A Homomorphic Encryption Based Location Privacy Preservation Scheme for Crowdsensing Tasks Allocation

Xiaodong Zheng (lnxiaodong@163.com)
Harbin Engineering University https://orcid.org/0000-0001-5758-5992
Qi Yuan
Qiqihar University
Bo Wang
Harbin Engineering University
Lei Zhang
Jiamusi University

Research Article

Keywords: crowdsensing, homomorphic encryption, location privacy, privacy preservation, tasks allocation

Posted Date: June 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-552689/v1

A homomorphic encryption based location privacy preservation scheme for crowdsensing tasks allocation

Xiaodong Zheng 1,2，Qi Yuan 3，Bo Wang1 ,2, and* Lei Zhang1,4
1 College of Computer Science and Technology,Harbin Engineering University, Harbin, 150001, P. R. China
2 College of Computer Control and Engineering,Qiqihar University, Qiqihar, 161006, P. R. China
3 College of Communication and electronic engineering,Qiqihar University, Qiqihar, 161006, P. R. China
4 College of Information Science and Electronic Technology, Jiamusi University, Jiamusi,154007, P. R. China
**Correspondence: lnxiaodong@163.com*

Abstract: In the process of crowdsensing, tasks allocation is an important part for the precise as well as the quality of feedback results. However, during this process, the applicants, the publisher and the authorized agency may aware the location of each other, and then threaten the privacy of them. Thus, in order to cope with the problem of privacy violation during the process of tasks allocation, in this paper, based on the basic idea of homomorphic encryption, an encrypted grids matching scheme is proposed (short for EGMS) to provide privacy preservation service for each entity that participates in the process of crowdsensing. In this scheme, the grids used for tasks allocation are encrypted firstly, so the task matching with applicants and publisher also in an encrypted environment. Next, locations used for allocation as well as locations that applicants can provide services are secrets for each other, so that the location privacy of applicants and publisher can be preserved. At last, applicants of task feedback results of each grid that they located in, and the publisher gets these results, and the whole process of crowdsensing is finished. At the last part of this paper, two types of security analysis are given to prove the security between applicants and the publisher. Then several groups of experimental verification that simulates the task allocation are used to test the security and efficiency of EGMS, and the results are compared with other similar schemes, so as to further demonstrate the superiority of proposed scheme.

Keywords: crowdsensing, homomorphic encryption, location privacy, privacy preservation, tasks allocation

0 Introduction

Recently, with the prosperous development of wireless communication as well as the application of 5G wireless system, mobile devices such as smart phone and tablet computer are widely used in nearly all aspects of people's daily life [1-3]. Then devices that equipped plenty of sensing modules (such as GPS, the temperature sensor, the gyroscope and so on) can be used to gather sensing information in adjacent, and the collected information can be sent to the research department or government to further used to improve the quality of people's daily life [4-6]. In general, the process of collecting sensing information from applicants is termed the crowdsensing and the definition of crowdsensing mainly refers to the method of obtaining sensing information from the distributed of mobile devices [7, 8]. During the process of crowdsensing, the publisher does not need to arrive at the sensing place and can get the update sensing information in real time, and the applicant can obtain a certain number of rewards from the publisher . As a result, both the publisher and the applicant will benefit from each other, as the publisher will get the sensing information with a lower expenditure on information collection and the applicant will get an incentive with just a bit of sensing information transformation [9-11].

However, along with the convenience of this type of service, the risk of privacy violence cannot be neglected [12-14]. For example, if a publisher wants to get the temperature variation of a specific region, he will send the sensing request to the authorized agency or broadcasts the sensing request to all potential applicants. Then applicants that are located in this region will feed back the sensing result to the publisher and get rewards from the publisher. During the process of publishing the request and getting feedback result from applicants, the publisher will learn the location of the applicant and the applicant will aware the sensing location of the publisher. In addition, these locations will also be released to the authorized agency. If an entity in the crowdsensing is un-trusted, the personal privacy of the publisher or the applicant will be violated [15, 16]. Thus, in order to cope with the problem of privacy violation in the process of crowdsensing, Yang et al. [17] utilized the model of

differential privacy to deal with the density of sensing applicants and generated a location privacy preservation scheme for the mobile crowdsensing. But this scheme mainly depends on adding dummies to generalize the real location, which will bring the lack of accuracy for feedback result. Thereafter, in order to cope with the same problem, Yang et al. [18] utilized the conception of distributed consensus from the blockchain to achieve the location privacy preservation of task allocation in crowdsensing. But the collaboration of this scheme will be affected by the willingness of collaborative users in the public chain. Then, in recent years, along with the deepening researches of crowdsensing, more schemes for preserving personal privacy are proposed. For example, the scheme of privacy preservation bus sharing [19], the scheme of sparse region sensing [20], the scheme for adjacent sensing [21] as well as the scheme for privacy recommending members [22] and so on.

Although schemes mentioned above can provide privacy preservation service for the publisher and the applicant in a certain extent, they do not achieve the closed loop of information transformation for entities in the crowdsensing, i.e. these schemes usually assume that the authorized agency is a trusted entity (such as the scheme of privacy preservation bus sharing [19] and the scheme of sparse region sensing [20]) or mainly designed for preserving the privacy of applicants (such as the scheme for adjacent sensing [21] and the scheme for privacy recommending members [22]). In addition, schemes used in recently usually neglect that the publisher also needs to preserve the personal privacy of his own. Thus, in order to cope with the violation of location privacy in the process of task allocation in crowdsensing and ensure during the process of task allocation no information about each entity can be released to any entity. In this paper, based on the conception of homomorphic encryption and encrypted grids matching, a privacy preservation scheme for the task allocation in crowdsensing has been proposed based on Paillier encryption system.

1 The system architecture and definitions

1.1 The system architecture

In general, a basic process of task allocation for crowdsensing can be conducted by three entities, and these entities can be called the publisher, applicants as well as the authorized agency. The publisher mainly refers to the user who wants to get crowdsensing results from others, and issues rewards as the incentive to a user who provides the precise result. The applicants refer to a group of users that send crowdsensing results to the publisher and obtain incentive. The authorized agency refers to an organization that generated by the government or large enterprises, which can provide central services to publish crowdsensing tasks or incentive rewards. As shown in Figure 1, in the process of crowdsensing, the publisher first sends the request of crowdsensing and the location grids of task allocation to the authorized agency. Next, the publisher publishes its public key to all applicants, so as to guarantee the security of feedback results. Once the authorized agency receives the request, this entity publishes the request to applicants. Then each applicant checks its location grid, if the grid matches request, this applicant sends the crowdsensing result to the authorized agency. Otherwise, the applicant waits for other tasks. The authorized agency collects all crowdsensing results and generates a set of these results, and then sends this set to the publisher. At the same time, the authorized agency issues incentives to all applicants who have sent the precise crowdsensing result by their locations.

