TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 02/11 pp260–269
Volume 21, Number 3, June 2016

An Efficient Location Privacy Protection Scheme Based on the Chinese Remainder Theorem

Jingjing Wang*, Yiliang Han, and Xiaoyuan Yang

Abstract: Traditional k-anonymity schemes cannot protect a user's privacy perfectly in big data and mobile network environments. In fact, existing k-anonymity schemes only protect location in datasets with small granularity. But in larger granularity datasets, a user's geographical region-location is always exposed in realizations of k-anonymity because of interaction with neighboring nodes. And if a user could not find enough adjacent access points, most existing schemes would be invalid. How to protect location information has become an important issue. But it has not attracted much attention. To solve this problem, two location-privacy protection models are proposed. Then a new generalized k-anonymity Location Privacy Protection Scheme based on the Chinese Remainder Theorem (LPSS-CRT) in Location-Based Services (LBSs) is proposed. We prove that it can guarantee that users can access LBSs without leaking their region-location information, which means the scheme can achieve perfect anonymity. Analysis shows that LPPS-CRT is more secure in protecting location privacy, including region information, and is more efficient, than similar schemes. It is suitable for dynamic environments for different users' privacy protection requests.

Key words: location privacy protection; generalized k-anonymity; location-based services; the Chinese remainder theorem

1 Introduction

More and more people join in the social networks to share information, and generate a lot of data thereby. But users do not want to open their data to the public or the service provider. Especially in modern wireless networks or mobile networks equipped with GPS facilities, Location-Based Services (LBSs) provide personalized services to mobile users based on their geographical locations.

However, this can be a double-edged sword. When users seek more benefits from LBSs, they have to accept abrogation of privacy as a cost. During LBS queries, user location information[1] is in danger of leaking out.

Most people do not want to publish their location, even their geographical regions too. Take an important government officer for example. His special position may require his location to be kept secret. His address is 239 Wall Street, New York City, New York, USA. Of course, he wants to keep his home address secret when he asks for LBSs. But he wants even his region—New York or USA—to be confidential. How to protect all the location information secret mentioned above in LBSs has become a hot topic. However, most new mobile social network applications, especially LBSs, publish user location information[2]. As a result, a large number of newer methods for location data confidentiality preservation are now being proposed; they employ randomization[3], space-vagueness[4], and time-vagueness[5]. Because of its moderate computation cost and easy implementation, space-vagueness is the most popular of these. k-anonymity is the main

• Jingjing Wang, Yiliang Han, and Xiaoyuan Yang are with the Department of Electronic Technology, the Key Laboratory of Network & Information Security of Armed Police Force, Engineering University of Armed Police Force, Xi'an 710086, China. E-mail: 344505421@qq.com; yilianghan@hotmail.com; xyyangwj@126.com.
∗ To whom correspondence should be addressed.

Manuscript received: 2016-02-09; accepted: 2016-04-26

approach to realize space-vagueness. It can obscure a precise location among k locations. To a user, k is static and is a measure of the degree of his privacy protection. To an attacker, k is dynamic, and restricts the precision of the location he can discover for a user. The bigger k is, the better the anonymity effect is.

2 Related Research and Problems

Location Privacy Protection Mechanisms (LPPMs) have been open challenges for decades. In the last ten years, many schemes have been proposed to allow users to make use of LBSs while mitigating privacy concerns[6–13]. These methods are used to let users submit false location data to prevent attackers from obtaining their real location information.

2.1 k-anonymity technology

In modern social networks, there has been considerable research on location privacy protection. As described in Ref. [6], location privacy protection technology based on k-anonymity has been impressively refined.

k-anonymity is an important information security concept. When publishing data, location data should be erased first; then the data should be generalized into clusters of k entries that published together[7]. k-anonymity is one way to realize space-vagueness. It can prevent an attacker from distinguishing the location of a particular user.

k-anonymity technology was first used in location privacy preserving by Gruteser and Grunwald in 2003[8], by generalizing user location into k adjacent access points in one region. Because it manifests as adjacent points in an area, it cannot protect the located region from leaking.

2.2 Location Privacy Protection Schemes (LPPSs) in LBSs

Research about how to protect location data in location-based services had a rather late start. In 2011, Huang et al.[9] proposed a method for location privacy preserving in location services, but the time for user to get location services increases greatly when k increases, it is very slow, and it also exposes the user's region.

In 2013, Damiani and Cuijpers[10] pointed out that the user's region is an important part of location privacy. They proposed a protection scheme based on privacy policy, namely using a controllable coordinate granularity mechanism to confine the precision of location gotten by the location provider, but it is not sufficiently efficient for mobile Internet use. In 2014, Yang et al.[11] proposed an LBS-oriented location privacy protection model and scheme. Their model took the central server as a completely trusted third party, which is always an additional threat to security and privacy in real applications. Focusing on the security problems of models and schemes, Peng et al.[12] proposed a method to judge the location privacy attacks in terms of the located region. In 2015 Wang et al.[6] proposed a location privacy protection approach named KAP, based on graph topology model, integrating the concept of k-anonymity. The approach does not take the location provider as a trusted party and has stronger security, but it needs a number of other access points' data around the user during every location service query. And if the user can not detect enough adjacent points, its algorithm will be invalid. Zhang et al.[1] put forward some classical methods and a model to evaluate LPPMs.

