Research Article

An Approach Based on Customized Robust Cloaked Region for Geographic Location Information Privacy Protection

Zhidong Shen^{1,2}, Siyuan Lu,² Huijuan Huang,² Meng Yuan,² Guoming Tang,³ Weiying Chen,¹ Taige Zhang,¹ and Ting Zhong¹

¹Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Wuhan 430079, China

²International School of Software, Wuhan University, Wuhan 430079, China

³Science & Technology on Information Systems Engineering Lab, National University of Defense Technology, Changsha, China

Correspondence should be addressed to Zhidong Shen; shenzd@whu.edu.cn

Received 22 October 2019; Revised 16 December 2019; Accepted 17 January 2020; Published 12 February 2020

Academic Editor: Luca Bedogni

Copyright © 2020 Zhidong Shen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Location-based services (LBS) have gained huge popularity because of the easy availability of modern mobile devices and the fast development of geographical information science (GIS). However, the lack of protection for private user positions might give rise to privacy concerns. This kind of problem is especially serious in mobile application environment because many mobile applications tend to use LBS. In this paper, we propose a new privacy preserving approach using customized robust cloaked region (RCR), depending on a peer-to-peer structure and the premise that users do not trust each other when sharing their geographical locations. Two algorithms are used to generate the RCR with high user density. The area of the RCR is controlled by the user's demanded degree of protection. To enhance the resistance to regional background knowledge attack, we incorporate a location semantic value into each unit of the user map. According to extensive simulations, our method can effectively obfuscate a user's geographical location into a highly indistinguishable region because of the disturbance of nearby users and different equally possible locations.

1. Introduction

With the development of the geographical information science (GIS) technique and the popularity of mobile devices, location-based services (LBS) have been widely employed. Extensive daily examples can be found in navigation services. Although these services are popular, LBS have caused many privacy concerns. Also, this question has become a popular issue and one that has drawn the attention of researchers. Among all solutions to geographical location preserving, spatial obfuscation has been proved the most effective one. In general, it deliberately generalizes the user's precise position into a region, so that adversaries can only retrieve a coarse location. Gruteser and Grunwald [1] firstly defined k -anonymity in the context of spatial obfuscation. Several methods such as the Casper framework presented by Mokbel et al. [2] and the nearest neighbor cloak and the

Hilbert cloak proposed by Kalnis et al. [3] are TTP-based. However, the anonymizer is highly likely to become the single attack target, and a severe privacy leakage will be incurred if the adversary gains access to it.

In addition, k -anonymity can also be implemented by a decentralized peer-to-peer (P2P) network [4]. Another P2P approach called MobiHide, presented by Ghinita et al. [5], uses Hilbert space-filling curves to hide the query issuer among a group of k users.

Although location anonymity suggests a practical way to obfuscate the user's precise location, it is vulnerable to many types of attack without careful consideration, and one of the most significant attacks is regional background attack (RBA). Specifically, an adversary could utilize the background knowledge to increase the precision of the user's position by excluding nonreachable area from the obfuscated area [6].

To address the abovementioned drawbacks when applying spatial obfuscation, this paper introduces a new approach using customized robust cloaked region (RCR). The proposed approach is based on a P2P structure and designed to address the privacy concern raised by sharing accurate positions. The main idea is that users no longer share their exact locations but an RCR based on nearby users' RCRs. Different from the cloaked region formed by applying k -anonymity, the number of users included in the RCR in our context is unknown, rather than a specific number. Instead, the size of the cloaked region is under control in accordance with the privacy preference of the user, which avoids an overlarge cloaked region or time delay caused by low user population [1].

The major contributions of this paper can be summarized as follows:

- (i) We propose the idea of replacing the precise position with a cloaked region when sharing location information in the peer-to-peer structure, which greatly reduces the risk of privacy breach.
- (ii) We propose two effective algorithms to calculate the cloaked region with a high user density.
- (iii) We enhance the robustness of the cloaked region by integrating location semantics and considering time parameter to reflect the fluctuation of population. This approach successfully defends against RBA, since a more reasonable cloaked region is produced and hard to be shrunk based on limited background knowledge.
- (iv) Extensive experiments are performed with real-world data, and the results have validated the correctness and superiority of the proposed approach and algorithm.

