Hindawi
Mobile Information Systems
Volume 2018, Article ID 3682382, 11 pages
https://doi.org/10.1155/2018/3682382

Research Article

Location Privacy Protection Research Based on Querying Anonymous Region Construction for Smart Campus

Ruxia Sun,[1] **Jinwen Xi**,[1] **Chunyong Yin**,[1] **Jin Wang**,[2] **and Gwang-jun Kim**[3]

[1]*School of Computer and Software, Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing, China*
[2]*School of Computer & Communication Engineering, Changsha University of Science & Technology, Changsha, China*
[3]*Department of Computer Engineering, Chonnam National University, Gwangju, Republic of Korea*

Correspondence should be addressed to Gwang-jun Kim; kgj@jnu.ac.kr

Received 15 June 2018; Revised 20 July 2018; Accepted 1 August 2018; Published 16 September 2018

Academic Editor: Jaegeol Yim

Along with the rapid development of smart campus, the deployment of novel learning applications for smart campus requires full consideration of information security issues. Location privacy protection is one of the most important issues, which considers the privacy protection and guarantees the quality of service. The existing schemes did not consider the area of the querying regions for location-based service provider (LSP) during the construction of the anonymous regions, which led to the low quality of service. To deal with this problem, the user's query range was introduced to present a novel anonymous region construction scheme. In the proposal, the anonymous server first generated the original anonymous subregions according to the user's privacy requirements, and then merged these subregions to construct the anonymity regions submitted to LSP based on the size of corresponding querying regions. The security and experiment analysis show that the proposed scheme not only protects the user's privacy effectively but also decreases the area of LSP querying regions and the region-constructing time, improving the quality of service for smart campus.

1. Introduction

With the development and construction of novel learning applications for smart campus, smart devices and services, smart meters, smart terminals, and the like are widely applied to offer real-time learning feedback to students through continuously monitoring and analyzing the status and activities of students with various devices and platforms. As a large number of smart meters and intelligence appliances are accessed, incorporating various technologies and enabling world-changing learning, the network border further extends to the user side. Security risks on the user side for smart campus will become more and more prominent. Data privacy issues, especially location privacy protection and the quality of service, must be considered [1].

Users' location privacy threats refer to the risks that an attacker can obtain unauthorized access to raw location data by locating a transmitting device and identifying the subject (person) using it. Examples of such risks include spamming users with unwanted advertisements, drawing sensitive inferences from victims' visits to various locations (e.g., students and teachers' offices), and learning sensitive information about them (identity, religious and political affiliations, etc.). Hence, location privacy protection for smart campus is becoming a critical issue [2].

However, location information is consistently sent to service providers without protection when users query LBSs, allowing location providers to collect location information from all users. The collected location information may expose users to customized advertisement or even be sold to third parties. A worse scenario is location information may be leaked to adversaries with criminal intents. Therefore, many researchers focus on creating location protection algorithms to protect the location privacy of users [3].

The European Union's Information Protection Supervision Organization recently said that high-tech equipment

such as smart meters that monitor household energy consumption will pose a huge threat to personal privacy. Smart meters may track personal information, and the vast amounts of information collected can have serious consequences for consumers.

With the popularization of mobile devices and location technology, location-based services (LBSs) are widely used in real life, which refers to the user accesses to its designated location information query and entertainment services through the mobile device [4]. However, the location-based service provider (LSP) may also collect and abuse user's service information while providing the convenient LBS for the user, to illegally obtain the user's confidential information. The location privacy protection in LBS has attracted the extensive attention of researchers [5–8].

In view of the popularization of mobile positioning devices, if these novel learning applications are used in smart campus, the combination of location information and services at different moments in personal privacy may reveal sensitive information such as the user's behavior habits and work nature. For example, if a mobile user is collected near a hospital, the user may be presumed to have any disease or health condition. If the user's starting location and ending location are analyzed since the last few days, the user's home address or work unit, the nature of work, and so on can be speculated. Therefore, the location data of mobile object brings convenience to people and brings about the threat of revealing privacy, which may contain other sensitive information such as home address, personal preference, personality habit, health status, working property, personal income, etc. If this information falls into the hands of normal institutions, it is a tool for information protection, if it falls into the hands of illegal institutions, it will be the weapon of innocent destruction. What we can do is to seek transparency in the use of personal information and to protect user's location privacy not to be exploited by unscrupulous businesses and illegal agencies.

As the most commonly used LBS privacy protection method, the basic idea of k-anonymity [9] is that when a user sends a LBS query, he will send his real location and querying content to a trusted anonymous server, then the anonymous server will remove the user's identification information and generate anonymity regions containing other $k-1$ users for the real location, finally sending them (all the generated anonymity regions and the real location) along with the querying content to the LSP. Compared with other LBS privacy protection methods (such as pseudo location [10], fuzzification [11], differential privacy [12], and cryptography-based methods [13, 14]), k-anonymity has the following advantages: (1) users can get accurate querying results; (2) the user's cost of computing and communication is small; and (3) this method can confuse the relevance between users and LBS queries. So, k-anonymity is widely used in LBS privacy protection [15, 16].

In smart campus, we can query the nearby points of interest. There are many sensors to be addressed for enabling such novel learning applications and services, which aims to enhance the service quality of novel learning applications. When k-anonymity method is used to protect the privacy of users in the above query, if the anonymous region generated by the anonymous server is too large, the querying cost of the location-based service provider (LSP) will increase and the service quality deteriorates [17–19]. To solve this problem, the existing methods [20–26] obtain n disjoint anonymous subregions by removing the part that does not contain the users in the regions, to reduce the area of anonymous regions and improve the service quality as shown in Figure 1. However, the quality of service in LBS queries based on k-anonymity is not only related to the size of the anonymous regions but also to the user's query range. If we use the existing methods to construct the anonymous regions, the quality of service cannot be effectively improved. As shown in Figure 2, when using the existing division methods to divide the initial anonymous region, LSP will repeatedly search the points of interest in some regions, reducing the quality of service, and r is the query radius. This paper also proves this point through experiments.

This paper proposes the LBS privacy protection scheme based on querying anonymous region construction. In which, firstly, the anonymous server generates k initial anonymous subregions according to the privacy protection requirements of users and merges the anonymous regions according to the corresponding querying regions so that the anonymous regions finally submitted to the LSP can reduce the querying cost of LSP and improve the service quality without reducing the user's privacy protection level. This is the first k-anonymity privacy protection scheme based on constructing the user querying anonymous regions. The main contributions of this paper are as follows:

(1) Based on the theoretical analysis, it is concluded that the existing anonymous region demarcation method cannot reduce the LSP querying cost and improve the service quality, and we prove it through experiments.