In conclusion, many existing studies are focused on the privacy problems introduced by the user's continuous location sharing in the worst cases, but most of the proposed solutions are not feasible in real applications, and some problems still need to be resolved urgently[13].

In summary, there are several problems in the research on location privacy preservation in LBS.

(1) In modern location services, the user does not get location data depending on his terminal devices as traditional GPS location technology does, but rather requests location services from a Location Provider (LP). When accessing a location service, the user's location data is generated in the LP. Then the data will be sent to the user through the network. Therefore, we should pay attention to helping the user get his location data securely, without leaking confidential location information. Typically, the LP is not to be trusted.

(2) In the existing schemes based on k-anonymity, the basic approach is to hide the user's real location data in an area including the user and at least $k-1$ other access points around him. It's important to note that through the schemes based on k-anonymity, the office's specific address is obscured with other $k-1$ different addresses in one region. The region may be "New York" or "USA" as mentioned in the example. But the officer's current region is always published by the existing schemes at the same time. So the privacy protection is not so perfect.

(3) In social networks, different users have various

kinds of location privacy protection requests. Even the same user may have diverse privacy protection needs at different times. How to protect and to what degree are important elements for user consideration. As such, a robust, dependable, secure, and efficient privacy protection system is needed. This system should support different privacy-preserving functions to respond to the user's different requests.

Our contributions in this paper are as follows:

(1) Two kinds of location-privacy protection models are given. One has a "third-party trust servicer", while the other does not. These models can be chosen freely according to the real environment.

(2) A generalized LPPS based on the Chinese Remainder Theorem (LPPS-CRT) is proposed. LPSS-CRT can guarantee the user's good LBSs without leaking any of their location information, even their regions. This is the main difference with other k-anonymity schemes. Our approach can respond to queries in response to different privacy-protection demands, with relatively low communication loads. It can achieve perfect privacy when needed, is more efficient than other similar schemes, and is suitable for environments in which user privacy-protection demands vary.

3 Our Scheme

To achieve perfect location privacy, we must protect the user's specific location. On the other hand, we must also pay attention to the user's region. And the method should be feasible and efficient. This will require a mathematical tool that obeys certain constraints.

(1) The tool can generalize a single locational identifer into k identifies, to support the basic function of k-anonymity.

(2) The generalized data should be such that an attacker cannot distinguish one locational entry from any other. Mathematically, the data consists of an equivalence set.

(3) The values of the k entries belonging to one equivalence class should not be similar; the k access points published should not be close to each other. If they are, the user's region information would be exposed, abrogating location privacy.

(4) The number k should be changeable, so as to meet different privacy demands in real applications.

(5) If some user asks for location services contiously for several times during a period of time, the generalized data published each time should be relatively reachable in a rational speed. Otherwise, the privacy of location would be exposed.

Surprisingly, the CRT meets all the above conditions. It can be used for location-data protection and has unexpected advantages in this application.

The CRT was first proposed in third–fifth century BC by Chinese militarist and mathematician Sun Tzu in his book *Sun's Mathematical Classic*, and was proved by Gauss in the 19th century. The CRT is a theorem about congruences in number theory and their generalizations in abstract algebra. It was a major contribution to the development of mathematics. The CRT has found wide application in modern cryptography, and is used in digital group signatures, threshold secret schemes, digital fingerprints, etc.

In this paper, in combination with the idea of generalized k-anonymity, a new kind of LPPS in location services based on the CRT is designed.

3.1 Basic definition

Definition 1 CRT: Suppose $\{m_1, m_2, \ldots, m_k\}$ are positive integers, and they are pairwise co-prime. Then, for any given sequence of integers: $\{a_1, a_2, \ldots, a_k\}$, there is some integer x that meets all the equations in Eq. (1) simultaneously. And there is no single solution of these equations; a solution set can involve different integers.

$$\begin{cases} x \equiv a_1 \bmod m_1, \\ x \equiv a_2 \bmod m_2, \\ \quad \cdots, \\ x \equiv a_k \bmod m_k \end{cases} \tag{1}$$

And if

$$\begin{aligned} M &= m_1 \times m_2 \times \cdots \times m_k, \\ M_i &= \frac{M}{m_i}, \\ M_i M_i^{-1} &= 1 \bmod m_i, \end{aligned}$$

then we can compute the solution set of Eq. (1):

$$X = M_1 M_1^{-1} a_1 + M_2 M_2^{-1} a_2 + \cdots + M_k M_k^{-1} a_k \bmod M \tag{2}$$

namely

$$\begin{aligned} X = {} & M_1 M_1^{-1} a_1 + M_2 M_2^{-1} a_2 + \cdots + \\ & M_k M_k^{-1} a_k + I \cdot M,\ I \in \mathbf{Z} \end{aligned} \tag{3}$$

Consider $X = \{x_1, \ldots, x_i, \ldots, x_j, \ldots\}$. Obviously, x_i is included in the solution set X.

Through Eq. (3), we see that I can be chosen randomly. If we choose I as different values, we can get different integers x_i to satify all the equations in

Eq. (1). According to abstract algebra, the solution set X is an equivalence class. We note that the elements of X are determined by different values of I. These elements have some relationship with each other too; that is, $M|(x_i - x_j)$.