2. Related Work

Historically, researchers have exerted great efforts toward the protection of location privacy. According to the protection goals, existing mechanisms can be classified into three main categories: protection for user identities, protection for position information, and protection for querying content. User identity protection usually modifies or hides real identities of mobile users while position protection adds noise to users' precise locations. Query content protection enables a data holder to share person-specific records in such a way that the released information remains practically useful but the identity of the individuals who are the subjects of the data cannot be determined [7].

2.1. Protection of User Identities. Anonymity and pseudonymity are two common ways to protect user identities. Anonymity makes an individual indistinguishable from all other individuals in a set, while in pseudonymity, an individual maintains a persistent identity (a pseudonym) that cannot be linked to their actual identities [8]. Among all the anonymity approaches, k -anonymity is the most widely used, which obfuscates user identities within a group of

users. Another common approach to protect user identities is mix-zones [9] where the user identity is mixed with all other users in the zone by changing the pseudonym. Freudiger [10] incorporated the mix-zone model with techniques of encryption to produce higher levels of privacy protection. However, such technique is limited to those LBS that do not require the user's identity. Also, with advancements in techniques of data mining, pseudonymity and anonymity are no longer safe enough since identity can often be inferred from the location [6].

2.2. Protection for Position Information. Existing works designed to preserve the privacy of position information can be mainly divided into two categories: position dummies and spatial obfuscation. The main idea of position dummies [11] is that when a user sends a service request to the Location Service Provider (LSP), in addition to its real location, it also sends multiple false positions randomly created. However, by monitoring long-term movement patterns of the user, the server may distinguish the genuine location from dummies. In order to reduce the computational cost, Hong and Landay [12] proposed a landmark approach using a landmark to replace the user's real position. On the other hand, spatial obfuscation approaches [1, 13–15] deliberately reduce the precision of position information sent to the LSP by replacing the user's position information with a larger region.

2.3. Protection for Query Content. K -anonymity is a widespread general privacy concept firstly proposed by Samarati and Seweeney [16] to protect the query content, which guarantees that each record is indistinguishable from at least $k-1$ other records with respect to certain identifying attributes [17]. The model of k -anonymity has been enhanced by various approaches to increase the level of privacy protection. The most prominent enhancements are strong k -anonymity [18], l -diversity [17], t -closeness [19], p -sensitivity [20], and (α, k) -anonymity [21]. In addition, several works use differential privacy to protect the location privacy in information datasets, where the presence of no single user could significantly change the outcome of the aggregated location information [22].

Nonexposure location anonymity, presented by Pan et al. [23], is the first study that explores the problem of location cloaking without exposing the accurate user locations. It is designed for k -anonymity, and cloaking is performed based on the proximity information among mobile users, instead of directly on their coordinates. The PROBE approach proposed by ML Damiani et al. [24] obfuscates the user location by taking into account the geographical context and the user's privacy preferences. The substantial difference with our system is that we not only perform spatial obfuscation based on location semantics but we also provide anonymity by making use of nearby users' locations. Moreover, we noticed that the population of a location varies greatly at different times; therefore, we incorporate time parameter with location semantics to ensure correctness.

3. Cloaked Region Generation Algorithm

In this section, we will firstly present several fundamental definitions and then demonstrate the workflow of the cloaked region generation. Finally, two cloaked region generation algorithms will be introduced.

Definition 1. A cloaked region is a rectangle represented by R_x and is defined as the following five-element tuple:

$$R_x = (x_l, y_l, x_r, y_r, id). \quad (1)$$

In this definition, (x_l, y_l) and (x_r, y_r) stand for the coordinate of the bottom left and upper right corner of the cloaked region, respectively. The element of id is used to uniquely identify a cloaked region, so two ids cannot be the same.

Definition 2. The size of the cloaked region $S(R_x)$ is defined as

$$S(R_x) = |x_r - x_l| * |y_r - y_l|. \quad (2)$$

Definition 3. A_{max} denotes the maximum allowed size, and A_{min} denotes the minimum allowed size for the cloaked region. Users are allowed to set up A_{max} and A_{min} to restrict the size of the cloaked region generated so as to satisfy varied degree of individual privacy protection requirement. In this algorithm, a gridding map named *userMap* is created upon actual geography [25–27]. The actual location of every user is represented by a cell in this map, and a single cell can accommodate more than one user.