(2) We think that the area of querying regions is the judgment criterion to merge the anonymous subregions and propose an anonymous region construction scheme based on the user's query range. Security analysis shows that the proposed scheme can effectively protect the users' location privacy.

(3) A large number of experiments show that this scheme can effectively reduce the querying cost of LSP and improve the service quality, without causing a large computational cost for anonymous servers.

2. Related Work

At present, the international research on the privacy protection for smart campus is still in the initial stage. The discussion on the privacy protection focuses more on the risk analysis of exposing personal privacy for the wireless applications and devices. Research in the United States is the leader, which publishes relevant documents on this issue.

Privacy laws in the United States do not explicitly address the smart campus and its related data, which is same as the regulations of the existing national Internet of Energy

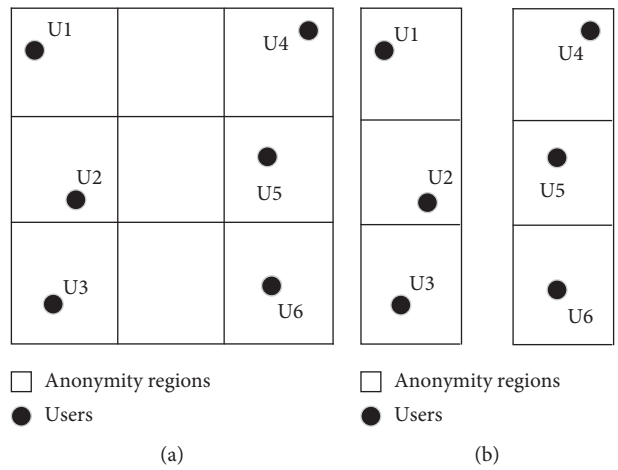

Figure 1: The existing anonymous regions division method. (a) Initial anonymity regions. (b) Anonymity regions after division.

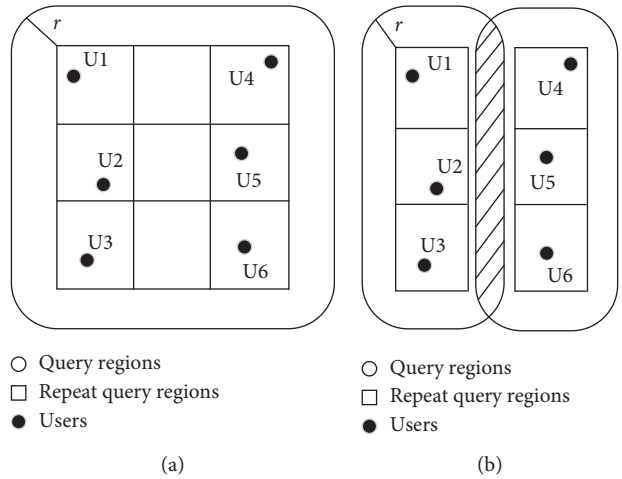

Figure 2: The querying regions of nearby points of interest. (a) Initial query regions. (b) Query regions after division.

and power transmission [27]. The existing laws and regulations need to be revised to appeal to the smart campus. At the same time, new data items in the Internet of Energy, as well as new ways of using existing data, require more research and public opinion to adapt to current laws or shape new laws.

US Internet of Energy Information Security is very concerned about privacy issues. NIST (National Institute of Standards and Technology) released "Guidelines for Smart Grid Cyber Security: Vol. 2, Privacy and Smart Grid" [28] in August 2010 and preliminarily analyzed the privacy issues of smart campus. For location privacy protection, many researchers have done a great deal of work.

Grutese applied k-anonymity to the field of location privacy protection for the first time [4]. The anonymous server divides the area with a quadtree structure and stores the users in the corresponding region in the node. When the user requests a service, a quadtree is retrieved upward from a leaf node corresponding to the user's location. If the current leaf node does not satisfy the k-anonymity, the parent node will be retrieved until no less than other $k-1$ users to obtain the anonymous region. However, if the number of users in a leaf node does not satisfy the privacy requirements, it needs to retrieve its parent node, which will result in a four-fold increase in the anonymity area, thereby degrading the user's service quality. To deal with the above problem, Mokbel et al. [29] improved the anonymous region construction method in [4]. In their scheme, if the current leaf node does not satisfy k-anonymity, firstly its sibling nodes will be retrieved, if it still does not satisfy the user's privacy requirements, then it retrieves the parent node. To further solve the problem of low quality of service due to the large anonymous region based on the quadtree structure, Li et al. [30] proposed to reduce the area of anonymous regions and improve the service quality by suppressing part of users' requests and deleting the most distant trails.

Subsequently, the researchers proposed Clique Cloak [31] and Hilbert Cloak [32], respectively, to construct the anonymous regions by searching for the nearest users who meet the privacy requirements. In [31], users can customize the privacy protection requirements, but the scheme uses an undirected graph to construct an anonymous region and the

cost of the solution is too large. It may happen that the anonymous region has not been successfully constructed but beyond the anonymous period. In [32], the anonymous server maps all users in two-dimensional space to a one-dimensional array according to the Hilbert curve and divides the users into several sets according to the value of k. When a user makes a service request, the anonymous region is constructed by using the user set to which the user belongs.

Recently, Jin et al. [2] thought the anonymization server could lead to a new class of attacks called location injection attacks which can successfully violate users' in-distinguishability (guaranteed by k-anonymity) among a set of users. Therefore, they propose and characterize location injection attacks by presenting a set of attack models and quantify the costs associated with them. Then, they propose and evaluate k-Trustee, which is resilient to location injection attacks and guarantees a low bound on the user's in-distinguishability. The experimental results show that the proposed cloaking algorithm guaranteeing k-Trustee is effective against various location injection attacks, but the quality of location services is poor and cannot be guaranteed.

In all the above solutions, if the k users used to construct the anonymous regions are far away from each other, the anonymous regions may still be too large, and the service quality may be low. Tan et al. [33] were the first ones to apply the idea of regionalization to the construction of anonymous regions, which divided the users in the anonymous regions into separate groups through the Hilbert space filling curve. When a user makes a server request, the anonymous server will use the locations of other users in the group to which it belongs to construct the anonymous region. Subsequently, Li and Zhu [34] also used the method of regionalization to research on reducing the area of anonymous regions and improving the quality of service. The anonymous server firstly constructs an anonymous region containing k users and then removes the anonymous regions that do not contain the users according to relationship among the users' locations to form multiple anonymous subregions that do not intersect each other.

However, the existing anonymous region construction schemes [35–37] ignore the impact of the users' query range on LBS querying service quality. When using these methods to construct the anonymous regions, the LSP will repeatedly search the points of interest in the partial regions, reducing the service quality.