Definition 2 Equivalence Class: In modern algebra, when a set has an equivalence relation defined on its elements, there is a natural grouping of elements that are related to one another, forming what are called equivalence classes.

Given a set x and an equivalence relation $\sim$ on X, the equivalence class of an element x in X is the subset of all elements in X which are equivalent to x.

Mathematically, the solution set X in Definition 1 is an equivalence set, and its different results with different I values all belong to the same equivalence class. Therefore, we can mix the user's real location data into an equivalence set including k elements by the CRT. Consequently, the k elements of the equivalence set would not be distinguishable by an attacker. We say that such a set has realized k-anonymity. If X is a set of location coordinates, and is an equivalence class, we say it can realize geo-indistinguishability.

Two kinds of k-anonymity schemes are shown in Figs. 1 and 2. In these two schemes, we take the location data of the user as a 3-tuple (longitude, latitude, time). For the sake of simplicity, we use the 3-tuple (x, y, t) to represent a location data instance.

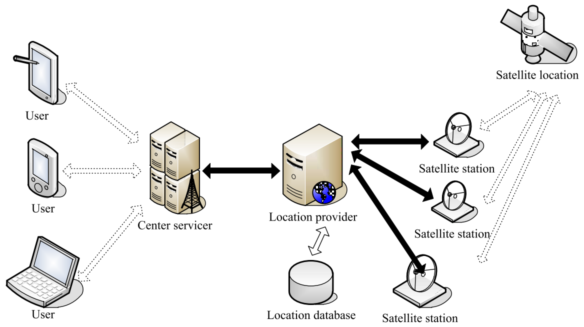


Fig. 1 Illustration for the LPSS-CRT model with central servicer.

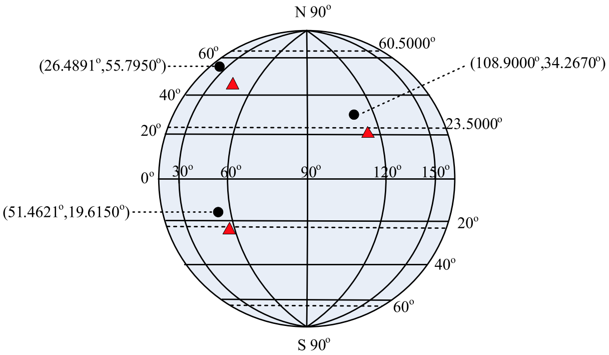


Fig. 2 Illustration for the toy example.

Definition 3 Generalized *k*-Anonymity (GKA): In location-based services, extend one user's accurate location data to k access points.

(1) These access points need not be adjacent neighbors in one region or interact with the user during location-based operations.

(2) These k access points must belong to an equivalence class, so that an attacker will not be able to tell which point is the user's real location.

From this definition, we see that GKA is more practical than the traditional definition[6–8] of k-anonymity in complex social networks. We say that if a scheme can realize GKA, it has realized k-anonymity.

3.2 LPPS-CRT with central service provider

Most research based on k-anonymity uses a trusted third party—a central servicer[8]. The main function of the central service provider is anonymity and agency query, which is necessary for getting LBS.

Notations used in this section are as folows:

x: longitude of the user's location;

x': integerizaion result of the longitude;

y: latitude of the user's location;

y': integerizaion result of the latitude;

t: time when the user asks for service;

v: velocity when the user asks for service;

c: query content of the user;

k: level of privacy protection requested by the user;

$I, I' \in \mathbf{Z}$: integers chosen in the process of CRT computation.

The LPSS-CRT model with central servicer is shown as Fig. 1

Step 1 Query: The user sends a message Q that includes his location information, query contents, and privacy request to the central anonymity servicer:

$$Q = \{x, y, t, v, c, k\}.$$

Step 2 GKA: The central servicer first expands the accurate location of the user into an equivalence set including k elements. Then it sends the query content and the k elements as a query set to the location provider.

We describe the scheme using an example illustrating the process in this step:

(1) Choose $m_1, m_2, \ldots, m_k$ randomly, satisfying the refinement condition:

$$m_1, m_2, \cdots, m_k \in \mathbf{Z}^+, \gcd(m_i, m_j) = 1, \ \ i, j \in \mathbf{Z}^+.$$

If the user does not want to leak his region information, make sure $m_1 \times m_2 \times \cdots \times m_k \gg 10^4$ (because here we take the precision of coordinates up to four decimal palces). Then the returned results are far apart from each other. If the user does not care about his current region information, $m_1, m_2, \ldots, m_k$ can be chosen randomly even its product is very small. In this case, the results returned are adjacent to each other.

To explain this process, we give a toy example, and choose geographic coordinates as the parameters in the example. In this example the parameters can be chosen freely.

