal sensing data to verify the data reconstruction accuracy. The results show that the region fitting effect of the test data and the real data is basically consistent. Meanwhile, the loss of reconstruction accuracy (global RMSE) is small.

This paper is organized as follows: Related works of security aggregate computation of MCS are presented in Section 2. Spatial ciphertext aggregation computing scheme and system architecture are discussed in Section 3. The secure computing protocol and a verifiable aggregating signature algorithm are proposed in Sections 4 and 5. Security analysis and the simulation results are shown in Section 6. The conclusion is drawn in Section 7.

2. Related Work

In order to realize the security spatial data aggregate computation of MCS, most existing researches focus on two

aspects: one is the security scheme of user sensing terminal and communication, for satisfying both data privacy and utility, such as localized differential privacy and identity authentication [17]. Wang and Sinnott [18] proposed a trajectories private publishing model with differential privacy, developed a private reference system for calibrating separate users trajectories, and constructed the enhanced noise prefix trees to publish data privately to ensure the accuracy and utility of the sensing data based on user trajectories. Tao et al. [19] proposed a hybrid authentication architecture by combining public key infrastructure (PKI) and combined public key (CPK); the users' security requirements of MCS system can be realized. He et al. [20] optimized the random noise adding mechanism based on the optimal distribution estimation algorithm in the network distributed computing architecture; the privacy of data can be guaranteed so as to realize the secure information exchange. Although it has the advantage of protecting privacy, the noise disturbance at the client side will seriously affect the accuracy of data aggregation calculation.

The other is the computational security for the sensing data aggregation and privacy-preserving outsourced computation [21]. Research methods include secure multiparty computing, homomorphic cryptography, secret sharing, and hybrid schemes [22]. Liu et al. [23] designed the security of square Euclidean distance and safety comparison and other lightweight building blocks and finally achieved the KNN classification algorithm based on spatial relationship in the cloud environment. Catak et al. [24] built a secure distance measurement method based on Paillier homomorphic password protocol and implemented a variety of clustering learning models for privacy protection. Deepak and Chandrasekaran [25] constructed a distributed data aggregation scheme for smart grid, which is based on the additive homomorphism of elliptic curve cryptosystem. In addition, combined with edge computing model, related research also realized data encryption aggregation based on spatial relationship. For example, based on mobile edge computing architecture and Paillier homomorphic encryption scheme, a distance-based secure location computing protocol is designed in [26], and base station is regarded as a semihonest participant to ensure the privacy of base station location information. Wu et al. [27] implemented a privacy-aware task assignment and statistical data aggregation scheme by fog node assistance, which is based on bilinear pair and homomorphic encryption. Similarly, Liu et al. [28] constructed a new cryptographic primitive to allow different providers data's outsource of the cloud server for secure storage and processing. These scholars used different cryptographic schemes to realize the secure computation of related applications and algorithms, but it is not suitable for the application of spatial data aggregation in MCS.

At present, scholars continue to deepen the research and exploration of homomorphic encryption system. Ciphertext computing ensures that users' private data can participate in the big data analysis in the cloud environment, which is very helpful for data mining and the secure implementation of machine learning algorithm [29]. In practical cloud environment applications, there are still efficiency problems in multiuser aggregation analysis and homomorphism scheme construction of multiple operations.

3. Problem Description and System Solution

This section discusses a hybrid scheme combining Paillier encryption mechanism [30] and spatial data homomorphism computation. The users upload the encrypted plaintext messages, the server performs homomorphism calculation on the ciphertext, and then the platform decrypts the results in order to protect the privacy of users. Paillier encryption mechanism only works on integer values, but most of real data sets contain continuous values. Therefore, for the real input numbers in the protocol, it is an obstacle in algorithm application. We map floating point coordinate data to discrete kilometer grid data, which has no influence on the spatial data analysis, so it does not affect the calculation efficiency and the validity of the proposed algorithm. For the standardization of academic papers, Table 1 lists some notations and their definitions in the research process of this paper.

3.1. Problem Description. Compound operation based on location data is widely used in information fitting and spatial analysis, especially in urban computing and spatial crowdsensing. The discrete sensing data of user can be used to achieve the overall regional data evaluation fitting . For example, air quality monitoring and road congestion analysis. As an important method in the regional information aggregation of MCS system, spatial interpolation can achieve the overall regional data evaluation fitting through user-provided discrete point data sampling fitting, as shown in Figure 1. The existing work studies the data reconstruction of the imperceptibility area under plaintext environment based on the limited perception data. Because of the privacy of edge data, how to use the sensing data provided by network users to calculate the value of unknown areas without exposing the user's personal information is a key problem in MCS. Therefore, the scheme design needs to consider four key factors:

(1) How to implement encrypted transmission.

(2) How to implement multiuser aggregated data ciphertext computation based on spatial relations.

(3) How to implement sensing data integrity verification.

(4) How to improve the efficiency of homomorphic aggregation operation and reduce computation complexity and storage space from the privacy protection perspective.

3.2. Proposed Scheme. The architecture diagram of the ciphertext-based privacy protection computing scheme for MCS system is shown in Figure 2. The cloud server receives the encrypted information of network node users and aggregates and sends this information to the task publisher. Consider an application scenario where a client with limited computing power wants to compute multinomial

Table 1: Notations.