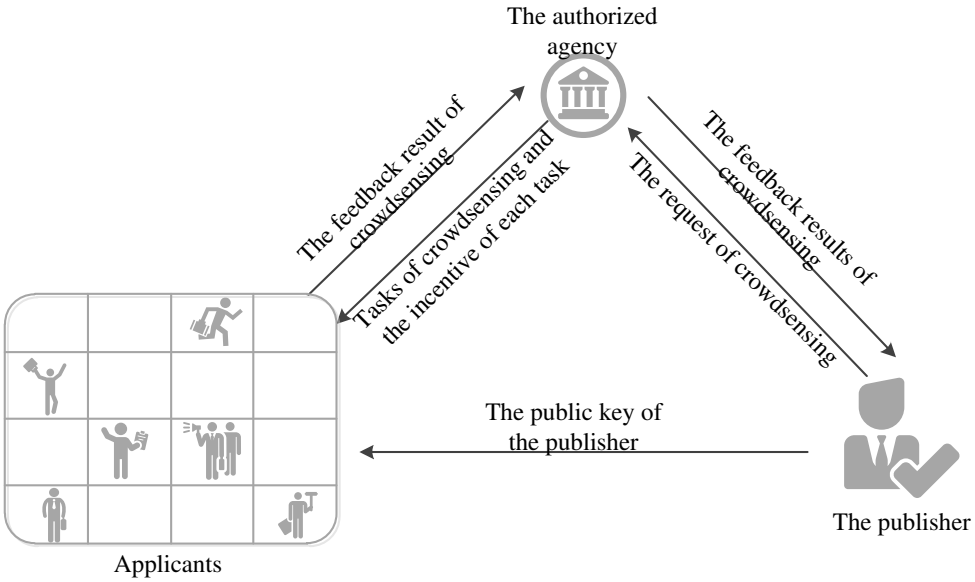

Fig.1 The system architecture of crowdsensing

Based on the system architecture and the brief process of crowdsensing shown in Figure 1, the location privacy of the publisher as well as the location privacy of the applicant will be violated in the step of tasks allocation and the step of results feeding back. In general, the location is a crucial condition that must be used to confirm where the crowdsensing task is needed, and where is the feeding back result the applicant can provide. With the precise location of the applicant, a malicious publisher can publish this information to others with commercial exploitation or pushes advertisement. With the precise location of the publisher, a malicious applicant can construct inaccurate feedback result and utilizes it to obtain the incentive. Furthermore, when the authorized agency matching a task he can also get information about locations from both publisher and applicant, so the location awareness matching further jeopardizes location privacy of both parts of entities in crowdsensing. In conclusion, the privacy violation of crowdsensing system is a serious problem, and it is also a complicated issue for each entity to preserve its own location privacy. Thus, a scheme used in tasks allocation that can provide location privacy for three entities simultaneously and guarantees each entity cannot get the precise location is needed.

1.2 The Paillier homomorphic encryption

As in the process of crowdsensing, the publisher, applicants and the authorized agency may aware the location of each other, and then threaten the privacy of them (as shown in Figure 1), a scheme is needed to find the best solution to publish a task with confidential location matching. As a result, three entities in the process of task allocation cannot obtain any information about each other. For confidential calculation, the strategy of homomorphic encryption is a better idea, as the result of calculation with encrypted parameters is identical to the result of calculation without encryption. Thus, in this paper, Paillier's public-key cryptosystem is considered to be used in the process of confidential location matching, and the workflow of this cryptosystem can be described as follows.

Firstly, two different large prime numbers p and q are selected randomly, and these numbers satisfy $\gcd(pq,(p-1)(q-1))=1$. Then calculates $N=pq$, and randomly selects a integer $g\in Z_{N^2}^{*}$, publishes the public key $pk=(N,g)$, preserves the private key $sk=(p,q)$.

The process of encryption: for the given public key pk and the plaintext m (which is an integer less than N). Then selects a random integer r from the set of $Z_{N^2}^{*}$, and calculates the ciphertext with $c=g^m r^N \pmod{N^2}$. For two parts of given plaintext $m_1, m_2 \in Z_{N^2}^{*}$, the encryption satisfies the following characters.

$$E(m_1)E(m_2)=E(m_1+m_2), \tag{1}$$

$$E(m_1)^{m_2}=E(m_1 m_2). \tag{2}$$

The process of decryption: for the ciphertext c, λ is the lowest common multiple (LCM) of $(p-1)$ and $(q-1)$, where $\lambda=lcm((p-1),(q-1))$, and then the plaintext can be get with

$$m=\frac{(c^{\lambda} \bmod N^2-1)/N}{(g^{\lambda} \bmod N^2-1)/N} \bmod N. \tag{3}$$

So based on characters of confidential addition method mentioned above, a confidential location matching scheme can be devised to solve the privacy problem in crowdsensing.

2 The scheme of encrypted grids matching

2.1 The basic idea and conception

In the process of crowdsensing tasks allocation, the region of sensing can be divided as a square area with multi-grids. Then the user who locates in the specified grid can be seen as an applicant, once he accepts the request of sensing task and can feed back sensing result for the publisher. So the applicant usually takes mobile sensing devices along with his moving and can communicate with the publisher and the authorized agency. In order to find plenty of applicants to satisfy the request of the publisher, the authorized agency also needed, and utilized to publish tasks and collect feeding back results.

Therefore, based on the process of task allocation as well as multiple grids of sensing region, the basic idea and conception of privacy preservation can be concluded as four steps. In the first step, the authorized agency sets the sensing region and grids, and the publisher and each applicant specifies the values of sensing and servicing grids with sensing request and sensing feedback ability. In the second step, the publisher and each applicant

encrypts the grid values with the public key of Paillier encryption system, and both entities send the encrypted result to the authorized agency. In the third step, the authorized agency calculates the result of matrix matching by the encrypted values from the publisher and applicants and sends the matching result to the publisher. The publisher checks whether the result of matrix matching satisfies the sensing request. If so, the publisher requests the authorized agency to collect encrypted results of crowdsensing from applicants. Otherwise, the publisher asks the authorized agency to publish the task once again and checks the result of matrix matching one more time. In the fourth step, the authorized agency collects feedback results from all applicants and sends the result set to the publisher, and then the whole process of crowdsensing is over.