A Toy Example:

Suppose

$$k = 3, (x, y) = (108.9000°, 34.2670°),$$
$$(m_1, m_2, m_3) = (13, 17, 113),$$
$$(n_1, n_2, n_3) = (13, 45, 92).$$

(2) Integerization:

$$(x, y) = (108.9000°, 34.2670°) \to$$
$$(x', y') = (x \times 10^4, y \times 10^4),$$
$$(x', y') = (1\,089\,000, 342\,670).$$

(3) Then we can get the equation sets based on the CRT:

$$\begin{cases} x' \equiv a_1 \bmod 13, \\ x' \equiv a_2 \bmod 17, \\ x' \equiv a_3 \bmod 113. \end{cases} \text{and} \begin{cases} y' \equiv b_1 \bmod 13, \\ y' \equiv b_2 \bmod 45, \\ y' \equiv b_3 \bmod 92. \end{cases}$$

Namely:

$$\begin{cases} X \equiv 3 \bmod 13, \\ X \equiv 14 \bmod 17, \\ X \equiv 19 \bmod 113. \end{cases} \text{and} \begin{cases} Y \equiv 3 \bmod 13, \\ Y \equiv 40 \bmod 45, \\ Y \equiv 62 \bmod 92. \end{cases}$$

(4) Compute the solution set of the equation sets above:

$$X = 15\,161 + 24\,973I,\ I \in \mathbf{Z};$$
$$Y = 19\,750 + 53\,820I',\ I' \in \mathbf{Z}.$$

(5) Choose two other data points $(x_i,\ y_j)$ randomly in the solution set, for example:

$$(x'_1, y'_1) = (264\,891, 557\,950), I_1 = 10, I'_1 = 10;$$
$$(x'_2, y'_2) = (514\,621, 1\,096\,150), I_2 = 20, I'_2 = 20 \tag{4}$$

The location coordinates extended from the real location are:

$$(x_1, y_1) = (26.4891°, 55.7950°),$$
$$(x_2, y_2) = (51.4621°, 109.6150°) =$$
$$(51.4621°, 19.6150°_{\text{s}}).$$

Here, the subscript "s" means south latitude. We can mark three access points computed in different areas in Fig. 2.

Besides, it is necessary to note that, the integers I and I' can be chosen randomly according to the user's request in reality. For example, assume the user is a businessman, who is always traveling to different cities to do his business all over the world. He wants to keep his location region as a business secret too and does not concern about the price of services. Then the integers I_2 and I'_2 that are greatly different from I_1 and I'_1 can be chosen for him in the continuous services, so that all the location coordinates extended from his real location are widely distributed in different regions all over the world. If the user lives in a single region, and does not mind other people knowing his location area, $|I_1 - I_2|$ and $|I_1 - I'_2|$ can be smaller. In this case, the extended location coordinates may be close to each other—even in the same region. In this way, our definition of generalized k-anonymity is realized. It can protect the user's privacy if needed, which is better than the results provided by the traditional definition of k-anonymity.

(6) Send the query message KAC to the location provider:

$$\text{KAC} = \{(x_1, y_1), (x_2, y_2), (x, y), c, t\}.$$

Step 3 LBS: The location provider offers the query results set QC to the central servicer:

$$\text{QC} = \{c_1, c_2, c, t'\}.$$

The central servicer finds the accurate result c in the set, and computes r according to the time interval and velocity:

$$r = v(t' - t).$$

Through the radius r and the original location shown in Fig. 3, the central servicer can judge the current area of the user. Then it can send c to the user in the located area timely. In this step, because the time of computing r is very short, so it almost can be neglected.

3.3 LPPS-CRT without central servicer

Although schemes with central servicer relieve users of the computation loads and have other advantages,

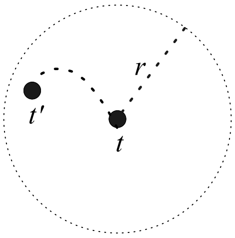


Fig. 3 Illustration for location-based service.

they have some weaknesses: (1) the central servicer is not always dependable, especially in complicated social networks; (2) the central servicer can be a performance bottleneck and the concentrated target may thus suffer many attacks; (3) the central servicer will require more computation and communication sources for agency query. So we must take the privacy preseving scheme without center servicer into account.

We give a description of this scheme in Fig. 4. Since the computation process of this scheme is similar to that of scheme in Section 3.2, we give a brief explanation.

Operation 1 GKA: The user chooses his level of privacy as k and the radius r of his region to be located next. r is the user's expected range of his next movements to ensure that the user can get accurate location results later.

Then the user will carry out the process of a k-anonymity computation based on the CRT to confuse the accurate location coordinates into an equivalence set including k different location coordinates. The method is the same as Step 2 of the scheme in Section 3.2, so here we do not give more details.

However, there are still some differences in this scheme. Since there is no central service r, if there is a communication interruption, the user cannot get a response to his location service request timely anymore. Since the user's terminal device has some computation and communication capabilities in modern era, we must take these resources into consideration to avoid this problem. He can interact with his neighbors by either single-hop or multi-hop mode to recover his communication with the location servicer and regain access to a location-based servicer.

Through the CRT computation, we can obtain the space-vagueness (k-anonymity) set of the location coordinates, which is

$$(X, Y) = \{(x_1, y_1), \ldots, (x_{k-1}, y_{k-1}), (x, y)\}.$$

Operation 2 Query: The user sends message Q directly to the location provider:

$$Q = \{(x_1, y_1), \ldots, (x_{k-1}, y_{k-1}), (x, y), c, r\}.$$

Operation 3 LBS: The location provider offers the query results set QC_i, $i \in 1, 2, \ldots, k$ to the k regions with radius r, which includes k different location coordinates received.

$$\begin{aligned}&\mathrm{QC}_1 = \{c_1, t'\} \text{ to } (x_1, y_1);\\&\mathrm{QC}_2 = \{c_2, t'\} \text{ to } (x_2, y_2);\\&\cdots\\&\mathrm{QC}_{k-1} = \{c_{k-1}, t'\} \text{ to } (x_{k-1}, y_{k-1});\\&\mathrm{QC}_k = \{c, t'\} \text{ to } (x, y).\end{aligned}$$

In the query results above, the message QC_k=$\{c, t'\}$ is the location service result for the user, whose location coordinates are (x, y).