Notations	Definition
Pk/Sk	Public key/private key in Paillier
$[m]_{\text{Enc}}$	Encryption algorithm of plaintext m
$[c]_{\text{Dec}}$	Decryption algorithm of ciphertext c
(x_i, y_i)	The location privacy information of data provider Mi
Z_i	The sensing data of data provider M_i
(X_j, Y_j)	The location information of requirement calculation point M_0
$\widehat{Z}_{j_{\text{Unknown}}}$	The value of requirement calculation point
S-SED	Secure square Euclidean distance calculation protocol
LeSAC	Location-encrypted spatial aggregation computing protocol

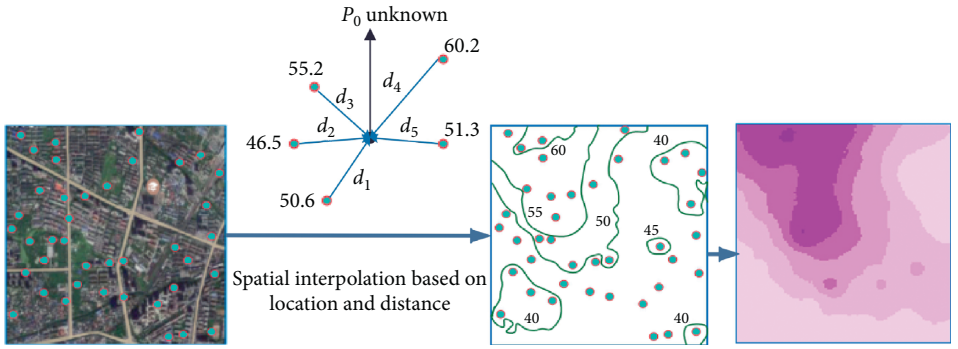

Figure 1: Discrete user data upload, aggregation, fitting, and regional evaluation.

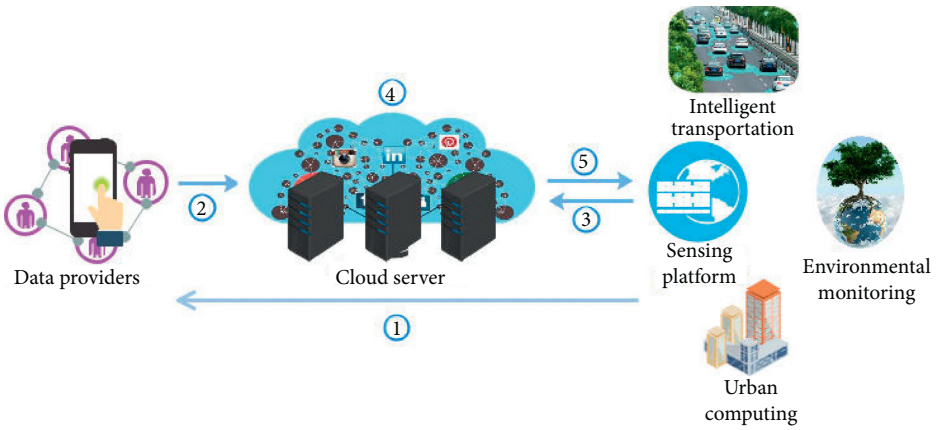

Figure 2: A location-encrypted verifiable spatial aggregation computing architecture of MCS. (1) Publish perception task. (2) Encryption and signature of the distributed upload data. (3) Issue computational request (functions, models, and parameters). (4) Homomorphic ciphertext computation and aggregation of signatures. (5) Decryption download and verification.

$f(m_1, \ldots, m_i)$ about sensitive messages $(m_1, \ldots, m_i)$. Then homomorphic encryption is a good way to delegate this computation; it can outsource heavy computing from the client to a server with sufficient computing power. Due to the particularity of MCS architecture, the outsourcing calculation based on the perceived user location does not care about the confidentiality of published information and only implements the aggregation homomorphism calculation based on the user location and distance. The specific flow is described as follows:

Step 1: certificate distribution agency generates the public and private keys of the encryption algorithm, and the generated public key is distributed to each user (data owner) of the network node. Data provider Mi collects data $[(x_1, y_1, Z_1), (x_2, y_2, Z_2), \ldots, (x_i, y_i, Z_i)]$, where Z_i is user sensing data. The data provider selects a random number $r_i \in Z_N$ and then uses the public key Pk to encrypt location coordinates and upload the encrypted coordinates and sensing data to the cloud computing center. Meanwhile, each data provider generates a signature σ_i.

Step 2: after receiving a computing request of $\widehat{Z}_{j_{\text{Unknown}}}$, the cloud platform traverses to find neighbor points through the Geohash index mechanism. A homomorphic secure Euclidean distance calculation protocol based on weighted perfect plane protocol to obtain D_{ij}

executes the enhanced spatial interpolation aggregation operation. This step also completes the aggregation of user signatures $\prod_{i=1}^{k} \sigma_i$.

Step 3: the cloud server completes the aggregation calculation of the value matrix of Z_j in the specific spatial region and then returns the spatial aggregation calculation results to the task publishing platform. The platform can verify the correctness of the data through signature verification function.

4. Secure Spatial Information Aggregation Computing Protocol

For the secure spatial information aggregation scheme under the MCS architecture, we first propose an enhanced spatial information aggregation algorithm and then construct an important spatial data ciphertext security computing protocol based on Paillier password system. Because the weight setting needs to realize the distance calculation of distributed users, we propose a secure square Euclidean distance calculation algorithm.