During this process, location information that the authorized agency receives from the publisher and applicants is encrypted with the Paillier encryption system, so the authorized agency knows nothing about the sensing grids without the private key. The feedback result also encrypted with the public key of the publisher, so the authorized agency also knows nothing about the feeding back result. For the publisher, there is nothing about location information about applicants sent to him, as the process of sensing grids matching is just conducted by the authorized agency. For applicants, the authorized agency also conducts sensing grids matching, and nothing about the publisher is sent to any applicant. Thus, as a result, during the whole process of tasks allocation in crowdsensing, there is nothing about the location of each entity can be obtained by any entities and the location privacy of them is preserved.

2.2 The process of location privacy preservation crowdsensing tasks allocation

Based on the basic idea of privacy preservation that mentioned in 2.1, in this paper the process of tasks allocation in crowdsensing can be converted to the problem of confidential matrix matching. Thus, the authorized agency needs to divide the area of crowdsensing into a region with multiple grids, and broadcasts the matrix of these grids to other entities such as the potential publisher and the potential applicants, and then these entities can leverage the matrix to finish the conduct of location matching.

For example, suppose a sensing region can be converted to a three-by-three matrix. In this matrix, elements that the publisher wants to get results of crowdsensing from are denoted as random numbers from 1 to 9, and elements not need to be sensed are denoted as 0. Thus, in the process of tasks allocation, the publisher utilizes this matrix and the specified grids with random number to finish the location matching. Next, the publisher encrypts each element in the matrix with the generated public key of Paillier encryption system and also broadcasts this public key to all applicants. As shown in equation 4, an area the publisher needs to sensing is converted to a matrix and denoted as A, and elements of this matrix the publisher want to get sensing results from are signed as random numbers from 1 to 9, others are signed as 0. Then the publisher leverages the public key to encrypt each element in the matrix of A and gets the encrypted matrix of $E(A)$.

$$E(A)=E(\begin{bmatrix}2&0&4\\0&8&0\\1&6&0\end{bmatrix})=\begin{bmatrix}E(2)&E(0)&E(4)\\E(0)&E(8)&E(0)\\E(1)&E(6)&E(0)\end{bmatrix} \tag{4}$$

After generating the encrypted matrix, the publisher sends this matrix and the incentive together to the authorized agency. Then the authorized agency broadcasts the request of sensing region to all users located in each grid and waits for applications sent from the user who is located in each grid and willing to participate in crowdsensing. For the user, if he wants to participate in crowdsensing, he has to check whether his location is in the grid that can provide feedback results. If so, the user converts to an applicant and construct the encrypted matrix with the public key of Paillier encryption system as shown in equation 5. Otherwise, this user cannot be seen as an applicant and has to wait for another task. Like the process of the publisher, the applicant also denotes the elements in the grid matrix which he can provide crowdsensing results with random numbers from 1 to 9 and others are signed as 0. Then the applicant encrypts the elements in this matrix with the public key sent from the publisher, and gets the confidential matrix $E(B)$. Next, the applicant sends the encrypted matrix to the authorized agency.

$$E(B)=(\begin{bmatrix}1&2&1\\0&0&0\\0&0&0\end{bmatrix})=\begin{bmatrix}E(1)&E(2)&E(1)\\E(0)&E(0)&E(0)\\E(0)&E(0)&E(0)\end{bmatrix} \tag{5}$$

Once various applicants apply for the crowdsensing simultaneously, the authorized agency will receive

several matrixes. Then the authorized agency constructs these matrixes into a matrix set and denotes it as $B_i, 0 \le i \le n$,(where n is the maximum of the publisher can reward these applicants), and then he calculates $C = A \cdot B_1 + A \cdot B_2 + ,...,+ A \cdot B_i$ and gets the encrypted result matrix without knowing anything about the value of each element. The encrypted result of calculation is then sent to the publisher and used to check whether satisfies the request of the publisher. The process of calculation can be seen in equation 6. The matrixes of $E(A)$ and $E(B)$ that used in equation 6 are derived from equation 4 and equation 5.

$$E(C) = E(A) \cdot E(B) = \begin{bmatrix} E(2\times1) & E(0\times2) & E(4\times1) \\ E(0\times0) & E(8\times0) & E(0\times0) \\ E(1\times0) & E(6\times0) & E(0\times0) \end{bmatrix} = \begin{bmatrix} E(2) & E(0) & E(4) \\ E(0) & E(0) & E(0) \\ E(0) & E(0) & E(0) \end{bmatrix} = E(\begin{bmatrix} 2 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}) \quad (6)$$

The publisher then decrypts the matrix C with the private key from Paillier encryption system. If the decrypt matrix contains all grids of the request, the publisher sends a message to the authorized agency to trigger the process of distributing incentive to the applicant who was concluded in this matrix set. Otherwise, the publisher sends a message to the authorized agency for initiating another request, and then the authorized agency releases the new request to the region which needs to crowdsensing, so as to select other applicants.

In the condition of the matrix do not encrypted by Paillier encryption system, suppose the non-zero elements in the matrix A construct the grids need to be sensed, the non-zero elements in the matrix B denote the grid that this applicant can provide the result of crowdsensing. As a result, the non-zero elements in the matrix C which is calculated by $A \cdot B$ can be seen as the matched grids. Once the publisher collects enough grids with non-zero elements from various applicants, he will get all grids that he wants to obtain the sensing result as well as applicants in these grids, so that the task allocation of crowdsensing will be finished successfully.

2.3 The process of information conduct (tasks publisher)

Suppose in matrix A, elements that the publisher constructed for each grid can be denoted as $a_{ij}, 0 \le i, j \le n$.

Based on the Paillier encryption system, the encrypted element can be denoted as $E(a_{ij})$, which means the element is calculated by the public key of the publisher. Then the publisher sends the matrix with encrypted elements to the authorized agency. The process of establishing the confidential matrix for the publisher is shown in Algorithm 1.

Algorithm 1： The establishment of confidential matrix for the publisher

Input： Location matrix $A = a_{ij}$

Output： Encrypted location matrix $E(A) = E(a_{ij})$

1) The publisher establishes a pair of public key and private key with Paillier algorithm;
2) **For** (j=0;j<=n;j++) **do**
3) **For** (i=0;i<=n;i++) **do**
4) $E(a_{ij}) = g^{a_{ij}} r^N (\mathrm{mod}\, N^2)$;
5) **End For**
6) **End For**
7) Return $E(a_{ij})$;

The process of getting public key from the Paillier encryption system has been mentioned in detail in the section 1.2, and in this paper we do not repeat it. In Algorithm 1, the confidential matrix of the publisher is constructed and denoted as $E(A)$ which can be used to calculate the matching grids, and the element of this matrix is encrypted with the public key, so without the private key of the publisher, no one can infer the precise

value of each element. Then this confidential matrix $E(A)$ is sent to the authorized agency.