The user can get his query content in his location region. It is important to say, the user's expecting radius r must be proper for the time interval spending on the interactive query process, or else the query content may be lost.

In this way, an LBS response is completed without a central servicer.

Further, because the location provider needs to support the computation loads for k access points simultaneously for protecting one user's location privacy in fact, if the privacy request k of the user is larger, the location privacy can charge more fees, which is very practical.

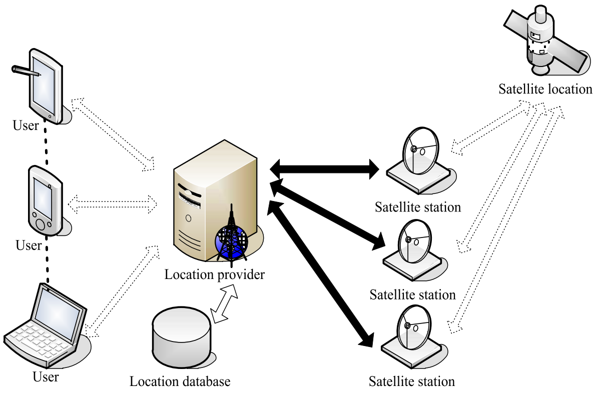


Fig. 4 Illustration of the LPSS-CRT model without central service *r*.

4 Discussion and Analysis

The privacy protection technology of location data not only needs to protect the user's location data, but also needs to balance the feasibility of services and overhead[14]. In this section, we discuss three aspects of the features and performance of our schemes, as suggested in Refs. [1, 11]: (1) degree of privacy preservation; (3) survivability of services; and (3) computation and communication overhead.

4.1 Degree of privacy preservation

As mentioned in Ref. [1], researchers have indicated that any notion of privacy is incomplete without explicit statements regarding the capabilities of an attacker. So when we begin to analyze a scheme, it is important to admit that the adversary can get some information from the interaction between the user and the LBS provider. In the following analysis, we will stipulate that the attacker can get some information. We can prove that LPPS-CRT can meet the conditions proposed by the classic papers: uncertainty, untraceability, and unpredictability.

Theorem 1: LPPS-CRT guarantees the user's location privacy request to a degree of at least θ.

The notations used in this section are as follows.

t_1, t_2, t': shown in Fig. 5. t_1 is the time of getting the former service; t_2 is the time of getting the latter service. For $t_1 < t' < t_2$, there is no service during the interval of time t_1 and t_2.

$P\{U_{\text{li}}^{t}\}$: probability that the user is at location "li" at time t.

D_t: all the location data published that can be collected by the attacker in the scheme with a central servicer at time t.

L_t: all the location data published that can be collected by the attacker in the scheme without a central servicer at time t.

θ: the highest location privacy leaking degree the user can accept, namely the best attack result that can be gained by an attacker.

Definition 4: If $P\{U_{\text{li}}^{t_2}|D_{t'}\} - P\{U_{\text{li}}^{t_2}\} \leqslant \theta$ for any time t', we say our scheme has guaranteed the user's location privacy request by a degree of at least θ. If $\theta = 0$, we say the scheme can guarantee perfect privacy for a user accessing the location service.

Proof: Each time a user asks for service, we get an equivalence class including his real location data by the CRT. In this equivalence class, there are k elements with equivalence relationship to each other. In theory, these k elements have a uniform probability of being chosen by the attacker, so we can say

$$P\{U_{\text{li}}^{t_2}\} = \frac{1}{k}.$$

For illustration, we discuss the situation of two successive user queries. If there are two successive requests for services in a period of time, the former occurs at time t_1 and the latter occurs at time t_2.

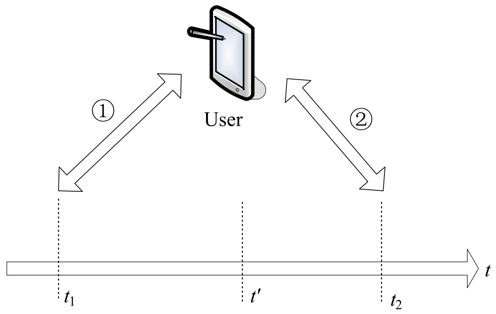


Fig. 5 Illustration for the time of two successive queries by a user r (① the former service; ② the latter service).