4.1. An Enhanced Distance-Based Interpolation Method for Spatial Data Aggregation. In order to reduce the computing energy consumption of the spatial sensing network, the scheme introduces bilinear interpolation method to enhance the weight value of the spatial. First of all, the sensing network randomly selects users A, B, and C, and the coordinate (x_A, y_A), (x_B, y_B), (x_C, y_C) is the secret information. From Euclidean distance formula, the calculation is as follows:

$$\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 = d_{ij}^2. \tag{1}$$

Considering the enhanced distance weight, if the sampling value of the near point makes a greater contribution to the unknown point, then calculate each user point contributed weight W_i for the unknown point value. Set $D_{ij} = d_{ij}^2$. According to the sensing data Z_i from the users known coordinate information, the unknown fitting value $\widehat{Z}_{j_{\text{Unknown}}}$ of space point (x_j, y_j) is inferred.

$$\begin{aligned} w_{ij} &= \frac{\left(1/D_{ij}\right)}{\sum_{i=1}^{n}\left(1/D_{ij}\right)}, \\ \widehat{Z}_{j_{\text{Unknown}}} &= \sum_{i=1}^{n} w_{ij} z_i. \end{aligned} \tag{2}$$

4.2. Retrieval of Private Information Based on Geohash Encoding. In order to compress and anonymously publish the high-dimensional data point set, we build a Geohash binary coding based on Geohash index, which can reduce the algorithm complexity of position traverse. The original data map to the corresponding buckets by the private hash retrieval method, and the data of the nearest location fall into the same block. Geohash represents a rectangular region. Participants can publish codes without exposing their exact coordinates. This helps users to protect their privacy. Therefore, in the neighboring point search, the unknown points (x_j, y_j) carry on the hash operation and complete the extraction of nearest neighbor point positions in the corresponding block to divide the large collections of distributed sensing users neighborhood calculation into the small data set.

4.3. Secure Square Euclidean Distance (S-SED) Calculation Protocol. The core of the proposed privacy protection scheme is the collaborative distance weight calculation. Therefore, we define a secure square computing protocol to realize homomorphic distance calculation. In the two-dimensional network space, the data provider uses the distributed public key to encrypt position coordinates $[x]_{\text{Enc}}$ and $[y]_{\text{Enc}}$. The goal of the perfect plane protocol is to calculate the perfect square expression $[(x - x_j)^2 + (y - y_j)^2]_{\text{Enc}}$ under the public key encryption; then,

$$\left([x]_{\text{Enc}} - x_j\right)^2 = \left([x]_{\text{Enc}}\right)^2 - 2x_j[x]_{\text{Enc}} + x_j^2. \tag{3}$$

For convenience, A means $([x]_{\text{Enc}})^2$ and B means $([x]_{\text{Enc}})$; according to the Paillier homomorphism property, $\forall m_1, m_2 \in \mathbb{Z}_n, k \in \mathbb{N}$, satisfying

$$\begin{aligned} \text{Dec}\left(\text{Enc}\left(m_1\right)\text{Enc}\left(m_2\right) \bmod n^2\right) &= m_1 + m_2 \bmod n, \\ \text{Dec}\left(\text{Enc}\left(m\right)^k \bmod n^2\right) &= km \bmod n. \end{aligned} \tag{4}$$

Then, under public key encryption, the cloud platform performs the following calculation:

$$A^* \left[B^{2x_j}\right]^{N-1} + x_j^2. \tag{5}$$

4.4. Location-Encrypted Spatial Aggregation Computing (LeSAC) Protocol. According to the system architecture shown in Figure 2, coordinate values are mapped one by one to the grid space domain in order to specify the coordinate information to the integer domain in the spatial sensing network. All data providers use the public key to encrypt the position coordinates. After the encryption is completed, each user ID, location information, and perception data are uploaded to the cloud server.

Finally, the secure aggregation calculation protocol in Algorithm 1 is as follows.

5. The Verifiable LeSAC Scheme Based on Aggregate Signatures

As mentioned above, location-concealed ciphertext aggregation computing protocol based on Paillier cryptosystem is designed and implemented, which can satisfy outsourced aggregation operation under the condition of invisible spatial information. In order to achieve the verifiable security of data outsourcing computing, based on the certificateless aggregate signature scheme proposed in [31], we realize the spatial data aggregation scheme satisfying privacy protection and results verifiable under the MCS architecture.

Input: Common input
(i) Paillier Encryption (Gen, Enc, Dec). Data provider M_i 's Pubic Key pubKey = n.
(ii) Data provider M_i 's sensory data *rec Da* ta_i.
Private input
(i) Data provider M_i 's location (x_i, y_i), the request point location (x_j, y_j).
(ii) The platform secretly holds the private key priKey = (λ, μ).
Output: The aggregate value of the request location point $\widehat{Z}_{j_{\text{unknown}}}$
(1) Geohash Boolean encod e(latitude, longitude) //Block partitioning indexes building
(2) For $j = 1$; $i = 0$; $j++$; $i++$
(3) do search neighbor regions of Z_j (left, right, up, down);
(4) until $(i = k)$;//k is the threshold of the number of nerghbor points
(5) EndFor;
(6) For j = 1 to finally
(7) For i = initial to k
(8) $[x_i]_{\text{Enc}} = (1+N)^{xi} \cdot r^N \bmod N^2$
(9) $[y_i]_{\text{Enc}} = (1+N)^{yi} \cdot r^N \bmod N^2$
(10) update add $([X_i]_{\text{Enc}}, [Y_i]_{\text{Enc}}, \text{rec}_{\text{Data}_i})$ for each $\widehat{Z}_{j_{\text{unknown}}}$
(11) $A^*[B^{2X_j}]^{N-1} + x_j^2 = ([x]_{\text{Enc}} - x_0)^2$
(12) $A^*[B^{2Y_j}]^{N-1} + y_j^2 = ([y]_{\text{Enc}} - y_0)^2$
(13) DistanceMatrix$[[\chi]] \longleftarrow$ calculation$[[D_{ij}]]_{\text{Enc}}$
(14) $z_j = \text{Dec}([[\widehat{Z}_{j_{\text{Unknown}}}]]_{\text{Enc}}) \longleftarrow$ (DistanceMatrix$[[\chi]]$, recData$_i$)
(15) Do loop until get sensing elements $[Z_j]$ of all the unknown point
(16) EndFor
(17) EndFor
(18) Platform output aggregation model results

Algorithm 1: Location-encrypted spatial aggregation computing (LeSAC) protocol.