2.4 The process of information conduct (tasks applicants)

Suppose that each applicant can get a public key $pk = (N, g)$ from the broadcast by the publisher. Each applicant can establish a square region with the grids of providing sensing result, and then this region can be converted to a two-dimensional matrix denotes as *B*, where the elements in this matrix correspond to the grids in this region and can be denoted as $b_{ij}, 0 \le i, j \le n$. Then the applicant encrypts each element with the public key pk and gets the confidential matrix $E(b_{ij})$, and then submits this matrix to the authorized agency. In Algorithm 2 we show the process of how the applicant encrypts the matrix.

Algorithm 2： The establishment of confidential matrix for applicants

Input： The location matrix of applicants $B = b_{ij}$

Output： Encrypted location matrix $E(B) = E(b_{ij})$

1) The applicant leverages the public key pk encrypts the location matrix of its own;
2) **For** (j=0;j<=n;j++) **do**
3) **For** (i=0;i<=n;i++) **do**
4) $E(b_{ij}) = g^{b_{ij}} r^N (\mathrm{mod}\, N^2)$;
5) **End For**
6) **End For**
7) Return $E(b_{ij})$;

As a number of applicants located in the sensing region, most of them can provide feedback results of different grids to the publisher, so the applicant who wants to participate in crowdsensing can utilize Algorithm 2 to encrypt his own matrix. Then these applicants send encrypted matrixes to the authorized agency, and then the authorized agency can calculate the matching of matrixes from the publisher and applicants to allocate the task of crowdsensing in confidentiality.

2.5 The process of information conduct (the authorized agency)

The authorized agency is the central entity for the task allocation and the central entity for privacy preservation, so most of the management (such as the grids matching and the message sending) is conducted in this entity. Once the authorized agency received the matrix from the publisher and matrixes from various applicants, he calculates the encrypted matrixes with the feature of Paillier encryption system, and gets the matched matrix $E(C) = E(A) \cdot E(B_1) + E(A) \cdot E(B_2) + ,..., + E(A) \cdot E(B_i)$, where $E(C) = E(c_{ij})$. Then the matching matrix $E(C)$ is sent to the publisher, and checked by the publisher to confirm whether this matrix satisfies the request. The algorithm of tasks matching with privacy preservation can be seen in Algorithm 3.

Algorithm 3： The tasks matching with location privacy preservation

Input： The location matrix $E(a_{ij})$, the matrix of applications located in $E(b_{ij})^k$

Output： The feeding back matrix with the secret calculation $E(c_{ij})$

1) **For** (j=0;j<=n;j++) **do**
2) **For** (i=0;i<=n;i++) **do**
3) $E(c_{ij}) = E(a_{ij}) \cdot E(b_{ij})^1 + E(a_{ij}) \cdot E(b_{ij})^2 + ,..., + E(a_{ij}) \cdot E(b_{ij})^k$;

4) **End For**

5) **End For**

6) Return $E(c_{ij})$;

Once the publisher received the result matrix $E(C)$, he has to decrypt this matrix with the private key of Paillier encryption system. Then the publisher checks whether the non-zero elements in this matrix cover all the grids that the publisher needs, if so, this matrix can provide feedback results of his needs and the publisher asks the authorized agency for collecting the results. Otherwise, he has to ask the authorized agency to release this request to applicants again, and asks the authorized agency to find other applicants for crowdsensing.

3 Security analysis

During the process of task allocation in crowdsensing, the information about personal location is submitted by two entities (the publisher as well as applicants). So in order to analyze the security of proposed scheme thoroughly, we should consider the security of sending the request of crowdsensing from the publisher as well as the security of sending the response for applying the task from applicants. That is we should ensure the privacy of two entities of the publisher and the applicant cannot be obtained by others.

3.1 The security of publisher

During the process of task allocation in crowdsensing, the threat of privacy violation for publisher will come from two entities, i.e. the authorized agency as well as applicants. For one thing, the authorized agency has to check whether grids of feeding back results satisfy the request of the publisher and in this process the authorized agency may get the precise location of the publisher requires. For another, applicants also want to know the grids that the publisher wants to get feedback results. So a privacy preservation scheme has to prevent these two types of entities from obtaining the location privacy. In fact, EGMS can provide the privacy preservation service for the publisher, and the security can be elaborated from two aspects. For resisting the privacy violation by the authorized agency, in EGMS, the information that the authorized agency gets is just the encrypted matrix $E(A)$, and without the private key $sk = (p, q)$ no one can decrypt the confidential elements in this matrix, let alone the precise location of the publisher wants to get the feedback result. In addition, even though the authorized agency gets the public key of this publisher, the authorized agency also difficult to encrypt a number to collide with the precise number that used to identify the grid that needs to be sensing. This is because different numbers can be used to denote the grid which the publisher requires to get feedback result in this matrix.

For the privacy violation from applicants, this type of entities also difficult to get the precise location of the publisher, as there is no interaction about information between the applicant and the publisher except the public key. Furthermore, if an applicant wants to leverage his own location grid to infer the region of the publisher wants to get feedback result, the result will be distorted and difficult to identify. Because the applicant can just sense the result from the grid that he located in, but the region that the publisher wants to get feedback result may be even larger and sometimes will contain a lot of grids. Suppose that, the number of grids that the publisher wants to get feedback result is denoted as k, and the number of grids that the applicant can sense is n (where k ? n), the probability of the applicant located in the grid that the publisher needs is n / k. If the number of k is large enough, it is difficult for the applicant infers the whole region of sensing, as the publisher can add redundant grids to generalize the real grid or the number of grids that he wants is larger. In addition, if the applicant can obtain the incentive from the authorized agency, the applicant also just knows that he is located in a grid that the publisher wants to get result but not the whole region of crowdsensing. If the applicant does not obtain the incentive, he will be even difficult to judge whether he is located in the grid that the publisher wants to get feedback result. Thus, we can conclude from the analysis mentioned above that it is difficult for an applicant to infer the privacy of the publisher by his own location, even though he has provided sensing result in crowdsensing.