3. Improved k-Value Location Privacy Protection Method

3.1. System Structure. This paper uses the centralized system architecture [38], composed of three parts: user, anonymous server, and LSP. When the user requests a service, the anonymous server blurs the real location of the user into an anonymous region and sends it to the LSP, which contains not only the real user but also other at least $k-1$ users. In this case, the correct rate that LSP correlates the service query with the right user will be not more than $1/k$, which achieves k-anonymity. The system structure is shown in Figure 3.

Assume that there is a secure communication channel between the user and the anonymous server. When the user queries the point of interest nearby, the secure channel is used to send the query request $q = \langle \mathrm{ID}, L(x, y), r, \mathrm{POI}, p\rangle$ to the trusted anonymous server. In which, ID represents the identity of the user; $L(x, y)$ represents the location coordinates of the user; r represents the radius of the user query; POI represents the point of interest of the user query; p represents the privacy protection requirement of the user current query; k represents the anonymous region generated by the anonymous server contains at least other $k-1$ users; and $A_{\min}$ represents the minimum area of anonymous regions generated by anonymous server.

After receiving the user's request, the trusted anonymous server will determine the identity through authentication and find other $k-1$ users to generate anonymous regions that the area is not less than $A_{\min}$ according to the user's privacy protection requirements $p = (k, A_{\min})$ and then send the query request $Q = \langle \mathrm{CR}, r, \mathrm{POI}\rangle$ obtained from the anonymization to the semitrusted LSP. CR represents the anonymous region generated by the anonymous server for the current user making the service query.

After receiving the anonymous query request sent by the anonymous server, the LSP will search in the database and return all the query candidate results to the anonymous server. After the anonymous server receives the query result from the LSP, it selects the query result according to the location $L(x, y)$ of the user, and finally returns the accurate query result to the user.

In this system model, we treat the LSP directly as an attacker. The attacking purposes are as follows: (1) the real location of the user could be identified in the anonymous region sent from the anonymous server and (2) the real user will be speculated from the query request.

In addition, in the above model, LBS query quality of service is mainly affected by the following four factors: (1) the time required by the anonymous server to generate anonymous regions; (2) the time that the anonymous server sends the anonymous query request to LSP; (3) the time required by the LSP to retrieve the database for the anonymous query request sent by the anonymous server; and (4) the time it takes for the LSP to send the retrieval result to the anonymous server, and the time required by the anonymous server to finalize the query result. Because the time taken for the LSP to send search results to the anonymous server and the time required for the anonymous server to finalize the query results is affected by the distribution of points of interest, the time required for the anonymous server to send the anonymous query request to the LSP is affected by the transmission bandwidth. Therefore, this paper evaluates the quality of service based on k-anonymous LBS queries only by the time it takes the anonymous server to generate the anonymous regions and the LSP to retrieve the database.

3.2. Querying Region of LSP. After receiving the anonymous query request $Q = \langle \mathrm{CR}, r, \mathrm{POI}\rangle$ sent by the anonymous server, the LSP first calculates the querying region QAR according to the anonymous region CR and the query radius

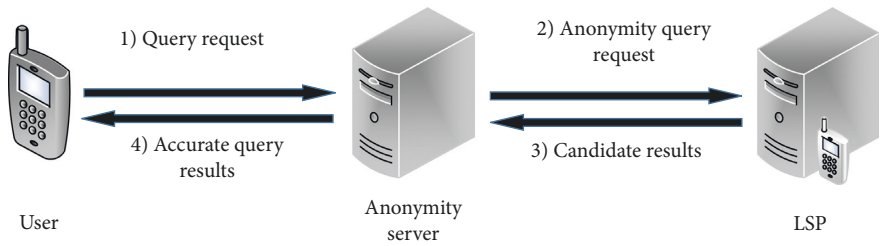

Figure 3: System structure.

r and then searches the interest point of the user query in the region QAR. The querying regions corresponding to the anonymous regions are as shown in Figure 4.

When the anonymous region is a circle, the area S_c of the querying region is

$$S_c = \pi(\tilde{r} + r)^2, \tag{1}$$

where $\tilde{r}$ represents the radius of the circular anonymous region and πr^2 is the area of the anonymous region.

When the anonymous region is a rectangle, the area S_r of the querying region is

$$S_r = ab + 2(a + b)r + \pi r^2, \tag{2}$$

where a, b is the side length of the rectangular anonymous region and ab is the area of the anonymous region CR.

It can be learned from the above analysis that after receiving the anonymous query request sent by the anonymous server, the time required by the LSP to retrieve the database is not only related to the size of the anonymous region CR generated by the anonymous server but also to the radius of the user's query, in other words, it is determined by the querying region. However, the existing anonymous region partitioning schemes only take the size of the anonymous regions into account, which makes these solutions not effectively improve the LBS query service quality and even further reduces the service quality.

4. Anonymous Region Construction Scheme Based on Query Range

In this paper, the anonymous server first generates k initial anonymous subregions according to the privacy protection requirements of the users and constructs the corresponding querying regions by calculating and then merges the anonymous subregions to finally obtain the anonymous subregion set. The solution can reduce the LSP query cost and improve the service quality without lowering the user privacy protection level.

4.1. Generation of Initial Anonymous Subregions. After receiving the service request sent by the users, the anonymous server will start to search the other $k-1$ users according to the user's privacy protection requirement $p = (k, A_{\min})$ and obtain their location information $L_1(x_1, y_1), \ldots, L_{k-1}(x_{k-1}, y_{k-1})$. Then, the server will generate k disjoint initial anonymous subregions $\mathrm{AR}_0, \mathrm{AR}_1, \ldots, \mathrm{AR}_{k-1}$ to satisfy

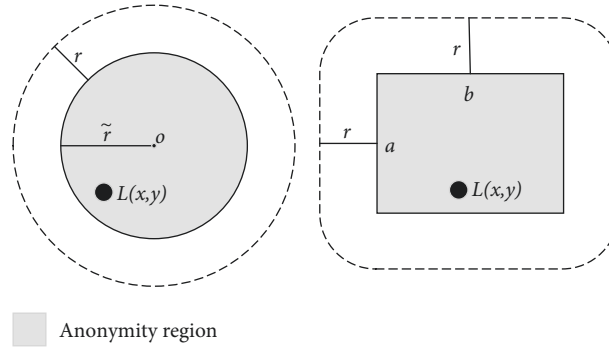

Figure 4: Querying region of LSP.

$$\begin{cases} \mathrm{centre}(\mathrm{AR}_i) \neq L_i(x_i, y_i), \\ \mathrm{area}(\mathrm{AR}_i) \neq A_{\min}, \end{cases} \tag{3}$$

where AR_i represents the initial anonymous subregion containing the i-th location $L_i(x_i, y_i)$, $0 \le i \le k-1$; $L_0(x_0, y_0)$ represents the location of the user who sent the query request; $\mathrm{centre}(\mathrm{AR}_i)$ represents the central location of initial anonymous subregion AR_i; and $\mathrm{area}(\mathrm{AR}_i)$ represents the area of initial subanonymous region AR_i.