(1) For a scheme with a central servicer

All the published location data the attacker can get at time t' can be expressed as

$$D_{t'} = \{(x_1, y_1), (x_2, y_2), (x, y)\}.$$

Even if the attacker gets the three access points' location information shown in Fig. 6, he has no way to identify one from another, which is uncertainty. If I is fixed in these successive queries in the process of the CRT computation, at t_2, the attacker can not tell the user's current location from access points A, B, and C. Because computed by the CRT, A, B, and C are all reachable relatively to their original location and velocity published at the former query, namely at time t_1. So the probabilities of A, B, and C are uniform to the attacker, which is untraceability. That is to say, the location data $D_{t'}$ does not help the attacker judge the location of the user for the second time at time t_2, which is unpredictability. Mathematically, the probability is

$$P\{U_{\text{li}}^{t_2}|D_{t'}\} = \frac{1}{k}.$$

(2) For a scheme without a central servicer

All the published location data the attacker can get at time t' is

$$L_{t'} = \{(x_1, y_1), (x_2, y_2), (x, y), r\}.$$

Compared with the scheme with a central servicer, the attacker can get the expected radius r additionally as more information to predict the user's next location.

By the same token, for each successive query of several service:

$$P\{U_{\text{li}}^{t'}\} = \frac{1}{k}.$$

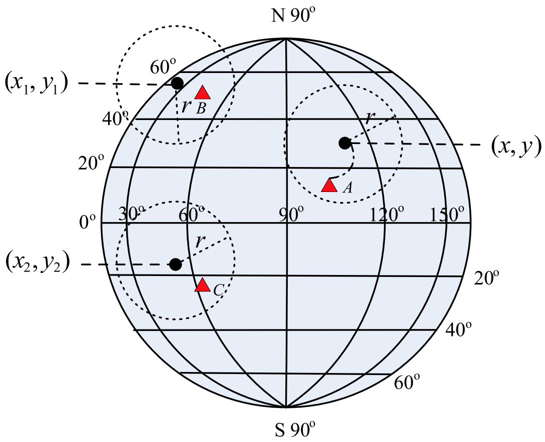


Fig. 6 Illustration for successive queries by a user. The dots stand for the extended location coordinates returned the first time; the triangles stand for the extended location coordinates returned the second time by the location service.

For example, as for two successive queries, we can see, in Fig. 6, that even if an attacker has information about the three different areas with radius r as the additional information to support his guess, i.e., to guess which area and access point in the three points A, B, and C is the actual current user location, he can not identify it. Because of radius r, the three areas are all respectively reasonable for the user's original location.

So we have the posterior probability,

$$P\{U_{\text{li}}^{t'}|L_{t'}\} = \frac{P\{U_{li}^{t'}L_{t'}\}}{P\{L_{t'}\}} = \frac{1}{k}.$$

Then through the difference privacy method proposed by Andrés and Bordenabe[15], we can prove the expression as follows:

$$P\{U_{\text{li}}^{t'}|L_{t'}\} - P\{U_{\text{li}}^{t'}\} = 0 \leqslant \theta.$$

The information entropy of the leaked location privacy to the attacker is

$$\sum_i P\{U_{\text{li}}^{t'}\}\log P\{U_{\text{li}}^{t'}\} - \sum_i P\{U_{\text{li}}^{t'}|L_{t'}\}\log P\{U_{\text{li}}^{t'}|L_{t'}\} = 0.$$

In fact, the different regions located by the radius also form an equivalence class. In this case, it also meets uncertainty, untraceability, and unpredictability conditions.

If we choose proper parameters in our scheme, we can prove that LPSS-CRT can guarantee perfect privacy, uncertainty, untraceability, and unpredictability, for users of LBSs.

4.2 Feasibility of services

Theorem 2: LPPS-CRT is feasible in real applications. It supports high quality of service and reliability of query results.

We prove its feasibility through two standards used universally: (1) quality of service; and (2) reliability of query results.

4.2.1 Quality of service

Definition 5: The quality of service can be judged by the proportion of the number of successful privacy requests n' as k changes from 1 to n, namely the success rate of anonymity.

Its math expression is

$$R_S = \frac{n'}{n} \times 100\%, \quad n' \leqslant n.$$

Proof: Because we used the classic math tool CRT to realize k-anonymity, when k is changed by the user, the only additional step required is to choose more or fewer values of I in the solution set according to different k:

$$X = M_1M_1^{-1}a_1 + M_2M_2^{-1}a_2 + \cdots + M_kM_k^{-1}a_k + I \cdot M.$$

And the only additional condition is that I is an integer. Whatever k is, the user can find enough elements with equivalence relationships, so as to realize changing privacy requests easily and smoothly, which means $R_S \approx 1$ if the service provider and the terminal devices run normally.

4.2.2 Reliability of query results

Definition 6: The reliability of query results can be judged by the relationship of the distance between the user's original location l_1 and the current location $l_{t'}$ when receiving a query result after a time interval and the radius of the user's location region.

For a scheme with a central servicer provider,

$$r = v(t' - t_1) = v \times \Delta t.$$

Here $|l_1l_{t'}|$ is the distance between location l_1 and location $l_{t'}$.

If $\dfrac{|l_1l_{t'}|}{v \times \Delta t} \leqslant 1$, we say the query result is reliable and accurate. Else, the query result can not be received by the user, because the user's new location has gone beyond the communication range accepted by the central servicer. In this case, the location-based service is invalid.

For a scheme without a central servicer, r is the expected radius of the user's location area, which is accepted by the location provider.

If $\dfrac{|l_1l_{t'}|}{r} \leqslant 1$, we say the query result is also reliable and accurate. Else, the query result can not be received by the user because the user's location has gone beyond the communication range accepted by the location provider. So the location service is also invalid. But we can avoid this case by choosing proper parameters in the scheme.