It can transform many signatures into one aggregate signature, only transmit and verify the aggregate signature instead of all the users' signatures, and obviously reduce the communication and computation costs, especially in situations of distributed multiusers. Figure 3 shows the construction of the scheme.

Combining with the characteristics of MCS system, we embed the aggregation signature in the process of spatial data aggregation outsourcing calculation, which can be verified to prevent privacy attacks such as data tampering. The verifiable aggregate signatures scheme is described below.

The Key Generation Center (KGC) runs the master key generation algorithm with the security parameter λ. It chooses pairing groups $\mathbb{PG} = (\mathbb{G}, \mathbb{G}_{\mathbb{T}}, q, e)$ and hash functions H_1, H_2, H_3, H_4. The message m of each participant is a tuple $(x_i, y_i, Z_i, \text{ID})$, and ID is data providers' identity.

KeyGen: the master secret key msk is randomly $s \in \mathbb{Z}_p$, and user's partial private key is $d_{\text{ID}} = s \cdot H_2$. If v_{ID} is secret value, then the privacy key is $(d_{\text{ID}}, v_{\text{ID}})$, and the public key is $(\text{pk}_{\text{ID}}, v_{\text{ID}})$.

Signature: each data provider with identity ID generates a signature for data interaction with the cloud server, and it selects an element $r \in \mathbb{Z}_p, U = rP$, and H_3 is $H_3(m, \text{pk}, U)$.

$$V = d_{\text{ID}} + H_3 \cdot r \cdot s \cdot P + H_3 \cdot v_{\text{ID}} \cdot Q + r \cdot Q. \quad (6)$$

Then, signature σ_i is (U_i, V_i), which is signed on the message of each data provider.

Aggregate: delegate computing server runs the aggregation algorithm, and public keys, message, and signatures of k data providers participate in aggregation.

$$\sigma = \prod_{i=1}^{k} \sigma_i. \quad (7)$$

Verify: the sensing platform verifies message signatures with master public key and the secret verification key. For $i = (1, \ldots, n)$, compute H_3, and verify $e(V, P)$. If yes, validation succeeds.

The security of this algorithm has been fully proved, and we lead it into the MCS spatial aggregation computing framework, which can better guarantee the security of data.

6. Security and Efficiency

In this section, we evaluate the security of the location-encrypted aggregation scheme and analyze the computation cost and communication overhead impacted by the increase of MCS network participating users n. In addition, we verify the availability of the algorithm using the accuracy of the region reconstruction with real sensing data.

6.1. Security Analysis. In MCS security outsourcing computing architecture, the security goal of the protocol is to protect the data provider's location privacy. The adversary includes external attacks, mobile terminals, platforms, and servers. The system assumes that users and servers are semihonest and that multiple users do not collude, meaning

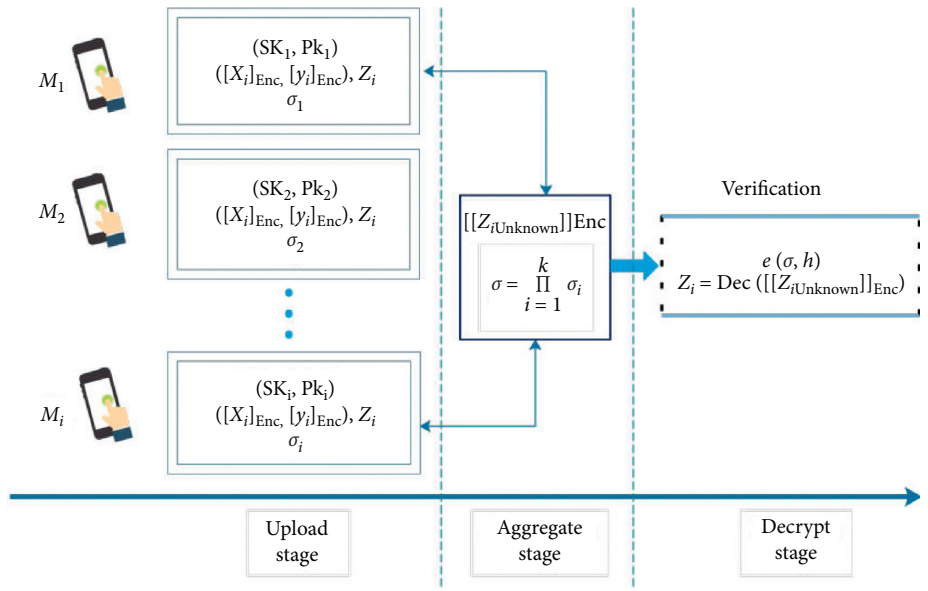

Figure 3: The verifiable scheme based on aggregate signatures.

that all entities will not cooperate to infer sensitive information about one party, but they may attempt to obtain additional information from server interaction calculations. In addition, a signature is attached to each data message to ensure that the data interaction is tamper-proof. The location ciphertext of node users in the scheme is transmitted to the central server, which transmits the aggregated computing model results to the platform through the network in the form of ciphertext. Obviously, the adversary cannot get any information in plaintext.