4.2. Merger of Anonymous Subregions. After generating the k initial anonymous subregions $\mathrm{AR}_0, \mathrm{AR}_1, \ldots, \mathrm{AR}_{k-1}$, the anonymous server will calculate the area $\mathrm{area}(\mathrm{QAR}_0)$, $\mathrm{area}(\mathrm{QAR}_1), \ldots, \mathrm{area}(\mathrm{QAR}_{k-1})$ of the corresponding querying regions $\mathrm{QAR}_0, \mathrm{QAR}_1, \ldots, \mathrm{QAR}_{k-1}$, respectively, according to the query radius of the user and judge whether the anonymous subregions need to be merged on the basis of the area of the querying regions. In order to effectively reduce the cost, LSP retrieves the points of interest, and after the merger of anonymous subregions, it should ensure

$$S_{\min} = \min \sum_{i=0}^{l} \mathrm{area}(\mathrm{QAR}'_i). \tag{4}$$

That is to say, the sum of the area of each querying region QAR'_i corresponding to each anonymous subregion AR'_i in the anonymous regions CS is the smallest. AR'_i represents the i-th anonymous subregion after the merger of anonymous subregions, $0 \le i \le l$, $0 \le l \le k-1$.

The merger process of anonymous subregions in the anonymous server is as follows:

(1) Filtering the anonymous subregions that require region consolidation

To ensure that the querying area of the anonymous region is minimized, the anonymous server chooses the anonymous subregions AR_i and AR_j to merge if the area of querying region is the smallest after the merger of them. The selected anonymous subregions AR_i and AR_j are needed to satisfy

$$\begin{aligned} &\forall i, j \in [0, k-1], \mathrm{QAR}_{i,j} = \operatorname{argmin}\left\{\operatorname{area}\left(\mathrm{QAR}_{i,j}\right)\right\}_{i \neq j}, \\ &\quad \operatorname{centre}\left(\mathrm{QAR}_{i,j}\right) \neq L_i(x_i, y_i), \\ &\quad \operatorname{centre}\left(\mathrm{QAR}_{i,j}\right) \neq L_j\left(x_j, y_j\right). \end{aligned} \tag{5}$$

(2) Selecting anonymous subregions to merge

For the two anonymous subregions AR_i and AR_j, $AR_{i,j} = \mathrm{gen}(AR_i, AR_j)$ represents the new anonymous subregion formed by merging the anonymous subregion AR_i with AR_j and the corresponding querying region is $QAR_{i,j}$. If $\mathrm{area}(QAR_i) + \mathrm{area}(QAR_j) \leq \mathrm{area}(QAR_{i,j})$, it does not merge AR_i with AR_j. Else it merges AR_i with AR_j to make the new anonymous subregion $AR_{i,j}$.

(3) Repeating the process until there is no need to merge anonymous subregions. In this case, the anonymous server will obtain the anonymous set

$$\mathrm{CS} = \{\mathrm{AR}'_0, \mathrm{AR}'_1, \ldots, \mathrm{AR}'_l\}. \tag{6}$$

In this scheme, when the anonymous server merges the anonymous subregions, it chooses the anonymous regions AR_i and AR_j to make the area of the querying region the smallest after merging. Therefore, after every merger of the anonymous subregions, the area of the new querying region $AR_{i,j}$ is the smallest, which ensures that the querying area of the final anonymous region set CS is the smallest.

In this paper, the locations of k users are used as the input to generate the corresponding anonymous subregions $AR_0, AR_1, \ldots, AR_{k-1}$ and merge them to obtain the final anonymous subregion set CS, which is submitted to the LSP (Algorithm 1).

5. Scheme Analysis and Experimental Results

5.1. Analysis of Security. In our scheme, we construct the disjoint initial anonymous subregions $AR_0, AR_1, \ldots, AR_{k-1}$ that the area of each initial anonymous subregion is less than $A_{\min}$ according to k real locations $L_0(x_0, y_0)$, $L_1(x_1, y_1), \ldots, L_{k-1}(x_{k-1}, y_{k-1})$ and ensure every real location is not located in the center of the initial anonymous subregion for smart campus, that is, $\mathrm{centre}(AR_i) \neq L_i(x_i, y_i)$. If the anonymous server directly sends the k initial anonymous subregions to the LSP, which cannot correctly identify the real location $L_i(x_i, y_i)$ of the user from the anonymous region AR_i. If the anonymous server uses the proposed scheme to merge the anonymous subregions and sends the final anonymous region set CS to the LSP, the area of every anonymous subregion AR_i satisfies $\mathrm{area}(AR_i) = A_{\min}$, so the area of every merged anonymous subregion AR'_i will satisfy no less than $A_{\min}$ and for any location $L_i(x_i, y_i) \in AR'_i$, and it satisfies that the center of AR'_i is not $L_i(x_i, y_i)$. In addition, the user's identity has been removed from the anonymous query request sent by the anonymous server. Though the LSP received the anonymous query request $Q = \langle \mathrm{CR}, r, \mathrm{POI} \rangle$, it is unable to know the user who made the service request. Therefore, this scheme proposed in this paper can effectively protect the privacy of users.

5.2. Analysis of Computational Complexity. When the anonymous server receives the service query request sent by the user and uses this solution proposed in this paper to generate an anonymous region for it, it firstly generates k initial anonymous subregions according to the privacy protection requirement of the user. The computational complexity of generating the initial anonymous subregions is $O(k)$. When the anonymous server determines whether any two anonymous subregions AR_i and AR_j need to be merged, it first needs to use the anonymous subregions AR_i and AR_j to generate a new anonymous region $AR_{i,j}$. In this case, the number of new anonymous regions that need to be generated is $C_k^2 = k(k-1)/2$, and the computational complexity is $O(k^2)$. Subsequently, the anonymous server will calculate the area of the querying region corresponding to each newly generated anonymous region $AR_{i,j}$, which requires $k(k-1)/2$ times of computation, and the computational complexity is $O(k^2)$. Finally, comparing the area of the querying region AR_i and AR_j with that of the newly generated anonymous region $AR_{i,j}$ to determine whether the both initial anonymous subregions are needed to merge, which requires $k(k-1)/2$ times of computation, and the computational complexity is $O(k^2)$. So the computational complexity of merging all the initial anonymous subregions is $O(k^2) + O(k^2) + O(k^2) = O(k^2)$. After obtaining the new anonymous region $AR_{i,j}$ merged from the initial anonymous subregions AR_i and AR_j, it also needs to judge whether the new anonymous subregions need to be merged again with other anonymous subregions. During the merging process of the anonymous subregions, the best case is that any one of k initial anonymous subregions does not need to be merged with others. In this case, the computational complexity required to implement this solution is

$$O_{\mathrm{best}} = O(k) + O\left(k^2\right) = O\left(k^2\right). \tag{7}$$

In contrast, the worst case is that any one of k initial anonymous subregions needs to be merged with others and finally merged to an anonymous region. In this case, it needs to repeat the above anonymous region merger and judge $k-1$ time, and the computational complexity required to implement this solution is

$$O_{\mathrm{worst}} = O(k) + O\left((k-1)k^2\right) = O\left(k^3\right). \tag{8}$$