4.3 Performance

High efficiency is the main advantage of our scheme when applied to social or mobile networks.

(1) In schemes with or without a central servicer, the space-vagueness degree k can be changed easily without increasing computation overhead.

To explain, we use the toy example and its figures given in Section 3.2. When the user changes his privacy request k as $k = 4$ in the current service ($k = 3$ in the former service), the only overhead to add is choosing another integer I randomly (for example, we chose $I = 30$) and performing the computation of Eq. (4) in the CRT, the equivalence set of four data are shown in Fig. 7a.

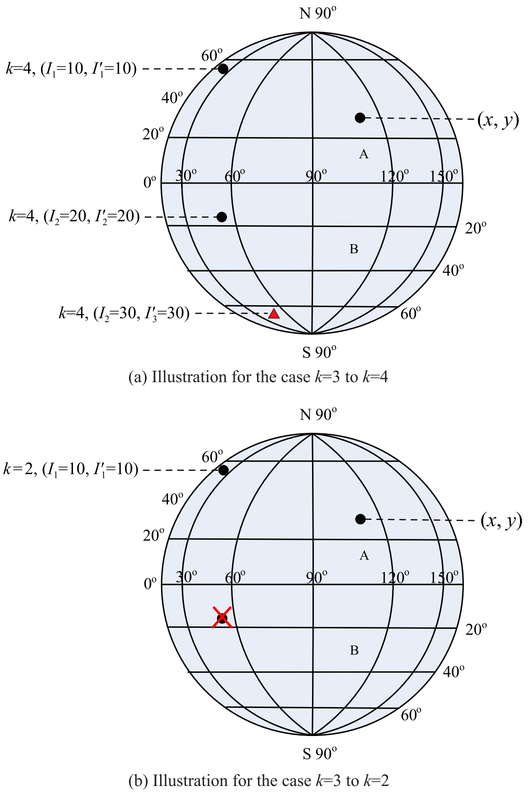
In another case, if the user wants to change k to 2, i.e., he'd like to reduce his privacy protection level, it is easier to realize. No more computation and new value of I are needed. On the contrary, we can directly delete one value of I already chosen in the last time service at once to realize smooth request change, the result is shown in Fig. 7b.

(2) Just as shown in Fig. 8, even k becomes larger, the computation and communication overhead of both center servicer and user's terminal devices is almost fixed because of the high efficiency of the CRT, which is better than the other existing classical schemes.

To get other access points data, the users in our scheme do not have to interact with other adjacent users for several times in LBS, which is more optimal than the existing schemes. Therefore, it can spare more communication time and network source.

5 Conclusion and Future Work

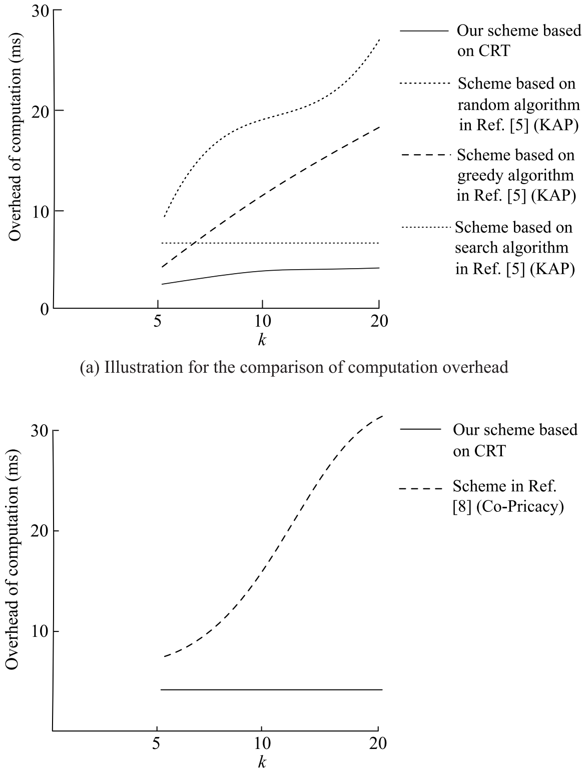
A new LPPS-CRT for perfect privacy in location services, based on the CRT, is proposed in this paper. In order to adapt to various practical environments, we design two models. One depends on a trusted third-party server and the other does not.



(a) Illustration for the case k=3 to k=4

(b) Illustration for the case k=3 to k=2

Fig. 7 Illustration for successive queries by the user in LBS. The black dots are the extended location coordinates in the last time service when k=3; the red triangle is the new fourth extended location data computed in the successive location service when k=4; the red cross means delete the access point's location data when k=2.



(a) Illustration for the comparison of computation overhead

(b) Illustration for the comparison of communication overhead

Fig. 8 Results for the overhead comparison of our scheme and classic schemes.

The scheme can be proved to achieve good LBSs for users without leaking the information of the user's location, even geo-location region. It thus overcomes the disadvantage of existing k-anonymity schemes that expose region information during interaction with neighbor nodes. It can also meet different location privacy preserving requests of users with high efficiency. Increasing the protection level of privacy does not increase the overhead significantly.

Our CRT-based scheme can resist the main attacks to location privacy, such as center points attacks, shape-center points offset attacks, indiscriminate attacks, etc. In summary, LPSS-CRT is an efficient, practical, and secure scheme.