The workflow of the cloaked region generation is depicted as follows:

- (i) The targeted user joins into a peer-to-peer distributed system and then searches for his neighbor through point-to-point communication protocol on his terminal.
- (ii) The neighbors share their cloaked regions to the targeted user. To ensure efficiency, when a neighboring user is asked to share his/her location, if he/she has generated an RCR for his history queries and still remains in it, the previously generated RCR will be delivered; otherwise, he/she will generate one randomly.
- (iii) A cloaked region is calculated according to the generation algorithm.
- (iv) Finally, a group member is randomly selected to send the intended query together with the generated cloaked region to the LSP.

To further clarify the abovementioned process, we present an example here. As shown in Figure 1, the targeted user u_1 has six neighbor $u_2 - u_7$ who share their own cloaked regions to u_1 through peer-to-peer communication protocol. As shown in Figure 2, the cloaked regions are represented by varied size rectangles generated according to predetermined A_{max} and A_{min} . As illustrated in Figure 3, a rectangular *userMap* will be generated which is centered on the precise

FIGURE 1: User distribution in actual geography.

FIGURE 2: RCR generated for neighbor.

FIGURE 3: User distribution in the gridding map.

location of the targeted user and covers all the cloaked regions of his neighbor. Note that the size of the *userMap* can be adjusted by users so that it can provide a particular individual request of privacy protection.

The cloaked region cell value assignment principle is defined as follows. It is based on the fact that the user will appear wherever in his cloaked region with equal probability. For example, the region M_2 in Figure 3 is a 4×5 rectangle if the unit length is defined as the length of the side of the individual cell. Then, the probability of user U_2 appearing at each cell is equal to $1/S(M_2)$, which is 5%. For simplicity, we assign the value of each cell as 100 times as its probability. If one cell is covered by more than one cloaked region, then we calculate the sum of the cell value in each cloaked region as its ultimate value. Figure 4 shows the result of the assignment according to this principle.

A high-quality cloaked region should contain as many users as possible, and meanwhile occupy a smaller area for the sake of the usability to the LBS. Therefore, it is crucial to provide a tradeoff between the sum of *Cell-Values* in a region and its size. To clarify this issue, we refer as *userDensity* to the measurement for the quality of a cloaked region.

Definition 4. The *userDensity* of a region is defined as the ratio of the amount of users in this region to the size of the region. According to the cloaked zone cell value assignment principle, the number of users in a certain region equals to the sum of the cell values in this region. Thus, the user density of a specified region M can be interpreted mathematically as

$$\text{userDensity}(M) = \frac{\sum_{c \in M} \text{CellValue}(c)}{S(M)}. \quad (3)$$

To obtain the cloaked region, we are supposed to look for the area which possesses the largest *userDensity*. Based on this purpose, we proposed two algorithms to generate the cloaked region.

3.1. Exhaustive Method. The exhaustive method, as given in Algorithm 1, aims to traverse all the rectangles containing the cell where the targeted user is located until the one with maximum user density is found and then return the final rectangle as the cloaked region to the user. This algorithm can guarantee that the generated cloaked region is of smaller size with more users. However, the drawback of this method is the high computational complexity and is thus time-consuming in practice.

3.2. Heuristic Method. Given the high computational complexity of the exhaustive method, we develop a heuristic method, Algorithm 2, to keep the balance between quality of privacy protection and cost of computation.

In the *userMap*, the higher the value a cell has, the more likely a user will appear in this cell. When a cell with a high value is located far from the targeted user, the

			5	5	5	5			10	10	10	10	10				
			5	5	5	5			10	10	10	10	10				
			5	5	5	5											
			5	5	5	5			5	5	5	5	5				
			5	5	9	9	4	9	9	5	5	5					
					4	4	4	U	9	5	5	5					
		4	4	4	8	8	4	9	9	5	5	5					
		4	4	4	8	8	4	4	4								
		4	4	4	8	8	4	4	9	5	5	5					
		4	4	4	4	4			5	5	5	5					
		4	4	4	4	4			5	5	5	5					
									5	5	5	5					
									5	5	5	5					

FIGURE 4: Gridding map after assignment.

cloaked region generated for the targeted user will shift to the high value cell which may result in a deviated cloaked region; also, there is a declined quality of LBS service. Thus, in order to guarantee an acceptable quality of service requiring less algorithm complexity, the distance between the targeted user and his neighbor should be taken into consideration when performing the cell value assignment.