3.2 The security of applicants

In the process of crowdsensing, an applicant usually wants to get the incentive from the publisher, but does

not want his precise location released to any entity (such as the publisher and the authorized agency). So during the process of confidential crowdsensing, the location privacy of each applicant needs to be preserved and not be released to both the publisher and the authorized agency. With regard to the confidential of EGMS, as the matrix of the applicant also encrypted by the public key of Paillier encryption system and without the private key $sk = (p, q)$ no one can decrypt the confidential elements in the matrix. Without the private key, no information about the precise location of the applicant can be released during the process of matrix matching. As a result, if the authorized agency is curious about the privacy of the applicant, EGMS can provide the service of privacy preservation for the applicant against the attack from the authorized agency.

In order to analyze the resistance of EGMS against attacks from the publisher, we assume that during the whole process of crowdsensing all information about the applicant is not sent to the publisher directly, and the publisher can just know the sensing result in the set of feedback results. In the process of task allocation, encrypted matrixes $E(b_{ij})^k$ are sent from various applicants to the authorized agency, and then the authorized agency calculates the matrix matching with another encrypted matrix $E(A)$ sent from the publisher, so during this process no information about the applicant is sent to the publisher. In addition, it is not certain about the applicant who provides which encrypted matrix, and the grid that the applicant can provide in the encrypted matrix is also ambiguous, as the process of task allocation is achieved by the encryption matrix matching. As a result, without the private key the authorized agency knows nothing about the applicant, let alone the publisher, as no information about the applicant is sent to the publisher directly. Thus, with the help of EGMS, the privacy of the applicant is preserved, and any entity such as the publisher and the authorized agency cannot get the privacy of the applicant.

In conclusion, based on the security analysis mentioned above, we can regard that during the whole process of crowdsensing with EGMS, and there is no plaintext information about each other interacted with three entities included in the crowdsensing system. Therefore, EGMS can provide the service of privacy preservation in the process of task allocation in the crowdsensing.

4 Experimental verification

4.1 Experiment settings

In order to verify the performance of our proposed scheme EGMS (such as the execution efficiency as well as the effectiveness of privacy preservation), in the experimental environment, several groups of experiments are given. Then in the following section, the results and reasons analysis are also given to further demonstrate the conclusion that we got in the security analysis. Simulation experiment is implemented on a laptop with I7 1.80GHz CPU, 8Gb memory and Windows 10 operation system, and we utilize Python 3.6 to test the algorithm proposed in this scheme and algorithms proposed from other references. The location data used in our simulation comes from Geolife, and we select the same region to generate the grids. Then we assume one out of ten users can feed back the result and can be seen as an applicant. Schemes used in comparison include the scheme of DPLP [12] which is based on the conception of differential privacy, the scheme of PAPP [13] which is based on the conception of path adjust, the scheme of differential-distortion [20] which is based on the conception of location distortion. Therefore, the results and reasons analysis are based on comparisons and analysis with these schemes. For evaluating the performance of execution efficiency, the running time, the matching probability of crowdsensing locations as well as the accuracy of feeding back results are used. For evaluating the performance of effectiveness of privacy preservation, we leverage the success ratio of guessing real locations and the value of entropy of information disclosure.

4.2 Experimental results and cause analysis

In the aspect of evaluating the performance of execution efficiency, we mainly focus on the running time, the matching probability of crowdsensing locations as well as the accuracy of feeding back results. Thus, we assume

in the simulation experiment all users in the crowdsensing are applicants that can feed back results to the authorized agency to obtain the incentive, and then we first simulate the running time of each scheme and get the comparison result. In general, the running time is calculated by the time from sending the request of crowdsensing to receiving the feedback result. Figure 2 shows the results of running time of various schemes. In this figure, compared with other schemes, we can see that, though EGMS utilizes the encryption system in calculating the process of confidential matrix matching, the running time of this scheme is shorter than other schemes.

The reason for this phenomenon can be concluded as the process of encryption or decryption can be conducted offline by the publisher or the applicant, and the authorized agency just implements additive operation on the basis of confidentiality. In addition, why the running time of EGMS is shorter than others also contains the following reasons. Firstly, matrix matching is calculated by the multiplication of matrix dot, which is less complex than matrix multiply, so the time complexity is much less than $O(n^3)$. Secondly, the process of matrix matching is focused on the non-zero element, but not the precise value, so during the process of matrix matching, the operation of XOR can be used to replace the multiplication of matrix dot, which can save much more time during the calculation. Thirdly, the number of grids is limited, so that the size of the matrix is restricted to a certain scale, and a small size of matrix also reduces the complex of calculating the element multiplication. For the running time of other schemes, as these schemes usually leverage adds dummy to satisfy the differential privacy to affect the result of location matching, or convert the road path during the process of data transmission, the running time of these schemes is higher than EGMS. This is because the calculation of adding dummies or the calculation of location shifting will increase the processing step of crowdsensing, and more steps of calculation usually bring more running time.

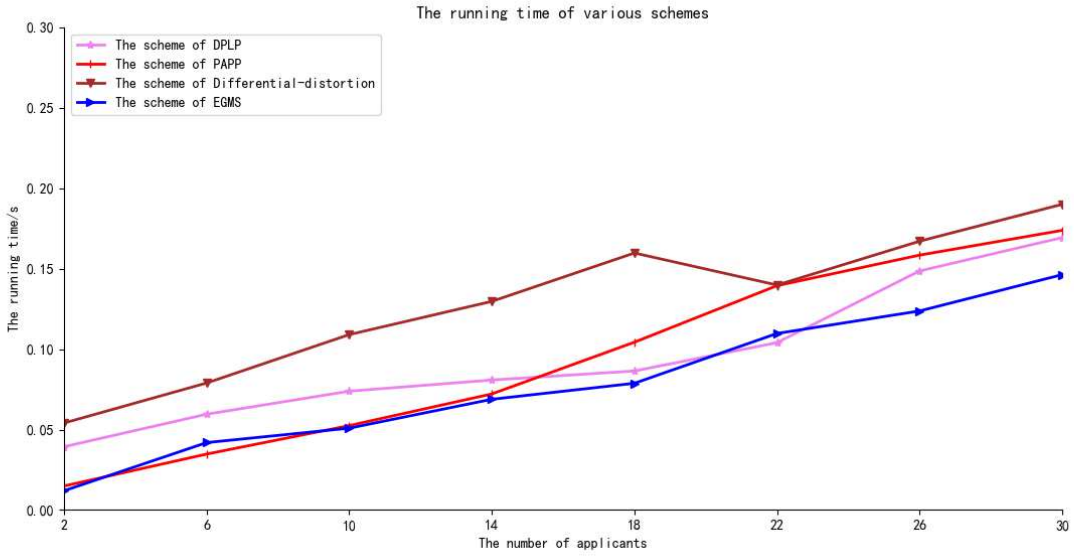

Fig.2 The running time of various schemes

Figure 3 shows the matching probability of crowdsensing locations of various schemes. In this figure, we can see that the scheme of EGMS performs better not only in the higher probability of matching, but also performs better in the stability of algorithm. The reason for this phenomenon can be summarized as the EGMS does not need to distort the location of the publisher or the applicant or add dummy to generalize the real location. During the whole process, all operations are just the location encryption and decryption, and no other location added in the calculation of location matching, so the matching probability is higher than other schemes (such as the scheme with dummy addition, the scheme with location transformation, the scheme with location distortion).