Future research is on how to choose the parameters in our scheme to get optimal results with perfect privacy and the different user's privacy preference in LBSs.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Nos. 61272492 and 61572521), the Shaanxi Province Natural Science Foundation of China (No. 2015JM6353), and the Basic Foundation of Engineering University of CAPF (No. WJY201521).

References

[1] X. Zhang, X. Gui, F. Tian, S. Yu, and J. An, Privacy quantification model based on the Bayes conditional risk in location-based services, *Tsinghua Science and Technology*, vol. 19, no. 5, pp. 452–462, 2014.

[2] N. Jabeur, S. Zeadally, and B. Sayed, Mobile social networking applications, *Communications of the ACM*, vol. 56, no. 3, pp. 71–79, 2013.

[3] A. Suzuki, M. Iwata, Y. Arase, T. Hara, X. Xie, and S. Nishio, A user location anonymization method for location based services in a real environment, in *Proc. of the 18th ACM SIGSPATIAL Int'l Symp. on Advances in Geographic Information Systems*, 2010, pp. 398–401.

[4] B. Gredik and L. Liu, Protecting location privacy with personalized k-anonymity: Architecture and algorithms, *IEEE Trans. on Mobile Computing*, vol. 7, no. 1, pp. 1–18, 2008.

[5] X. Pan, J. Xu, and X. Meng, Protecting location privacy against location-dependent attacks in mobile services, *IEEE Trans. on Knowledge and Data Engineering*, vol. 24, no. 8, pp. 1506–1519, 2012.

[6] Y. Wang, H. Zhang, and X. Yu. KAP: Location privacy-preserving approach in location services, (in Chinese), *Chinese Journal on Communication,* vol. 35, no. 11, pp. 182–190, 2014.

[7] M. Wernke, P. Skvortsov, and F. Durr, A classification of location privacy attacks and approaches, *Personal and Ubiquitous Computing*, vol. 18, no. 1, pp. 163–175, 2012.

[8] 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 (MOBISYS 2003)*, San Franciso, CA, USA, 2003, pp. 31–42.

[9] Y. Huang, Z. Huo, and X. Meng, CoPrivacy: A collaborative location privacy preserving method without cloaking region, (in Chinese), *Chinese Journal of Computers,* vol. 34, no. 10, pp. 1977–1985, 2011.

[10] M. L. Damiani and C. Cuijpers, Privacy challenges in third-party location services, in *IEEE 14th International Conference on Mobile Data Management (MDM 2013)*, Milan, Italy, 2013, pp. 213–225.

[11] S. Yang, C. Ma, and C. Zhou, LBS-oriented location privacy protection model and scheme, (in Chinese), *Chinese Journal of Communication,* vol. 35, no. 8, pp. 116–127, 2014.

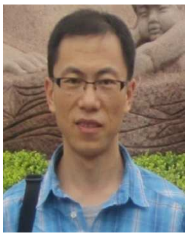
[12] Z. T. Peng, K. Kaji, and N. Kawaguchi, Privacy protection in Wi-Fi based location estimation, in *the 7th International Conference on Mobile Computing and Ubiquitous Networking (ICMU 2014)*, Singapore, 2014, pp. 62–67.

[13] X. Lin, S. P. Li, and Z. H. Yang, Attacking algorithms against continuous queries in LBs and anonymity measurement, (in Chinese), *Chinese Journal of Software*, vol. 20, no. 4, pp. 1058–1068, 2009

[14] L. Wang and X. Meng, Location privacy preservation in big data era: A survey, (in Chinese), *Chinese Journal of Software,* vol. 5, no. 4, pp. 693–712, 2014.

[15] M. E. Andrés and N. E. Bordenabe, Geo-indistinguishabilty: Differential privacy for location-based system, in *Proceeding of the 2013 ACM SIGSAC Conference on Computer Communications Security*, 2013, pp. 901–914.

Jingjing Wang received the BS (2008) and MS degrees (2011) in cryptography from the Engineer University of Armed Police Force. Currently she is working toward the PhD degree at Engineer University of Armed Police Force. She is also a teacher in Engineering University of Armed Police Force. She is a member of the CCF (China Computer Federation). Her research interests include applied cryptography, location-based privacy protection, and secure multi-party computation.



Yiliang Han received the BS (1999) and MS degrees (2005) in computer science from the Engineer University of Armed Police Force and the PhD degree (2011) in computer science and technology from Xi'an Jiaotong University. He is currently an associate professor in the Department of Electronic Technology at the Engineering University of Armed Police Force. He is a member of the CCF (China Computer Federation). He won the award of "Excellent Talents in Military" in 2008. Now he is also a PhD supervisor. His research interests include applied cryptography, security of computer networks, and ubiquitous computing.

Xiaoyuan Yang received the BS (1982) and MS degrees (1991) in information and electronic system from Xidian University. He is currently a professor and PhD supervisor at Engineering University of Armed Police Force. He is also the director of the Key Lab of Cryptography and Information Security of Armed Police Force. He won the award of "Excellent Professional and Technical Personnel" from the whole military and "Excellent Teachers" from China Education Ministry. He has published about 120 research papers, 4 research books, and 4 teaching materials. His research interests include cryptography protocols and post-quantum cryptography.