Definition 5. The *dCellValue* for a cell c denotes the distance between c and the targeted user. It is defined as the ratio of the sum of the cell values in the rectangle M determined by cell c and C_{center} to the number of the cells in M . The *dCellValue* can be depicted in the following formula:

$$\text{dCellValue}(c) = \frac{\sum_{c \in M} \text{CellValue}(c)}{|M|}. \quad (4)$$

4. Robust Cloaked Region Generation Algorithm

The aforementioned cloaked region generation algorithm is more based on theory rather than daily life scenario. Thus, it fails to develop immunity to adversaries with background knowledge. In this section, we will introduce the customized coefficient *Diversity* which will greatly enhance the defensiveness of the region generation algorithm.

4.1. Location Semantics. In our daily lives, each location possesses certain semantics in which a group of representative users distribute, and this allows adversaries with background knowledge, in particular, map knowledge, to speculate the precise user spatial information by excluding nonreachable areas from the cloaked area. The cloaked

Definition: cl = bottom left cell of the cloaked region,
 cr = upper right cell of the cloaked region,
 $Ccenter$ = centric cell of the $userMap$,
 $M = (cl, cr)$ (a rectangle determined by cl and cr),
 $maxUserDensity$ = the maximum value of $userDensity$

Input: Nearby users' cloaked regions and customized size constraints $Amin$ and $Amax$

Output: Cloaked region for the target user

```

(1) Initialize the  $userMap$  by calculating  $CellValue$ ;
(2) Initialize  $maxUserDensity$  as 0;
(3) for cell  $i$  in the rectangle area  $(i, Ccenter)$  do
(4)   for cell  $j$  in the rectangle area  $(Ccenter, j)$  do
(5)     assign  $M = (i, j)$ ;
(6)     if  $M$  contains  $Ccenter$ , and  $S(M) > Amin$ ,  $S(M) < Amax$  then
(7)       calculate  $userDensity$  for  $M$ ;
(8)       if  $userDensity(M) > maxUserDensity$  then
(9)         assign  $maxUserDensity = userDensity(M)$ ;
(10)        assign  $cl = i$  and  $cr = j$ ;
(11)      end
(12)    end
(13)  end
(14) end
(15) return the cloaked region determined by  $cr$  and  $cl$ 

```

ALGORITHM 1: Generate cloaked region using exhaustive method.

Input: Nearby users' cloaked regions and customized size constraints $Amin$ and $Amax$

Output: The cloaked region for the target user

```

(1) Initialize the  $userMap$  with the  $dCellValue$ ;
(2) Initialize cloaked region  $M$  to contain  $Ccenter$  only;
(3) while search for cell  $c$  with maximum  $dCellValue$  do
(4)   extend  $M$  to cell  $c$ ;
(5)   if  $S(M) > Amax$  then
(6)     withdraw  $M$  to its previous state and set the  $dCellValue$  of  $c$  to 0;
(7)   else
(8)     for each cell in  $M$  do
(9)       set  $dCellValue$  to 0;
(10)    end
(11)  end
(12) end
(13) return the cloaked region  $M$ 

```

ALGORITHM 2: Generate cloaked region using the heuristic method.

region generation algorithm can be more defensive if location semantics is taken into consideration, since the area generated will contain fewer places where users are less likely to appear.

Definition 6. Location semantics K is defined as the degree a user is likely to appear in a certain location. K is a number between 0 and 1, where 0 stands for “zero likely” with *possibility of zero* and 1 stands for “very likely” with *possibility of one*. K is a variable that indicates the spatial distribution of population and the user's preference to a location. It varies for different users, for example, if the user is a doctor, the semantics of hospital for this user is much

higher than other people who do not work in the hospital (the hospital is sensitive for a patient, but not for a doctor).