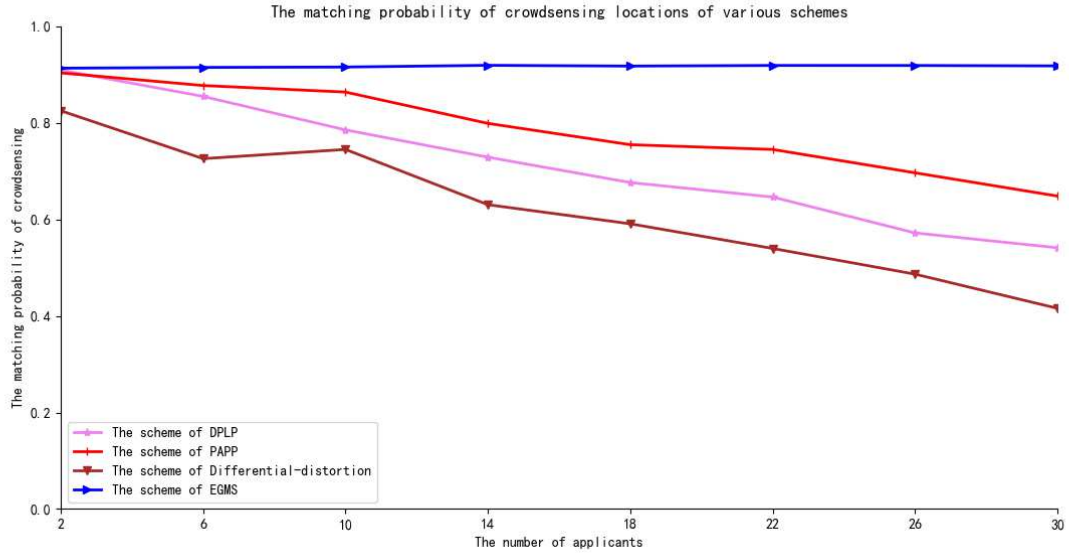

Fig.3 The matching probability of crowdsensing locations of various schemes

Figure 4 shows the accuracy of feeding back results of various schemes. In this figure, we can see the accuracy of feeding back results of EGMS is higher than other schemes. As there is no dummy added and no location distortion used in preserving the location privacy of the publisher and the applicant, all results are calculated by the precise grid of requiring and sensing. Thus, based on the feature of encryption, the scheme of EGMS performs the best among all schemes used in comparison, as other schemes usually utilize dummy addition or location transformation or location distortion which will mislead the accuracy of feedback result.

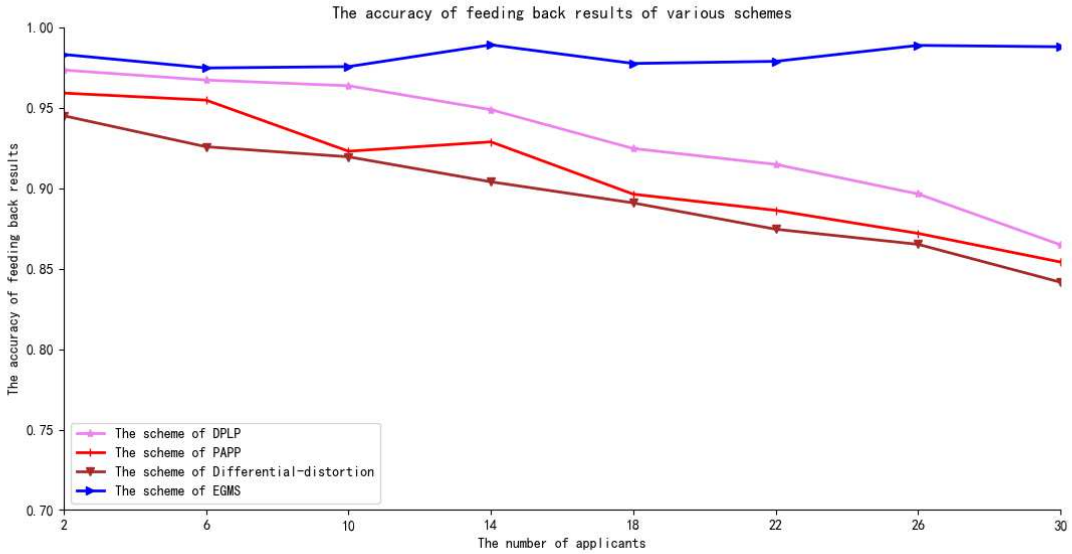

Fig.4 The accuracy of feeding back results of various schemes

For evaluating the performance of effectiveness of privacy preservation, we utilize the success ratio of guessing real locations and the value of entropy of information disclosure as metrics, and all results of comparison are based on above two aspects. Figure 5 shows the success ratio of guessing real locations. From this figure, we can see that for all schemes except EGMS, along with the increasing number of applicants in the task allocation, the risk of successfully identified by an adversary is gradually decreasing. This is because more applicants will generalize the precise location of each other and the adversary will be difficult to successfully identify the

specified user among similar applicants. However, the increasing number of applicants does not affect the performance of EGMS, as this scheme utilizes encryption to conduct the confidential matrix matching which does not depend on the obfuscation of the generalized locations. In addition, as no one can get the precise location without the private key of Paillier encryption system, the adversary cannot guess anything meaningful for any entity, and this feature further increases the privacy preservation level of EGMS.