Theorem 1. *Based on the IND-CPA security of Paillier encryption scheme, the LeSAC protocol satisfies the indiscriminability of selecting plaintext.*

Proof. Paillier encryption scheme is $\Pi_p = (\text{Gen}, \text{Enc}, \text{Dec})$, the LeSAC computing protocol is Π_L, and hypothetical adversary A_L advantage is δ; then

$$\Pr\left[\text{Pub}K_{A,\pi_L}^{\text{IND-cpa}}(n) = 1\right] \leq \frac{1}{2} + \delta. \tag{8}$$

Based on the advantage of opponent A_L, an opponent A_P is constructed to attack Paillier encryption scheme. The challenge game is shown in Algorithm 2.

Bayes formula is adopted to solve

$$\begin{aligned}
&\Pr\left[\text{Pub}K_{A,\pi_L}^{\text{IND-cpa}}(n) = 1\right] \\
&= \frac{1}{2}\Pr[d = d' | d = 0] + \frac{1}{2}\Pr[d = d' | d = 1] \\
&= \frac{1}{2}\left(\frac{1}{2} + \delta\right) + \frac{1}{2}\left(\frac{1}{2}\right) \\
&\leq \frac{1}{2} + \text{negl}'(n).
\end{aligned} \tag{9}$$

When the adversary guesses $b = 0$, it wins the game with a probability of $(1/2 + \delta)$, where δ is a negligible function. Otherwise, when $b = 1$, the calculated value obtained by Algorithm 2 is independent of the user's location identity, and the server can win with only 1/2 probability.

In a word, ciphertext calculation request has no correlation with the real identity of the user. This algorithm has semantic security, and the adversary cannot know the location information of participating users. The proposed protocol is meeting security requirements. □

6.2. Efficiency Analysis. We first analyze the communication complexity. The protocol includes two communication transmissions: one from the data supply group to the cloud computing server and the other from the cloud computing server to the local server. From the interaction stage of the protocol, it can be seen that the data provider needs to send two encrypted ciphertexts and a plaintext information value to the computation center. The ciphertext space is $\mathbb{Z}_{N^2}$, and each ciphertext is $\log_2 N^2$ bit. Since n party users participate in the protocol, the total communication complexity of encryption protocol is $2n \log_2 N^2 + 2$ bit.

Meanwhile, we develop Python model code to implement our privacy protection data aggregation scheme, and experiments simulate the computational cost of aggregation calculation with distributed user location data and sensing data. The experimental environment is as follows: Intel (R) Core (TM) i5-6500 CPU 3.30 GHz eight-core processor and 8 GB RAM memory. In our scheme, the number n of MCS system sensing participants directly affects the communication efficiency of aggregation computation and aggregation signature. As shown in Figure 4, with the number of users participating in the sensing task n increasing, the computational cost increases linearly. Generally, in our scheme, lightweight outsourcing ciphertext computing can be implemented for energy-constrained mobile users and platforms.

Input: the key (P_K, S_K) of challenger, the plaintext (m_0, m_1), random value b
Output: Security judgment of the algorithm
(1) Challenger C_p setup a instance of Π_p, runs [Gen], gets key (P_K, S_K), and sends P_K to adversary A_p
(2) A_p sends P_K to A_L
(3) A_L outputs a pair of locations (x_0, y_0) , (x_1, y_1), sends them to A_p
(4) A_p computes d_0, d_1, and sends to C_p as Plaintext (m_0, m_1)
(5) C_p chooses a uniform bit $b \in\in (0, 1)$, and sends $c = \mathrm{Enc}(P_K, m_b)$ to A_L
(6) A_L generates a interact script, outputs b'
(7) If $b' = b$, the experiment outputs 1, which means the adversary A success, and 0 otherwise

Algorithm 2: Security analysis experiment.

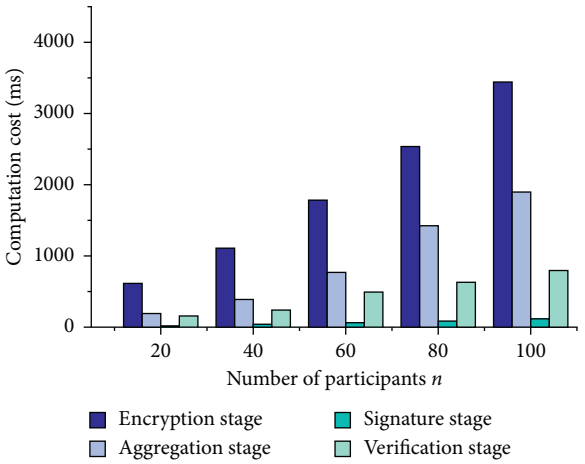

Figure 4: Computation cost of each stage.

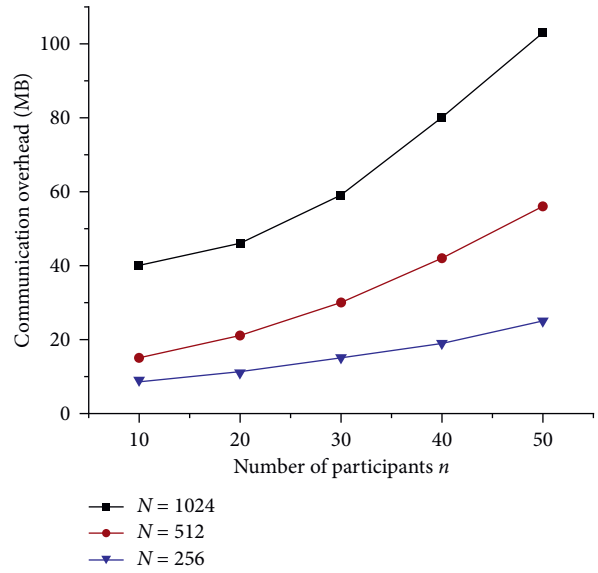

Figure 5: Performance comparison.