Definition 7. The probability of a user showing up in cell g in a cloaked region M is defined as

$$p(g) = \frac{K(g)}{\sum_{g' \in M} K(g')}, \quad (5)$$

where $K(g)$ stands for the location semantics of cell g .

Definition 8. Assume that cell g is covered by several cloaked regions $m_1, m_2, m_3, \dots, m_n$ ($n = 1, 2, \dots$), with the

probability of each region $p_1(g), p_2(g), p_3(g), \dots, p_n(g)$. Let L be the sum of the probability of g in all the cloaked regions it appears, which is denoted as

$$L(g) = p_1(g) + p_2(g) + \dots + p_n(g). \quad (6)$$

Definition 9. Given a cloaked region M , its user density can be defined as

$$\text{userDensity}(M) = \frac{\sum_{g \in M} L(g)}{|M|}, \quad (7)$$

where g denotes the cell in M and $|M|$ stands for the size of M .

4.2. Time Parameter. In practice, not only does the location semantics play an important role in privacy protection, but the time parameter can also enhance the defensiveness, since the crowdedness of a place varies in a day; for instance, people are more likely to gather in the office during the day while in their residence during the night.

Definition 10. The user's preference to a certain position on a time point or during a time period is denoted as T .

4.3. Size Customization. As there is a trade-off between the benefits gained from LBS and possibility that private information might be revealed (at least partially), a user can choose the size of the cloaked region to be generated by assigning the values of $A_{\max}$ and $A_{\min}$, as in Algorithm 3. Bigger $A_{\max}$ can result in a larger obfuscated area, and vice versa.

Definition 11. A customized coefficient named *Diversity* can be defined by a four-element tuple consisting of location semantics K , time parameter T , $A_{\max}$, and $A_{\min}$:

$$\text{Diversity} = \{K, T, A_{\max}, A_{\min}\}. \quad (8)$$

5. Experiments

The experiment is divided into two parts. The first part aims to compare the performance of two cloaked region generation algorithms, namely, the exhaustive method and the heuristic method; the second part is to demonstrate the correctness and superiority of our RCR generation algorithms.

Our method is applied during a specific moment in time for simplification. Therefore, we implement our algorithms under static scenarios without taking mobility into account. However, our approach could be extended to moving users with further improvements and necessary modifications.

The experiments were conducted on Windows 7 Operating System with Intel Core (TM) i5-4200U 1.60 GHz CPU and 8 GB RAM. We obtained experimental data from the Brinkhoff generator, which is an object-oriented data generation system developed by researcher Thomas

Brinkhoff in 2000. Brinkhoff's generator takes a gridding map as input, adopts a discrete timing model, and produces new objects at each time stamp. The generator runs on JDK 1.8.

5.1. Comparison Experiments between Exhaustive and Heuristic Methods. In the comparison experiments between the exhaustive method and heuristic method, we created a $2500 * 2500$ gridding map upon the actual map of Oldenburg in Germany, as shown in Figure 5 and adopted it as the input map. The *userMap* was divided into $20 * 20$ rectangles, and the number of the simulated users varied from 1000 to 5000. Since the number of neighbors around the targeted user and the size restriction (set up by the targeted user) both have big impacts on the cloaked region generated, we examined these two variables, respectively, in the experiment. The parameter setting is given in Table 1.

For performance evaluation, we measure the average value of the following metrics: (1) user density, (2) size of the RCR generated, and (3) running time of the algorithm.

5.1.1. User Density Comparison. As shown in Figures 6 and 7, the user density of the cloaked region generated by the exhaustive method is always higher than that generated by the heuristic method, as the working mechanism of the exhaustive method is to find the cloaked region with maximum user density under the specified size restriction. In terms of user density of the cloaked region generated by the two algorithms, the exhaustive method has a better performance than the heuristic method.

5.1.2. Size Comparison. As shown in Figures 8 and 9, the size of the cloaked region generated by the heuristic method is always larger than that generated by the exhaustive method. Since the heuristic method adopts greedy algorithm, cloaked region continues to extend to the cell with maximum *dCellValue* until it reaches its size upper bound. As a larger cloaked region leads to the declined quality of LBS service, the exhaustive method has a better performance.