5.3. Experimental Results and Analysis. In this paper, a network-based generator of moving objects (NGMO) [39] is used to generate the experimental data. This generator is often

Input: k pieces of locations $L_0(x_0, y_0), L_1(x_1, y_1), \ldots, L_{k-1}(x_{k-1}, y_{k-1})$; the query radius r; the privacy requirement $A_{\min}$;
Output: anonymous region set CS;
(1) for $i = 0$ to $k-1$ do
(2) $\mathrm{AR}_i \leftarrow \mathrm{Gen}(L_i(x_i, y_i))$;
(3) $\mathrm{centre}(\mathrm{AR}_i) \neq L_i(x_i, y_i), \mathrm{Area}(\mathrm{AR}_i) = A_{\min}$;
(4) $\mathrm{CS} \leftarrow (\mathrm{AR}_i)$;
(5) $\mathrm{QAR}_i \leftarrow \mathrm{Gen}(\mathrm{AR}_i, r)$, calculating $\mathrm{Area}(\mathrm{QAR}_i)$;
(6) end for
(7) for $i, j = 0$ to ij and $i \neq j$ do
(8) $\mathrm{AR}_{i,j} \leftarrow \mathrm{Gen}(\mathrm{AR}_i, \mathrm{AR}_j)$;
(9) $\mathrm{CR} \leftarrow \mathrm{AR}_{i,j}$;
(10) $\mathrm{QAR}_{i,j} \leftarrow \mathrm{Gen}(\mathrm{AR}_{i,j}, r)$, calculating $\mathrm{Area}(\mathrm{QAR}_{i,j})$;
(11) if $\mathrm{Area}(\mathrm{QAR}_i) + \mathrm{Area}(\mathrm{QAR}_j) \leq \mathrm{Area}(\mathrm{QAR}_{i,j})$ then
(12) $\mathrm{CS} \leftarrow \mathrm{CS} \backslash \{\mathrm{AR}_{i,j}\}$;
(13) end if
(14) if $\mathrm{Area}(\mathrm{QAR}_i) + \mathrm{Area}(\mathrm{QAR}_j) > \mathrm{Area}(\mathrm{QAR}_{i,j}), \mathrm{QAR}_{i,j} = \mathrm{argmin}\{\mathrm{Area}(\mathrm{QAR}_{i,j})\}_{i \neq j}$, $\mathrm{centre}(\mathrm{QAR}_{i,j}) \neq L_i(x_i, y_i)$ and $\mathrm{centre}(\mathrm{QAR}_{i,j}) \neq L_j(x_j, y_j)$ then
(15) $\mathrm{CS} \leftarrow \mathrm{CS} \backslash \{\mathrm{AR}_i, \mathrm{AR}_j\}$;
(16) end if
(17) end for
(18) return CS

Algorithm 1: Anonymous subregion generation algorithm based on querying region division.

used in the existing research of LBS privacy protection, which is based on the Oldenburg map of German city and simulates the location information of users by setting parameters such as the number of moving objects. We set the privacy requirement as k, the generated initial anonymous subregion is rectangular, and its area satisfies $A_{\min} = 160000\ \mathrm{m}^2$. At the same time, to assess the LSP query cost, this paper simulates 500000 points of interest, such as restaurants, hotels, hospitals, parking lots, and so on. In addition, R-tree structure is used to access these points of interest because R-tree is the best-balanced tree for storing high-dimensional data, which can effectively improve the searching efficiency in high-dimensional space [40]. The experimental environment is Intel (R) Core(TM) i7-6700HQ CPU @ 2.60 GHz, 20.0 GB RAM. The algorithms are programmed by Python and the programs run in the Windows 10 Enterprise. The experimental data sets are shown in Table 1.

5.3.1. Problems Existing in Real Anonymity Region Constructions Schemes. To prove that the existing anonymous region construction scheme cannot effectively improve the service quality of LBS query, this paper selects Casper scheme [29] and Fragment scheme [34], respectively, to search nearby points of interest. As the most commonly used anonymous region construction method, the Casper method generates at least an anonymous region that contains all k users. Fragment is to deal with the anonymous region generated by the Casper scheme to reduce the area of the anonymous regions by removing the part that does not contain the user's location according to the location of the user in the anonymous region.

This paper compares the time required to generate the anonymous region, the area of the anonymous region, and the area of the querying region of LSP in the above two schemes to prove that the existing anonymous region division methods will further reduce the quality of service when there is an overlap querying region between both anonymous regions. In this part of the experiment, this paper sets the user's query radius $r = 500$ m. The experimental results are shown in Figures 5 and 6.

As is shown in Figure 5, compared with the Casper scheme, the Fragment scheme uses a region division method to reduce the area of anonymous regions, for example, when $k = 25$, the area of anonymous regions generated by the Casper scheme is $5.71 \times 10^7\ \mathrm{m}^2$, and the time of generating the anonymous regions is 180.275 ms. The area of anonymous regions generated by the Fragment scheme is $3.41 \times 10^7\ \mathrm{m}^2$, and the time of generating the anonymous regions is 175.331 ms. However, in the Casper and Fragment schemes, the time required for the LSP to query the points of interest is, respectively, 9.940 s (the corresponding area of querying regions is $7.234 \times 10^7\ \mathrm{m}^2$) and 10.501 s (the corresponding area of querying regions is $7.329 \times 10^7\ \mathrm{m}^2$). Therefore, when the anonymous server adopts the Casper and Fragment schemes to generate the anonymous regions, the cost time when the user obtains the accurate query results (without considering the transmission delay) is $9.940 + 0.180 = 10.120$ s and $10.501 + 0.175 = 10.676$ s, respectively. When the anonymous server adopts the existing region division schemes to construct the anonymous regions, the time for the user to obtain the query result increases instead.

5.3.2. Analysis of Our Scheme's Effectiveness. In this section, we compare our proposed scheme with Casper scheme and prove that the proposed scheme can effectively reduce the query cost of the LSP and improve the query service quality in the smart campus. The experimental results are shown in Figures 5 and 6.

Table 1: Experimental data sets.