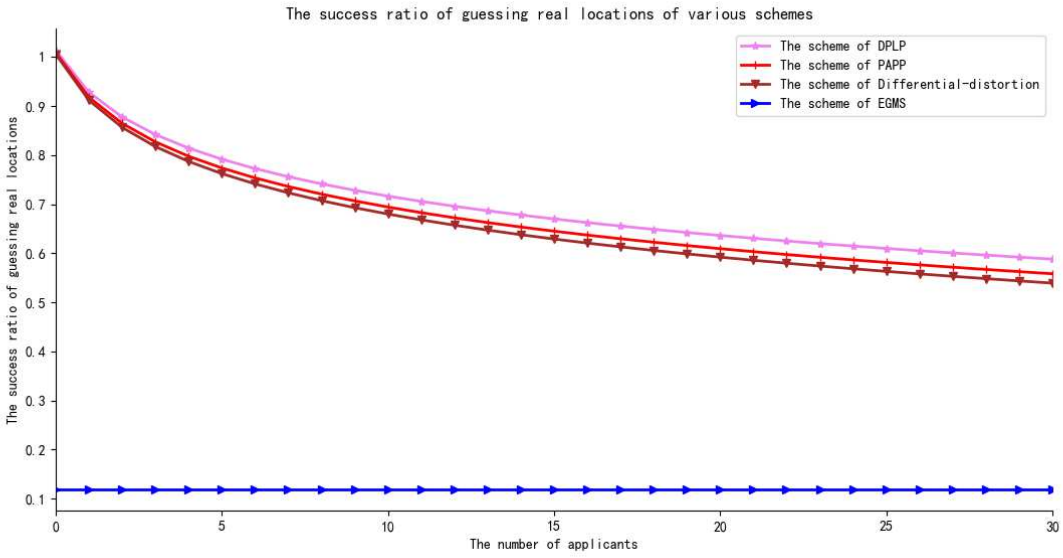

Fig.5 The success ratio of guessing real locations of various schemes (Adversary)

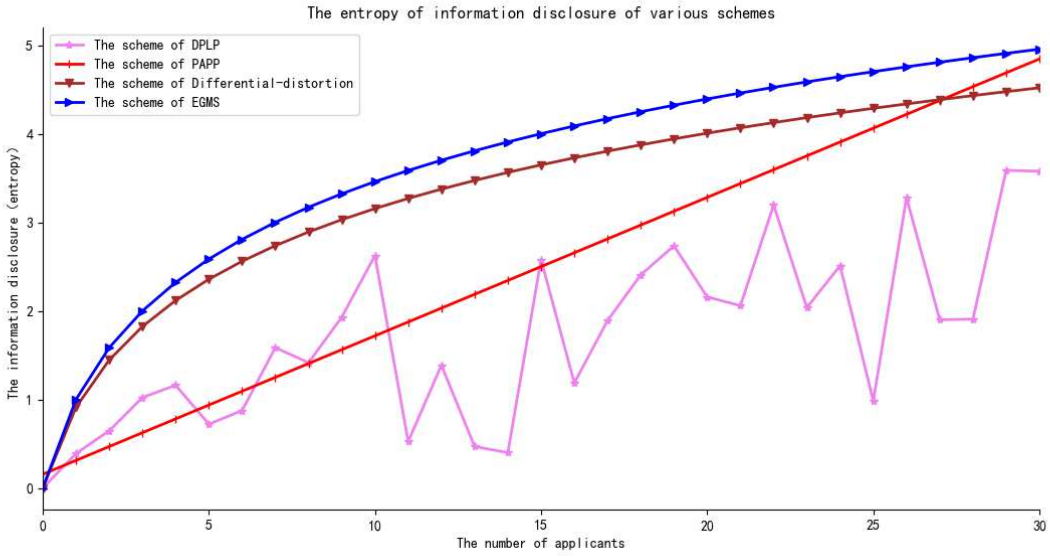

Fig.6 The entropy of information disclosure of various schemes

In general, entropy is usually used for measuring the uncertainty of a result in an information environment. So it can be used as a metric to measure the uncertainty of an adversary guessing the precise location of a user. Suppose that the success ratio of an adversary identifies a special user is $p(i)$, and based on the equation of entropy we have $H(i) = -\sum_{i=1}^{n} p(i)\log_2 p(i)$, where $H(i)$ denotes the uncertainty of this special user. Then based on the principle of maximum entropy, we can conclude that the higher value of entropy the higher

uncertainty of the adversary identifies the real locations, the more secure of the personal privacy of the publisher and the applicants. Thus, based on the value of entropy, we measure the performance of privacy preservation with this metric and show the result of comparison in Figure 6.

Figure 6 shows the entropy of information disclosure of various schemes. From this figure, we can see that the entropy value of EGMS is the highest. This is because all process steps of this scheme are conducted in the confidential environment, and no information about each entity is released to any entity, so that the uncertainty of the adversary is the highest and the information disclosure is the lowest one among these schemes. For other schemes (such as the scheme with dummy addition, the scheme with location transformation, the scheme with location distortion), as they are derived from the conception of generalization or the conception of obfuscation, dummies or distortion locations are used to achieve above operations. But dummies or distortion locations will bring in attribute correlation, and the correlation will be used by the adversary to increase the success ratio of identifying the real location. As a result, these schemes have lower values of entropy than EGMS and performance worse than EGMS, which in turn provides a lower level of privacy preservation for the publisher and the applicant.

In conclusion, based on comparisons with other schemes, we can conclude that EGMS performs better than other similar schemes, no matter in the performance of execution efficiency or in the performance of effectiveness of privacy preservation, so it can be used in the actual environment.

5 Conclusion and future works

In order to cope with the privacy problem in the task allocation in crowdsensing, in this paper, an encrypted grids matching scheme that based on the Paillier encryption system has been proposed. We first elaborate the implementation of our proposed scheme with three algorithms conducted by three entities in the crowdsensing respectively. Then we analyze the security of the proposed scheme by the perspective of the publisher and the perspective of the applicant, which further illustrates the effectiveness of our proposed scheme. At last, several groups of comparison experiments are given, and results with reasons analysis are also given to further verify the superiority of our proposed scheme.

Though the scheme mentioned in this paper can provide privacy preservation service for the publisher and the applicants, this scheme still has some inherent problems not resolved effectively. For example, if there are fewer applicants than the request of task allocation, how many times and how long is the interval of resubmitting the same request? If the applicant sends a fake result, what kind of punishment is the authorized agency can implement to this applicant? So the future works will be focused on how to solve the problems mentioned above.

Acknowledgments

We would like to present our thanks to anonymous reviewers for their helpful suggestions.This work was supported by the Natural Science Foundation of Heilongjiang Province of China under Grant LH2020F050, National Natural Science Foundation of China(No. 61872204). Science Research project of Basic scientific research business expenses in Heilongjiang Provincical colleges and universities of China (No. 135309453).