5.1.3. Running Time Comparison. As shown in Figures 10 and 11, the running time of the exhaustive method is much longer than that of the heuristic method, and the difference grows larger with the increase in region size. The longer running time of the exhaustive method lies in the fact that it continues searching the rectangles containing the targeted user until the one of the maximum user density is found. Also, as $A_{\min}$ and $A_{\max}$ get larger, the exhaustive method will take an even longer time to find the desired rectangle. That is why the running time of the exhaustive method will soar as the allowed size of the generated RCR grows larger. As shown in Figure 11, the running time of the heuristic method decreases as $A_{\min}$ and $A_{\max}$ grow larger. This is because when the number of users remains still, the user density descends as the size of the cloaked region grows

Definition: cl = the bottom left cell of the RCR,
 cr = the upper right cell of the RCR,
 $M = (cl, cr)$ (a rectangle determined by cl and cr),
 C_{center} = the cell in which the targeted user is located,
 $maxUserDensity$ = the maximum value of $userDensity$

Input: Nearby users' cloaked regions, customized size constraints A_{min} and A_{max} , location semantic K under time T

Output: the RCR for the target user

- (1) Update $userMap$ with location semantics K ;
- (2) Assign the calculated $P(g)$ for each cell in the $userMap$;
- (3) **for** each rectangle $M = (i, j)$ in the $userMap$ **do**
- (4) **if** M contains C_{center} , and $S(M) < A_{max}$, $S(M) > A_{min}$ **then**
- (5) **calculate** $userDensity(M)$;
- (6) **if** $userDensity(M) > maxUserDensity$ **then**
- (7) $maxUserDensity = userDensity(M)$; $cl = i$; $cr = j$;
- (8) **end**
- (9) **end**
- (10) **end**
- (11) **return** RCR determined by cr and cl

ALGORITHM 3: Integrated robust cloaked region (RCR) generation algorithm.

FIGURE 5: Map of Oldenburg city in Germany.

TABLE 1: Experiment parameters.

The number of neighbors	Size restriction of cloaked region
10, 20, 30, 40, 50	[25, 49]
	[9, 16] [16, 25] [25, 36] [36, 49]
30	[49, 64]
	[64, 81] [81, 100]

larger. As a result, the RCR extension time will decline at the same time. When looking into the running time for generation of RCR, we find that the heuristic method has a better performance.

5.1.4. RCR Generation Algorithm Experiment. In the RCR generation algorithm experiment, we created a $25 * 25$

gridding map upon on district map of the city of Wuhan, China (as shown in Figure 12), and adopted it as the input map. The number of simulated neighbor is increased from 10 to 50, and the size of the RCR is restrained to [25, 81]. After updating the gridding map with location semantics shown in Table 2, we can get the gridding map as shown in Figure 13. For performance evaluation, we measure the average values of the following metrics: (1) user density of the RCR, (2) size of the RCR, and (3) the defensiveness of RCR generation algorithm to adversaries.

As shown in Figure 14, as the number of neighbors increases from 10 to 50, the user density of RCR does go up as it is supposed to be.

As indicated in Figure 15, the average size of the RCR generated grows larger as the number of neighbors rises. That is because as the number of neighbor increases, the

FIGURE 6: Relationship between user density and the number of neighbors.

FIGURE 7: Relationship between user density and size restriction.

FIGURE 8: Relationship between size and number of neighbors.

RCR tends to contain the area which is of high user density but locates far from the targeted user. Consequently, the size of the RCR grows larger.

Table 3 demonstrates the ability of our RCR generation algorithm to prevent the user's spatial privacy from being revealed. An RCR of high quality ought to contain areas of similar location semantics K so that the adversary with

background knowledge cannot use map matching to exclude irrelevant areas and thus shrink the cloaked region under intended size. As shown in Table 3, our RCR is capable of generating a high-quality cloaked area as defense from adversaries with background knowledge. For example, when t is weekday daytime, the probability of RCR-containing lake is only 1.5%, whereas the probability rises to 18.3% at night. This

FIGURE 9: Relationship between size and size restriction.

FIGURE 10: Relationship between running time and number of neighbors.

FIGURE 11: Relationship between running time and size restriction.