In addition, we choose $N = 256$, $N = 512$, and $N = 1024$ as the basic cryptoparameters for algorithm simulation and compared the communication overhead of training phase for security protocol with the increase of participants n; the experimental results are shown in Figure 5. The experiment shows that the communication overhead increased with increasing the number of MCS network participating users n.

To further verify the efficiency of the proposed scheme, we compared many kinds of secure computing schemes suitable for MCS multisource data aggregation analysis. As showed in Table 2 Xu et al. [32] proposed a high throughput secure multiparty $(n \geq 3)$ computation protocol. Experimental investigation shows that when the data source is complex and the number of participants increases, the communication efficiency of our scheme which uses the Geohash encoding indexing mechanism is obviously better than that in [32]. Kong et al. [33] proposed a novel efficient location privacy-preserving data sharing scheme by homomorphic encryption and proxy reencryption technique. But it makes a loss in calculating energy consumption due to the reencryption of sensing data aggregation. The classical fully homomorphic encryption (FHE) scheme [34] and the optimized BGV scheme [35] can also realize ciphertext data aggregation calculation. However, its structure is complicated, the asymptotic complexity is t·polylog (λ) after optimization, and it is still far from practical application.

6.3. Accuracy Analysis. This section uses the real data set of OpenSense (Zurich) [36] to obtain the environment pollution terminal data; it can assess the location-encrypted data reconstruction algorithm we proposed based on Paillier encryption. The data sets include the latitude and longitude information of data acquisition terminal, the data acquisition time, and the ozone concentration particulate matter content. Ignoring the temporal variation characteristics, we analyze the accuracy of cipher collaborative computing compared with the actual data reconstruction.

In order to extract the evenly regional distribution sensing data, the experiment first carries out k-means clustering on the original data sets and found 160 cluster centers of the ozone data and the 119 cluster centers of particulate matter pollution data. Then, look for the nearest data points through the mean shift function, as shown in Figure 6(a). Figure 6(b) shows the regional data reconstruction, which is carried out by the distance-based attribute weight method proposed in Section 3. In order to verify the data accuracy after ciphertext calculation, the training samples were used after ciphertext aggregation calculation by Algorithm 1, and the data reconstruction effect was shown in Figure 6(c).

Table 2: Comparison of overhead and functionality.

	Total runtime of server (s)			Data privacy (s)	Verifiable (s)	Efficiency optimization
	$n = 5$	$n = 10$	$n = 50$			
Xu et al.	3.2	8.2	15	√	×	×
Kong et al.	0.35	0.82	2.62	√	√	×
Ours	0.42	0.76	1.32	√	√	√

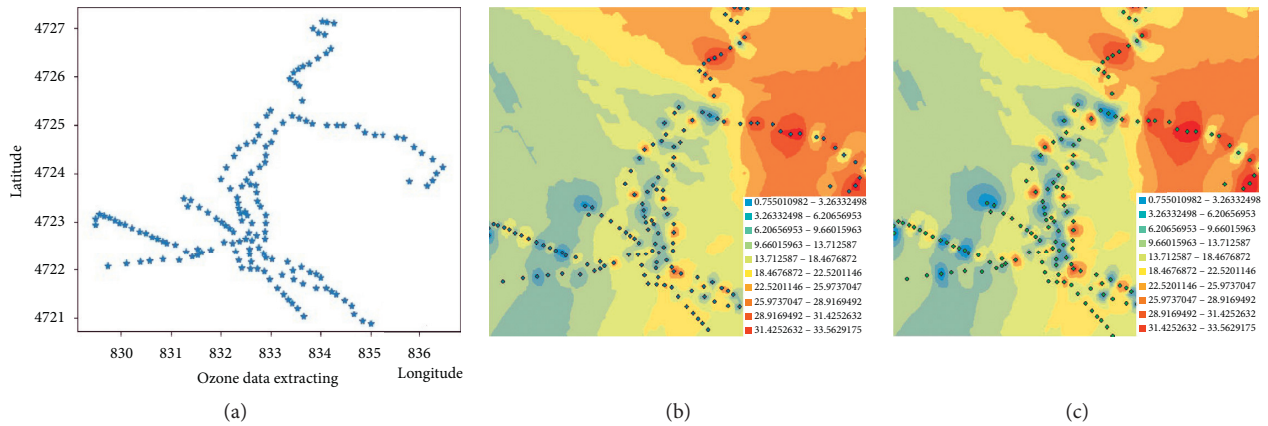

Figure 6: Ozone data reconstruction effect comparison. (a) Real data extraction. (b) Real data reconstruction. (c) Location-encrypted data reconstruction.