Status	Id	Reward number	Type id	Time stamp	X-axis	Y-axis	Speed
New point	105	1	0	0	4229.0	16335.0	298.0
New point	106	1	0	0	19065.0	9922.0	132.0
New point	107	1	0	0	3670.0	20230.0	298.0
New point	108	1	1	0	5565.0	18047.0	298.0
New point	109	1	0	0	10567.0	17947.0	298.0
Point	109	2	0	1	10275.4	17638.7	672.0
Point	108	2	1	1	5487.7	17812.9	504.5

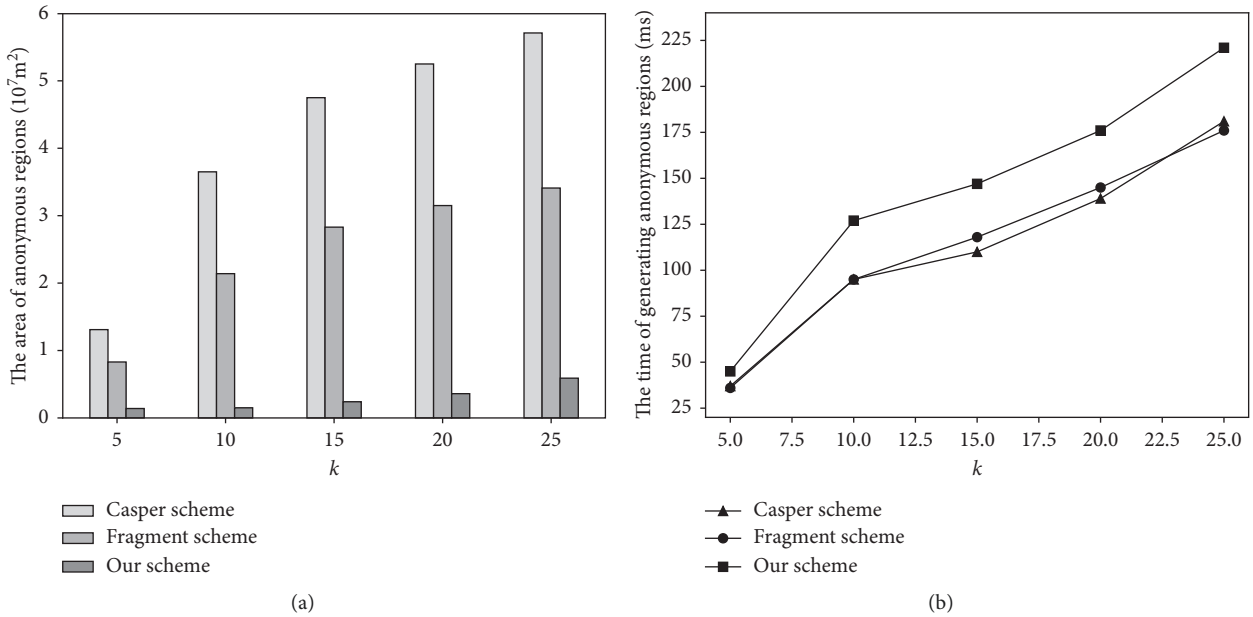

Figure 5: Computational cost of anonymity server. (a) The area of anonymous regions. (b) The query cost of LSP.

As is shown in Figures 5 and 6, compared with the existing scheme, the anonymous regions and querying regions generated by our proposed scheme significantly reduced and the same as the time required for LSP to query the points of interest. For example, when $k = 25$, the area of anonymous regions generated by our scheme is 5.89×10^6 m^2, which decreases by 2.80×10^7 m^2 compared to the Casper scheme, and the time to generate the anonymous regions increases from 180.275 ms to 221.034 ms, increasing by 40.759 ms. The area of querying regions generated by our scheme is 2.493×10^7 m^2, which decreases by 4.741×10^7 m^2 compared to the Casper scheme, and the time for query processing increases from 9.940s to 1.961s. Therefore, compared with the Casper scheme, our scheme can effectively improve LBS query service quality.

As can been seen from Figures 5(b) and 6(b), compared with the Casper scheme, as the value of k increases, our scheme needs more extra time to construct the anonymous regions (the time difference between our scheme and the Casper scheme), but less time for query processing. At the same time, the time that the anonymous server generates the anonymous regions is in the milliseconds level, which has little impact on the user delay. The time of query processing is in the second level, which greatly affects the user delay. The query processing time of the LSP determines the user's service quality to a great extent.

The paper also has made a brief analysis about the influence of the user-specified query radius on the scheme. The experimental results are shown in Figures 7(a) and 7(b). As the user query radius r increases, the number of anonymous subregions that need to be merged also increases, increasing the computational cost of the anonymous server. As a result, the area of querying regions and the time of LSP query processing also increase, which obviously influences the LBS query service quality.

In summary, the proposed scheme not only reduces the computational cost of the anonymous server but also significantly reduces the area of the LSP querying regions and its query time, which effectively improves the LSB query service quality for smart campus.

6. Conclusions

In this paper, we merged novel learning applications technology with mobile technology to research on the location privacy protection technology for smart campus and proposed a new location privacy protection scheme based on querying anonymous region construction which faces the

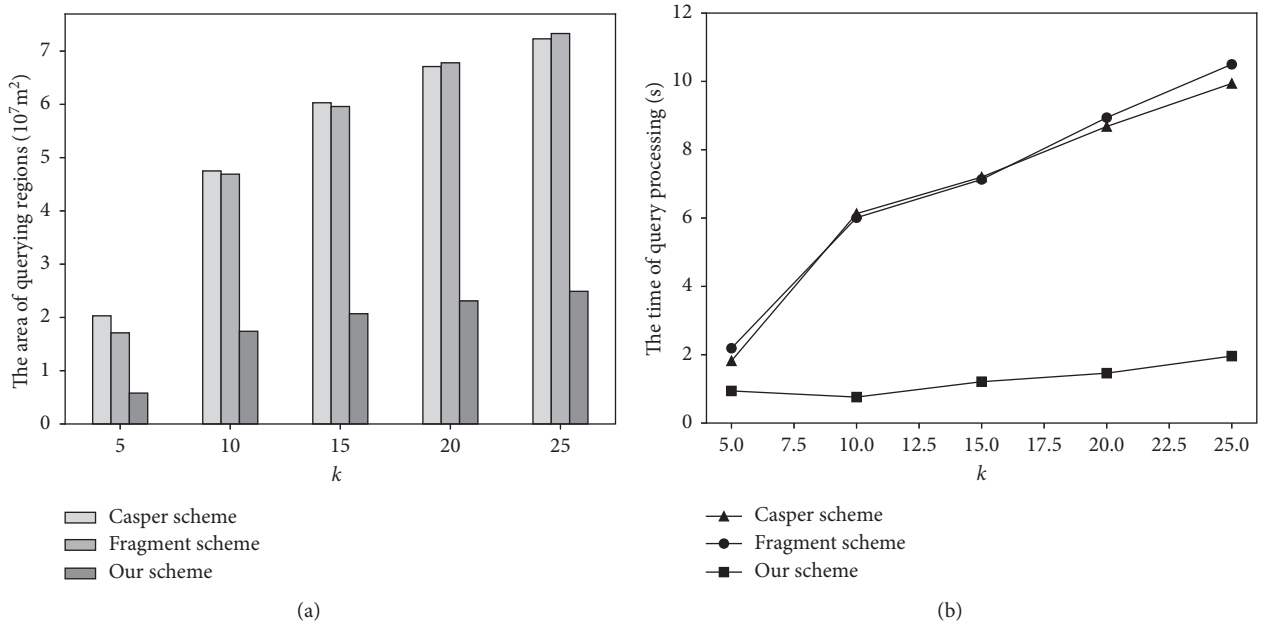

Figure 6: The area of querying regions and query cost of LSP. (a) The area of querying regions. (b) The query cost of LSP.