FIGURE 12: Map of Xudong ave., Wuhan, Hubei 430077, China.

is because at daytime, people are less likely to be around a lake as they may need to work or study. This is also a reason to why K for lake at daytime is small. As a consequence, the generated obfuscated area tends to cover less lake area. However, when it

is night time, the K for mall and office building becomes the same as that for lake, since people are more likely to stay at home on a weekday evening. As K for the lake, park, office building, and mall are the same, the probability for the

TABLE 2: Location semantic K under different time periods T .

Location	Location semantics K			
	$T=\text{weekday daytime}$	$T=\text{weekday night}$	$T=\text{weekend daytime}$	$T=\text{weekend night}$
Road	0.7	0.3	0.9	0.4
Lake	0.1	0.1	0.1	0.1
Park	0.5	0.1	0.7	0.1
Downtown	0.7	0.1	0.9	0.1
Residential area	0.3	0.9	0.3	0.9

FIGURE 13: Location semantic gridding map.

FIGURE 14: Relationship between user density and number of neighbors.

FIGURE 15: Relationship between size and number of neighbors.

TABLE 3: User showing up probability in the cloaked region under different time periods T .

Location	User showing up probability				Size percentage on the map (%)
	T weekday daytime (%)	T weekday night (%)	T weekend daytime (%)	T weekend night (%)	
Road	58.7	25.9	64.8	32.7	42.24
Lake	1.5	18.3	1.8	19.1	9.12
Park	24.1	18.9	26.7	17.6	6.72
Downtown	29.7	19.8	55.9	16.2	20.8
Residential area	22.1	42.7	29.4	49.3	14.4
Office building	32.5	19.6	23.4	22.8	6.72

generated RCR-containing lake goes up, so that the adversary cannot exclude the lake region from RCR easily.

6. Conclusions

We proposed a new geographical location preserving mechanism. The mechanism was based on peer-to-peer structure and used RCR to eliminate the trust concern within a user group. The RCR is generated under customized size constraints with an attempt to have as high user density as possible. Two methods developed to calculate the RCR are the exhaustive method and the heuristic method.

There are still many worthy focal points for further research based on the work in this paper: (1) the mechanism we propose mainly focuses on the snapshot query scenario, and if the user is making continuous LBS queries, the algorithm will not be useful. (2) Location semantics is just mentioned as a concept in this article, and to make it practical in real-time application, we need to work out a semantic labeling framework. (3) The protection model proposed can be further strengthened to resist many other possible attacks; for example, it is still likely to reveal the user's identity with the user's profile attributes, even if protected by an anonymity user group; therefore, more sophisticated strategies should be applied when conducting anonymity.

Data Availability

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

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the National Key R&D Program of China (Grant no.2018YFC1604000/2018YFC1604002) and the Natural Science Foundation of Hubei Province (Grant no.2017CFB663).

References

- [1] 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*, ACM, San Francisco, CA, USA, May 2003.
- [2] M. F. Mokbel, C.-Y. Chow, and W. G. Aref, "The new Casper: query processing for location services without compromising privacy," in *Proceedings of the 32nd International Conference on Very Large Data Bases*, VLDB Endowment, Seoul, Korea, September 2006.
- [3] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias, "Preventing location-based identity inference in anonymous spatial queries," *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 12, pp. 1719–1733, 2007.
- [4] C.-Y. Chow, M. F. Mokbel, and X. Liu, "Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments," *GeoInformatica*, vol. 15, no. 2, pp. 351–380, 2011.
- [5] G. Ghinita, P. Kalnis, and S. Skiadopoulos, "MOBIHIDE: a mobile peer-to-peer system for anonymous location-based queries," in *Proceedings of the International Symposium on Spatial and Temporal Databases*, Springer, Boston, MA, USA, July 2007.
- [6] J. Krumm, "Inference attacks on location tracks," in *Proceedings of the International Conference on Pervasive Computing*, Springer, San Francisco, CA, USA, 2007.
- [7] W. Wei, F. Xu, and Q. Li, "Mobishare: flexible privacy-preserving location sharing in mobile online social networks," in *Proceedings of the 2012 Proceedings IEEE INFOCOM*, IEEE, Orlando, FL, USA, March 2012.
- [8] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in *Proceedings of the International Conference on Pervasive Computing*, Springer, Munich, Germany, May 2005.
- [9] Z. Liu, "New privacy-preserving location sharing system for mobile online social networks," in *Proceedings of the 2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, IEEE, Compiegne, France, October 2013.
- [10] J. Freudiger, M. Raya, M. Flegyhzi, P. Papadimitratos, and J.-P. Hubaux, "Mix-zones for location privacy in vehicular networks," in *Proceedings of the ACM Workshop on Wireless Networking for Intelligent Transportation Systems (WiN-ITS)*, Vancouver, Canada, August 2007.
- [11] H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in *Proceedings of the ICPS'05 International Conference on Pervasive Services*, IEEE, Santorini, Greece, July 2005.
- [12] J. I. Hong and J. A. Landay, "An architecture for privacy-sensitive ubiquitous computing," in *Proceedings of the 2nd*