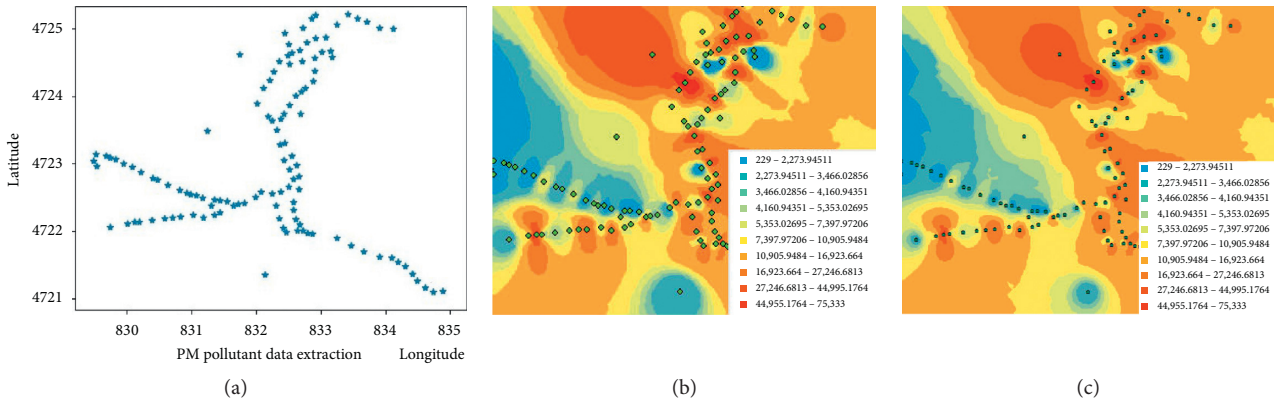

Figure 7: PM pollutant data reconstruction effect comparison. (a) Real data extraction. (b) Real data reconstruction. (c) Location-encrypted data reconstruction.

In addition, we quantify the fitting accuracy in terms of root mean square error (RMSE); it refers to the error between the measured value and the real value of spatial data points. The number of samples, the sparsity of spatial point distribution, the indexing mechanism, and the grid partitioning accuracy in our scheme all affect the data reconstruction accuracy. The experimental result shows that when using the LeSAC scheme, the global RMSE of ciphertext data (the ozone data) reconstruction increases by 2.52%; in other words, the reconstruction accuracy of data only loses by 2.52%.

We apply a comparative experiment for the particulate matter content data set shown in Figure 7 by the same approach; and the global RMSE of the data sets under LeSAC encryption protocol increases by 3.95%.

As shown in the Figures 6 and 7, the location encryption collaborative computing method we proposed can well realize the reconstruction of area data, while the loss of reconstruction accuracy is small. Therefore, in mobile crowd sensing network, data aggregation computation through encryption can meet the application requirements of data availability.

7. Conclusion and Future Work

In this paper, we proposed a spatial data aggregation scheme for privacy protection of MCS network. In the proposed scheme, the enhanced neighborhood inverse distance weighted aggregation protocol is presented to compute and fit the unknown point information, which achieves lower

calculation consumption. Based on Paillier homomorphic encryption system, we implemented a secure ciphertext computing protocol in order to protect the location privacy of data providers in the communication process. In additional, we added an efficient aggregate signature algorithm to achieve data security and verification. Moreover, the security analysis and efficiency analysis show that the proposed protocol satisfies the IND-CPA security and has high communication efficiency. The proposed scheme was tested on the real mobile crowd sensing data sets, and the result shows that the accuracy results meet the availability requirements of regional data interpolation fitting.

In the future, focusing on the privacy protection in data flow and aggregation computing, we plan to study the cooperative federated learning model of multiuser and multidimensional data. Based on secure multiparty computing, secret sharing model, and the application of homomorphic encryption, we will realize multiple federated computing methods of mobile crowd sense application.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. U1836205 and 61662009), the Science and Technology Program of Guizhou Province (nos. [2018]3001, [2019]2004, and [2020]4Y177), the Guizhou Science Contract Plat Talent (no. [2020]5017), the 13th Five-Year National Cryptography Development Foundation (no. MMJJ20170129), the Innovative Talent Team of Guizhou Colleges and Universities (no. [2013]09), the Youth Science and Technology Talents Growth Project (no. [2018]260), the Guizhou Science Contract (no. [2019]1249), the Research Project of Guizhou University for Talent Introduction (no. [2020]61), and the Cultivation Project of Guizhou University (no. [2019]56).

References

[1] P. Sun, Z. Wang, L. Wu et al., "Towards personalized privacy-preserving incentive for truth discovery in mobile crowdsensing systems," *IEEE Transactions on Mobile Computing*, pp. 1–151, 2020.

[2] G. Xu, H. Li, S. Liu, M. Wen, and R. Lu, "Efficient and privacy-preserving truth discovery in mobile crowd sensing systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 3854–3865, 2019.

[3] B. Guo, Z. Wang, Z. Yu et al., "Mobile crowd sensing and computing," *ACM Computing Surveys*, vol. 48, no. 1, pp. 1–31, 2015.

[4] T.-Y. Wu, Z. Lee, M. S. Obaidat, S. Kumari, S. Kumar, and C.-M. Chen, "An authenticated key exchange protocol for multi-server architecture in 5G networks," *IEEE Access*, vol. 8, pp. 28096–28108, 2020.

[5] R. Rana, C. T. Chou, N. Bulusu, S. Kanhere, and W. hu, "Ear-phone: a context-aware noise mapping using smart phones," *Pervasive and Mobile Computing*, vol. 17, pp. 1–22, 2015.

[6] D. Hasenfratz, O. Saukh, S. Sturzenegger, and L. Thiele, "Participatory air pollution monitoring using smartphones," in *Proceedings of the the 2nd International Workshop on Mobile Sensing*, pp. 1–5, Jaipur, India, January 2012.

[7] H. Schulze, "Crowd research partners "insider threat report": hopes and fears revealed," 2017, https://www.csoonline.com/article/3238867.html.

[8] P. Zhou, W. Chen, S. Ji, H. Jiang, L. Yu, and D. Wu, "Privacy-preserving online task allocation in edge-computing-enabled massive crowdsensing," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7773–7787, 2019.

[9] G. Xu, H. Li, C. Tan, D. liu, Y. Dai, and K. Yang, "Achieving efficient and privacy-preserving truth discovery in crowd sensing systems," *Computers and Security*, vol. 69, pp. 114–126, 2016.