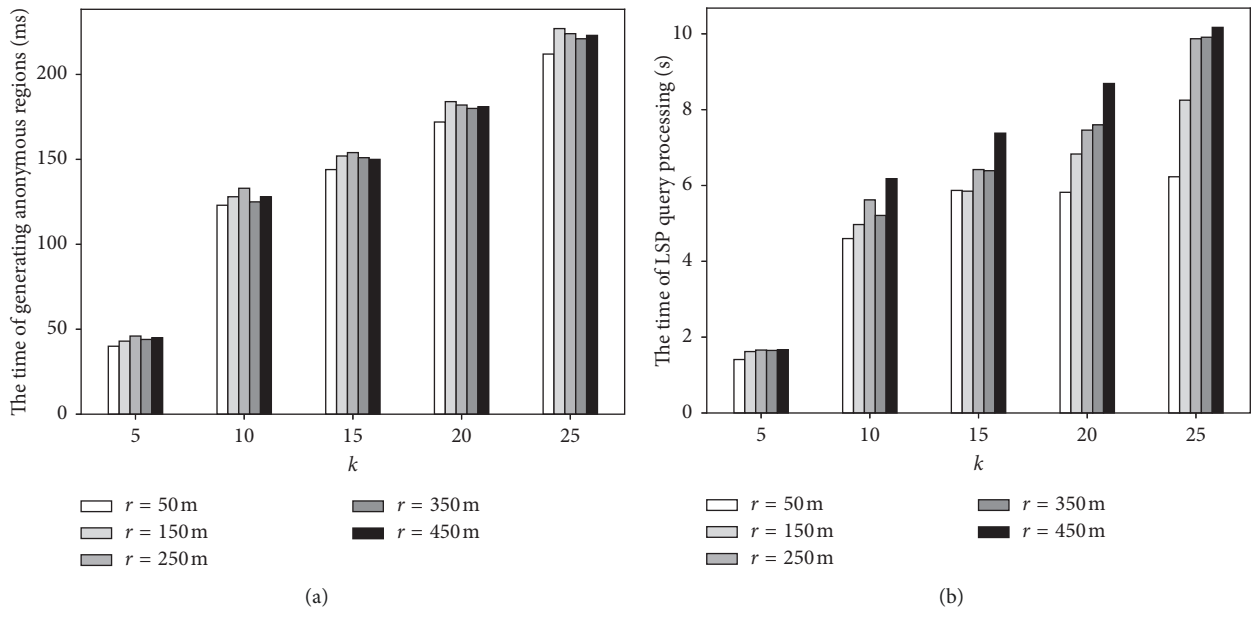

Figure 7: The impact of user query radius on the scheme of this paper. (a) The time of generating anonymous regions. (b) The time of LSP query processing.

challenges to achieve the higher quality of smart campus services. Through theoretical analysis and experimental results in the location privacy protection based on k-anonymity, this paper proves that the existing anonymous region construction schemes based on the idea of region division cannot effectively improve the LBS query service quality. The root cause of this problem is that the service quality of sensors in the smart campus is not only related to the anonymous area size constructed by the anonymous server but also to the range of the LSP query. To deal with the above problem, this paper introduces the user's query range into the construction process of the anonymous region. The anonymous server first generates k initial anonymous subregions based on the user's privacy protection requirement, and then the anonymous subregions will be judged if it needs to be merged based on the size of the querying region. The scheme analysis shows that our proposed scheme can effectively reduce the query cost of LSP and improve the service quality while protecting user's location privacy for smart campus.

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 funded by the National Natural Science Foundation of China (61772282, 61772454, and 61811530332). It was also supported by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX18_1032), Natural Science Foundation of Jiangsu Province (BK20150460), and Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (CICAEET). It was also funded by the open research fund of Key Lab of Broadband Wireless Communication and Sensor Network Technology (Nanjing University of Posts and Telecommunications), Ministry of Education.

References

[1] J. S. Turner, "New directions in communications," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 1, pp. 11–23, 1995.

[2] L. Jin, C. Li, B. Palanisamy, and J. Joshi, "k-trustee: location injection attack-resilient anonymization for location privacy," *Computers & Security*, vol. 78, pp. 212–230, 2018.

[3] Y. Huang, Z. Cai, and A. G. Bourgeois, "Search locations safely and accurately: a location privacy protection algorithm with accurate service," *Journal of Network and Computer Applications*, vol. 103, pp. 146–156, 2018.

[4] S. Xiao, "Consideration of technology for constructing Chinese smart grid," *Automation of Electric Power Systems*, vol. 9, no. 33, pp. 1–4, 2009.

[5] T. Peng, Q. Liu, and G. Wang, "Enhanced location privacy preserving scheme in location-based services," *IEEE Systems Journal*, vol. 11, no. 1, pp. 219–230, 2017.

[6] X. Chen and Y. Mu, "Preserving user location privacy for location-based service," in *Proceedings of 11th International Conference on Green, Pervasive, and Cloud Computing (GPC 2016)*, pp. 290–300, Xi'an, China, May 2016.

[7] R. Schlegel, C. Y. Chow, Q. Huang, and D. S. Wong, "User-defined privacy grid system for continuous location-based services," *IEEE Transactions on Mobile Computing*, vol. 14, no. 10, pp. 2158–2172, 2015.

[8] K. G. Shin, X. Ju, Z. Chen, and X. Hu, "Privacy protection for users of location-based services," *Wireless Communications IEEE*, vol. 19, no. 1, pp. 30–39, 2012.

[9] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st International Conference on Mobile Systems, Applications and Services*, pp. 31–42, ACM, San Francisco, CA, USA, May 2003.

[10] B. Niu, Z. Zhang, X. Li, and H. Li, "Privacy-area aware dummy generation algorithms for location-based services," in *Proceedings of 2014 IEEE International Conference on Communications (ICC)*, pp. 957–962, IEEE, Sydney, NSW, Australia, June 2014.