- International Conference on Mobile Systems, Applications, and Services*, ACM, Boston, MA, USA, 2004.
- [13] B. Hoh, "Preserving privacy in gps traces via uncertainty-aware path cloaking," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, ACM, Alexandria, VA, USA, October 2007.
 - [14] Y. Cai and Ge Xu, "Cloaking with footprints to provide location privacy protection in location-based services," U.S. Patent No. 9,239,935, 2016.
 - [15] C. A. Ardagna, M. Cremonini, E. Damiani, S. De Capitani di Vimercati, and P. Samarati, "Location privacy protection through obfuscation-based techniques," in *Proceedings of the IFIP Annual Conference on Data and Applications Security and Privacy*, Springer, Redondo Beach, CA, USA, July 2007.
 - [16] P. Samarati and L. Sweeney, "Generalizing data to provide anonymity when disclosing information," *PODS*, vol. 98, 1998.
 - [17] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: privacy beyond k -anonymity," *ACM Transactions on Knowledge Discovery from Data*, vol. 1, no. 1, p. 3, 2007.
 - [18] C. Zhang and Y. Huang, "Cloaking locations for anonymous location based services: a hybrid approach," *GeoInformatica*, vol. 13, no. 2, pp. 159–182, 2009.
 - [19] N. Li, T. Li, and V. Suresh, "t-closeness: privacy beyond k -anonymity and l -diversity," in *Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering*, IEEE, Istanbul, Turkey, April 2007.
 - [20] A. Solanas, F. Seb, and J. Domingo-Ferrer, "Micro-aggregation-based heuristics for p -sensitive k -anonymity: one step beyond," in *Proceedings of the 2008 International Workshop on Privacy and Anonymity in Information Society*, ACM, Nantes, France, March 2008.
 - [21] R. C.-W. Wong, " (α, k) -anonymity: an enhanced k -anonymity model for privacy preserving data publishing," in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY, USA, August 2006.
 - [22] Z. Montazeri, A. Houmansadr, and H. Pishro-Nik, "Defining perfect location privacy using anonymization," in *Proceedings of the 2016 Annual Conference on Information Science and Systems (CISS)*, IEEE, Princeton, NJ, USA, March 2016.
 - [23] X. Pan, X. Meng, and J. Xu, "Distortion-based anonymity for continuous queries in location-based mobile services," in *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, Seattle, WA, USA, November 2009.
 - [24] M. L. Damiani, E. Bertino, and C. Silvestri, "Protecting location privacy against spatial inferences: the PROBE approach," in *Proceedings of the 2nd SIGSPATIAL ACM GIS 2009 International Workshop on Security and Privacy in GIS and LBS*, ACM, Seattle, WA, USA, January 2009.
 - [25] T. D'Roza, G. Bilchev, and B. George, "An overview of location-based services," *BT Technology Journal*, vol. 21, no. 1, pp. 20–27, 2003.
 - [26] J. Camenisch, D. A. Ortiz-Yepes, and F.-S. Preiss, "Strengthening authentication with privacy-preserving location verification of mobile phones," in *Proceedings of the 14th ACM Workshop on Privacy in the Electronic Society*, ACM, Denver, CO, USA, October 2015.
 - [27] J. W. Hager, *Datums, Ellipsoids, Grids, and Grid Reference Systems*. No. DMA-TM-8358.1, Defense Mapping Agency Hydrographic/Topographic Center, Washington DC, USA, 1992.