Declarations

Funding:the Natural Science Foundation of Heilongjiang Province of China under Grant LH2020F050, National Natural Science Foundation of China(No. 61872204). Science Research project of Basic scientific research business expenses in Heilongjiang Provincical colleges and universities of China (No. 135309453).

Conflicts of interest/Competing interests:no

Availability of data and material:no

Code availability :no

Authors' contributions:Xd.Zheng gave the idea, L.Zhang. and B.Wang did the experiments, Xd.Zheng. and Q.Yuan. interpreted the results, Xd.Zheng wrote the paper.

References

[1] L. Zhang, D. Liu, M. Chen, H. Li, C. Wang, Y. Zhang, and Y. Du, "A user collaboration privacy protection scheme with threshold scheme and smart contract," *Information Sciences,* vol. 560, pp. 183-201, 2021/06/01/, 2021.

[2] P. Huang, X. N. Zhang, L. K. Guo, and M. Li, "Incentivizing Crowdsensing-Based Noise Monitoring with Differentially-Private Locations," *Ieee Transactions on Mobile Computing,* vol. 20, no. 2, pp. 519-532, Feb, 2021.

[3] L. Zhang, M. Chen, D. Liu, L. He, C. Wang, Y. Sun, and B. Wang, "A ε-sensitive indistinguishable scheme for privacy preserving," *Wireless networks,* vol. 26, no. 07, pp. 5013–5033, 2020.

[4] I. J. Vergara-Laurens, L. G. Jaimes, and M. A. Labrador, "Privacy-Preserving Mechanisms for Crowdsensing: Survey and Research Challenges," *Ieee Internet of Things Journal,* vol. 4, no. 4, pp. 855-869, Aug, 2017.

[5] Y. Wang, Z. Cai, X. Tong, G. Yang, and G. Yin, "Truthful incentive mechanism with location privacy-preserving for mobile crowdsourcing systems " *Computer Networks,* vol. 135, pp. 32-43, 2018.

[6] L. Zhang, J. Li, S. Yang, and B. Wang, "Privacy Preserving in Cloud Environment for Obstructed Shortest Path Query," *Wireless Personal Communications,* vol. 96, no. 2, pp. 2305-2322, 2017.

[7] C. Zhao, S. S. Yang, and J. A. McCann, "On the Data Quality in Privacy-Preserving Mobile Crowdsensing Systems with Untruthful Reporting," *Ieee Transactions on Mobile Computing,* vol. 20, no. 2, pp. 647-661, Feb, 2021.

[8] V. Sadhu, S. Zonouz, V. Sritapan, and D. Pompili, "CollabLoc: Privacy-Preserving Multi-Modal Collaborative Mobile Phone Localization," *Ieee Transactions on Mobile Computing,* vol. 20, no. 1, pp. 104-116, Jan, 2021.

[9] J. Shu, X. Jia, Y. Kan, and W. Hua, "Privacy-Preserving Task Recommendation Services for Crowdsourcing," *IEEE Transactions on Services Computing,* vol. PP, no. 99, pp. 1-1, 2018.

[10] Y. H. Zhang, M. Li, D. J. Yang, J. Tang, G. L. Xue, and J. Xu, "Tradeoff Between Location Quality and Privacy in Crowdsensing: An Optimization Perspective," *Ieee Internet of Things Journal,* vol. 7, no. 4, pp. 3535-3544, Apr, 2020.

[11] L. Zhang, S. Yang, J. Li, and L. Yu, "A Particle Swarm Optimization Clustering-Based Attribute Generalization Privacy Protection Scheme," *Journal of Circuits, Systems and Computers* vol. 27, no. 11, pp. 641-654, 2018.

[12] J. Wei, Y. Lin, X. Yao, and J. Zhang, "Differential Privacy-based Location Protection in Spatial Crowdsourcing," 2019.

[13] G. C. Luo, K. Yan, X. Zheng, L. Tian, and Z. P. Cai, "Preserving adjustable path privacy for task acquisition in Mobile Crowdsensing Systems," *Information Sciences,* vol. 527, pp. 602-619, Jul, 2020.

[14] J. Xu, B. J. Cui, R. S. Shi, and Q. L. Feng, "Outsourced privacy-aware task allocation with flexible expressions in crowdsourcing," *Future Generation Computer Systems-the International Journal of Escience,* vol. 112, pp. 383-393, Nov, 2020.

[15] S. H. Zou, J. W. Xi, H. G. Wang, and G. A. Xu, "CrowdBLPS: A Blockchain-Based Location-Privacy-Preserving Mobile Crowdsensing System," *Ieee Transactions on Industrial Informatics,* vol. 16, no. 6, pp. 4206-4218, Jun, 2020.

[16] X. J. Zhu, E. Ayday, and R. Vitenberg, "A Privacy-Preserving Framework for Outsourcing Location-Based Services to the Cloud," *Ieee Transactions on Dependable and Secure Computing,* vol. 18, no. 1, pp. 384-399, Jan, 2021.

[17] M. M. Yang, T. Q. Zhu, Y. Xiang, and W. L. Zhou, "Density-Based Location Preservation for Mobile Crowdsensing With Differential Privacy," *Ieee Access,* vol. 6, pp. 14779-14789, 2018.

[18] M. Yang, T. Zhu, K. Liang, and W. Zhou, "A blockchain-based location privacy-preserving crowdsensing system," *Future Generation Computer Systems-the International Journal of Escience,* vol. 94, pp. 408-418, May, 2019.

[19] Y. Y. He, J. B. Ni, B. Niu, F. H. Li, and X. M. Shen, "Privbus: A privacy-enhanced crowdsourced bus service via fog computing," *Journal of Parallel and Distributed Computing,* vol. 135, pp. 156-168, Jan, 2020.

[20] L. Y. Wang, D. Q. Zhang, D. Q. Yang, B. Y. Lim, X. Han, and X. J. Ma, "Sparse Mobile Crowdsensing With Differential and Distortion Location Privacy," *Ieee Transactions on Information Forensics and Security,* vol. 15, pp. 2735-2749, 2020.

[21] D. Yuan, Q. Li, G. L. Li, Q. Wang, and K. Ren, "PriRadar: A Privacy-Preserving Framework for Spatial Crowdsourcing," *Ieee Transactions on Information Forensics and Security,* vol. 15, pp. 299-314, 2020.

[22] Z.J. Zhao, Z.B. Ying, Z. Yang, X.M. Liu, and J.F. Ma, "Recommendation of Platoon Members by

Combining the Blockchain and Vehicular Social Network,"*Journal of Xidian University,* vol. 47, no. 5, 2020.