[11] X. Shu, D. Yao, and E. Bertino, "Privacy-preserving detection of sensitive data exposure," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 1092–1103, 2015.

[12] C. Yin, J. Xi, R. Sun, and J. Wang, "Location privacy protection based on differential privacy strategy for big data in industrial Internet-of-Things," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3628–3636, 2017.

[13] Z. Fu, F. Huang, K. Ren, J. Weng, and C. Wang, "Privacy-preserving smart semantic search based on conceptual graphs over encrypted outsourced data," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 8, pp. 1874–1884, 2017.

[14] Z. Fu, K. Ren, J. Shu, X. Sun, and F. Huang, "Enabling personalized search over encrypted outsourced data with efficiency improvement," *IEEE Transactions on Parallel & Distributed Systems*, vol. 27, no. 9, pp. 2546–2559, 2016.

[15] C. Yin, S. Zhang, J. Xi, and J. Wang, "An improved anonymity model for big data security based on clustering algorithm," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 7, pp. 1–13, 2017.

[16] K. Vu, R. Zheng, and J. Gao, "Efficient algorithms for K-anonymous location privacy in participatory sensing," in *Proceedings of 31st Annual IEEE International Conference on Computer Communications (INFOCOM, 2012)*, pp. 2399–2407, IEEE, Orlando, FL, USA, March 2012.

[17] T. Ma, Y. Zhang, J. Cao et al., "KDVEM: a k-degree anonymity with vertex and edge modification algorithm," *Computing*, vol. 97, no. 12, pp. 1165–1184, 2015.

[18] R. Zhang, Y. Zhang, and C. Zhang, "Secure top-k query processing via untrusted location-based service providers," in *Proceedings of 31st Annual IEEE International Conference on Computer Communications (INFOCOM, 2012)*, pp. 1170–1178, IEEE, Orlando, FL, USA, March 2012.

[19] R. Zhang, J. Sun, Y. Zhang, and C. Zhang, "Secure spatial top-k query processing via untrusted location-based service providers," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 1, pp. 111–124, 2015.

[20] C. Y. Chow, M. F. Mokbel, and W. G. Aref, "Casper: query processing for location services without compromising privacy," *ACM Transactions on Database Systems (TODS)*, vol. 34, no. 4, pp. 1–48, 2009.

[21] C. Yin, J. Xi, and R. Sun, "Location privacy protection based on improved-value method in augmented reality on mobile devices," *Mobile Information Systems*, vol. 2017, Article ID 7251395, 7 pages, 2017.

[22] C. Yin and J. Xi, "Maximum entropy model for mobile text classification in cloud computing using improved information gain algorithm," *Multimedia Tools and Applications*, vol. 76, no. 16, pp. 16875–16891, 2017.

[23] H. Rong, T. Ma, M. Tang, and J. Cao, "A novel subgraph K+-isomorphism method in social network based on graph similarity detection," *Soft Computing*, vol. 22, no. 8, pp. 2583–2601, 2017.

[24] B. Gu, X. Sun, and V. S. Sheng, "Structural minimax probability machine," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 7, pp. 1646–1656, 2017.

[25] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 26, no. 7, pp. 1403–1416, 2015.

[26] R. Lu, X. Lin, X. Liang, and X. Shen, "A dynamic privacy-preserving key management scheme for location-based

services in VANETs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 1, pp. 127–139, 2012.

[27] A. R. A. Bouguettaya and M. Y. Eltoweissy, "Privacy on the web: facts, challenges, and solutions," *IEEE Security & Privacy*, vol. 99, no. 6, pp. 40–49, 2003.

[28] A. Lee, "Guidelines for smart grid cyber security," NIST Interagency/Internal Report (NISTIR)-7628, National Institute of Standards and Technology, Gaithersburg, MD, USA, 2010.

[29] M. F. Mokbel, C. Y. Chow, and W. G. Aref, "The new Casper: a privacy-aware location-based database server," in *Proceedings of IEEE 23rd International Conference on Data Engineering (ICDE 2007)*, pp. 1499-1500, IEEE, Istanbul, Turkey, April 2007.

[30] X. Li, E. Wang, W. Yang, and J. Ma, "DALP: a demand-aware location privacy protection scheme in continuous location-based services," *Concurrency and Computation: Practice and Experience*, vol. 28, no. 4, pp. 1219–1236, 2016.

[31] B. Gedik and L. Liu, "Protecting location privacy with personalized k-anonymity: architecture and algorithms," *IEEE Transactions on Mobile Computing*, vol. 7, no. 1, pp. 1–18, 2008.

[32] P. Kalnis, G. Ghinita, K. Mouratidis, and P. Dimitris, "Preventing location-based identity inference in anonymous spatial queries," *IEEE Transactions on Knowledge & Data Engineering*, vol. 19, no. 12, pp. 1719–1733, 2007.

[33] K. W. Tan, Y. Lin, and K. Mouratidis, "Spatial cloaking revisited: distinguishing information leakage from anonymity," in *Proceedings of 11th International Symposium on Spatial and Temporal Databases*, pp. 117–134, Aalborg, Denmark, July 2009.

[34] T. C. Li and W. T. Zhu, "Protecting user anonymity in location-based services with fragmented cloaking region," in *Proceedings of 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE)*, pp. 227–231, IEEE, Zhangjiajie, China, May 2012.

[35] D. Steiert, D. Lin, Q. Conduff, and W. Jiang, "Poster: a location-privacy approach for continuous queries," in *Proceedings of the 22nd ACM on Symposium on Access Control Models and Technologies*, pp. 115–117, ACM, Indianapolis, IN, USA, June 2017.

[36] E. H. Yilmaz, E. Ferhatosmanoglu, and R. C. Aksoy, "Privacy-preserving aggregate queries for optimal location selection," *IEEE Transactions on Dependable and Secure Computing*, vol. 2017, no. 99, p. 1, 2017.

[37] I. Memon, "Authentication user's privacy: an integrating location privacy protection algorithm for secure moving objects in location based services," *Wireless Personal Communications*, vol. 82, no. 3, pp. 1585–1600, 2015.

[38] Y. Wang, D. Xu, and F. Li, "Providing location-aware location privacy protection for mobile location-based services," *Tsinghua Science and Technology*, vol. 21, no. 3, pp. 243–259, 2016.

[39] T. Brinkhoff, "A framework for generating network-based moving objects," *GeoInformatica*, vol. 6, no. 2, pp. 153–180, 2002.

[40] M. Hadjieleftheriou, Y. Manolopoulos, Y. Theodoridis, and V. J. Tsotras, "R-trees–a dynamic index structure for spatial searching," in *Encyclopedia of GIS*, pp. 993–1002, Springer, Boston, MA, USA, 2008.