[10] C. Liu, Y. Tian, J. Xiong, Y. Lu, Q. Li, and C. Peng, "Towards attack and defense views to k-anonymous using information theory approach," *IEEE Access*, vol. 7, pp. 156025–156032, 2019.

[11] W. Zhang, M. Li, R. Tandon, and H. Li, "Online location trace privacy: an information theoretic approach," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 235–250, 2019.

[12] S. B. Othman, A. A. Bahattab, A. Trad, and H. Youssef, "Confidentiality and integrity for data aggregation in WSN using homomorphic encryption," *Wireless Personal Communications*, vol. 80, no. 2, pp. 867–889, 2015.

[13] H. Huang, T. Gong, P. Chen, R. Malekian, and T. Chen, "Secure two-party distance computation protocol based on privacy homomorphism and scalar product in wireless sensor networks," *Tsinghua Science and Technology*, vol. 21, no. 4, pp. 385–396, 2016.

[14] J. Xiong, R. Bi, M. Zhao, J. Guo, and Q. Yang, "Edge-assisted privacy-preserving raw data sharing framework for connected autonomous vehicles," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 24–30, 2020.

[15] K. Niu, C. Peng, Y. Tian, and W. Tan, "Spatial ciphertext aggregation computing scheme for mobile crowd sensing privacy protection," in *Proceedings of the 2020 International Conference on Networking and Network Applications (NaNA)*, pp. 388–393, Haikou City, China, December 2020.

[16] J. Xiong, R. Ma, L. Chen et al., "A personalized privacy protection framework for mobile crowdsensing in IIoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 4231–4241, 2020.

[17] B. Ying and A. Nayak, "Lightweight remote user authentication protocol for multi-server 5G networks using self-certified public key cryptography," *Journal of Network and Computer Applications*, vol. 131, pp. 66–74, 2019.

[18] S. Wang and R. O. Sinnott, "Protecting personal trajectories of social media users through differential privacy," *Computers & Security*, vol. 67, pp. 142–163, 2017.

[19] D. Tao, P. Ma, and M. S. Obaidat, "Anonymous identity authentication mechanism for hybrid architecture in mobile crowd sensing networks," *International Journal of Communication Systems*, vol. 32, no. 14, pp. 1–16, 2019.

[20] J. He, L. Cai, and X. Guan, "Preserving data-privacy with added noises: optimal estimation and privacy analysis," *IEEE*

Transactions on Information Theory, vol. 64, no. 8, pp. 5677–5690, 2018.

[21] X. Liu, B. Qin, R. H. Deng, and Y. Li, "An efficient privacy-preserving outsourced computation over public data," *IEEE Transactions on Services Computing*, vol. 10, no. 5, pp. 756–770, 2017.

[22] J. H. Cheon and J. Kim, "A hybrid scheme of public-key encryption and somewhat homomorphic encryption," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 1052–1063, 2017.

[23] L. Liu, J. Su, X. Liu et al., "Toward highly secure yet efficient KNN classification scheme on outsourced cloud data," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 67–94, 2019.

[24] F. O. Catak, I. Aydin, O. Elezaj, and S. Yildirim-Yayilgan, "Practical implementation of privacy preserving clustering methods using a partially homomorphic encryption algorithm," *Electronics*, vol. 9, no. 229, pp. 1–15, 2020.

[25] K. Deepak and K. Chandrasekaran, "Investigating elliptic curve cryptography for securing smart grid environments," in *Proceedings of the 2020 Third ISEA Conference on Security and Privacy*, pp. 1–7, Guwahati, India, March 2020.

[26] H. Jiang, H. Wang, Z. Zheng, and Q. Xu, "Privacy preserved wireless sensor location protocols based on mobile edge computing," *Computers & Security*, vol. 84, pp. 393–401, 2019.

[27] H. Q. Wu, L. Wang, and G. Xue, "Privacy-aware task allocation and data aggregation in fog-assisted spatial crowdsourcing," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 589–602, 2019.

[28] X. Liu, R. H. Deng, K. K. R. Choo, and J. Weng, "An efficient privacy-preserving outsourced calculation toolkit with multiple keys," *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 1–14, 2016.

[29] K. Bonawitz, V. Ivanov, B. Kreuter et al., "Practical secure aggregation for privacy-preserving machine learning," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS'17)*, pp. 1175–1191, New York, NY, USA, October 2017.

[30] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proceedings of the Cryptology Eurocrypt'99*, pp. 223–238, Prague, Czech Republic, May 1999.

[31] G. Wu, F. Zhang, L. Shen, F. Guo, and W. Susilo, "Certificateless aggregate signature scheme secure against fully chosen-key attacks," *Information Sciences*, vol. 514, pp. 1–14, 2019.

[32] Y. Xu, C. Peng, W. Tan, Y. Tian, M. Ma, and H. Ding, "High-throughput secure multiparty multiplication protocol via bipartite graph partitioning," *Peer-to-Peer Networking and Applations*, vol. 14, pp. 1–17, 2021.

[33] Q. Kong, R. Lu, M. Ma, and H. Bao, "A privacy-preserving sensory data sharing scheme in internet of vehicles," *Future Generation Computer Systems*, vol. 92, pp. 644–655, 2019.

[34] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pp. 169–178, ACM, New York, NY, USA, June 2009.

[35] Z. Brakerski, C. Gentry, and V. Vaikuntanathan, "Fully homomorphic encryption without bootstrapping," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pp. 309–325, ACM, New York, NY, USA, January 2012.

[36] J. J. Li, B. Faltings, O. Saukh, D. Hasenfratz, and J. Beutel, "Sensing the air we breathe-the opensense zurich dataset," in *Proceedings of the AAAI'12*, pp. 323–325, Toronto, Canada, July 2012.