A Novel Personalized TTP-free Location Privacy Preserving Method

Nianhua Yang\textsuperscript{1, 2}, Yuru Cao\textsuperscript{2}, Qing Liu\textsuperscript{2} and Jiming Zheng\textsuperscript{2}

\textsuperscript{1}Department of Computer Science \& Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
\textsuperscript{2}School of Business Information Management, Shanghai University of International Business and Economics, Shanghai 201620, China

yangnh@sjtu.edu.cn, caoyuru2003@suibe.edu.cn, liuq@suibe.edu.cn, zhengjiming@suibe.edu.cn

Abstract

This paper proposes a novel TTP(Trusted Third Party)-free location privacy preserving method. A cloaking region is used to hide the precise position of the location-based service requestor (a mobile user). The cloaking region is formed through collaboration among neighbors. A mobile user can not only set personalized anticipated privacy requirements but also set minimum privacy requirements according to different contexts. The anticipated requirements will be satisfied if the time is allowed. Otherwise, the system will pursue the minimum standards for privacy requirements. This approach can satisfy k-anonymity, l-diversity and cloaking granularity simultaneously for privacy preserving.

Keywords: location privacy preserving, context sensitive, personalization, TTP-free

1. Introduction

With the development and popularity of wireless communication and mobile positioning techniques, location-based services (LBSs) have gained close attention in recent years. LBSs are value-added services provided by service providers based on the geographical location of the mobile device (requestor). LBSs can be divided into several types, such as navigation services, requesting the nearest business locations, receiving traffic alerts of notifications [1]. However, a user is required to send geographical location to the service provider in order to get LBSs. So the user’s location privacy may be revealed or misused by the service provider. Location privacy preserving has become a critical problem for the reason that most people are reluctant to use LBSs if their location privacy is in danger in spite of its convenience [1].

In order to protect location privacy in LBSs, different researches have been conducted based on k-anonymity [2], l-diversity [3] or cloaking granularity [4] approaches. For location privacy preserving, the exact geographic position value of a user is extended to a cloaking region. The anonymization server, which delegates the user, sends the LBS requests to the service provider based on the cloaking region substituting for the exact geographic position. For computing a cloaking region, methods based on k-anonymity, l-diversity or cloaking granularity hold different privacy metrics. k-anonymity based methods \cite{4-10} extend a cloaking region until ‘k-1’ other users are included. l-diversity based methods extend the cloaking region until ‘l-1’ different locations are included. Cloaking granularity requires the cloaking region to be larger than a user-specified threshold. So a service provider or an adversary is difficult to infer the exact geographic position of the requestor.

Although these before mentioned methods guarantee location privacy in some degree, each of them has a critical limitation. In a k-anonymity based cloaking region, some requestors...
may hold the same location value. And in an $l$-diversity based one, requestors may be in a very small populated area. They can’t be able to prevent the location disclosure. A granularity based cloaking region can’t defend against attacks for requestor identifies in the case where the locations are publicly known and there is only one requestor in the cloaking region [11].

![Figure 1. TTP-based System Architecture](image)

Most of previous location privacy preserving methods [5, 6, 8-10, 12, 13] rely on a centralized trusted third party based (TTP-based) architecture, such as shown in Fig. 1. A mobile user (requestor) sends its geographic location and query requirement to the centralized trusted anonymizing proxy. The location cloaker in the trusted anonymizing proxy extends the geographic location into a cloaking region according to the cloaking algorithm. Then, the cloaking region and corresponding query requirement are sent to the LBS provider by the trusted anonymizing proxy. After getting the query results for the cloaking region from the LBS provider, the trusted anonymizing proxy will select the exact results for the mobile user according to its exact geographic location from the results of the cloaking region by the query processor. At the end, the exact results will be returned to the mobile user. However, the trusted anonymizing proxy knows every exact geographic location of the requestor, and it may pose serious privacy threats. Furthermore, the trusted anonymizing proxy could be a system bottleneck in the TTP-based architecture.

In order to overcome these inherent drawbacks in the TTP-based system architecture, more and more researches have paid attention to decentralized location privacy preserving approaches [7, 14-18]. Most of these decentralized approaches are realized by collaboration among neighbors. Though these decentralized approaches can avoid performance bottleneck and hostile attacks on the anonymizing proxy, most of them just consider part of privacy metrics among $k$-anonymity, $l$-diversity and cloaking granularity.

As different users may have different privacy requirements, privacy requirement parameters in a location anonymization model should be personalized [6, 12]. Furthermore, the ability for forming a cloaking region in a system may be different under different contexts. In a populated area, a $k$-anonymity based cloaking region can be easily formed even if $k$ is very large. But it is hard to form the cloaking region in a sparsely populated region even if $k$ is very small. So the privacy requirement parameters in a location anonymization model should also be context-sensitive.

The paper is an extension of our previous work [19] which is included in the proceedings of ACN 2013. It proposes a novel context-sensitive personalized collaborative location privacy preserving method (CSPC). It corporates privacy metrics of $k$-anonymity, $l$-diversity and cloaking granularity which is represented as the side length of the cloaking region. The variable $s$ represents the side length of the cloaking region. It allows each mobile user in the system define a range for $k$, $l$ and $s$. When the upper bound of each of these parameters can’t be satisfied in a given time limitation, the system will test if the present cloaking region
satisfies the lower bounds of these parameters. If one of the lower bounds can’t be satisfied, it indicates that the cloaking region can’t be created. The region is obtained not by a centralized anonymizing proxy but by collaborations of the mobile users in the cloaking region.

Main contributions of this paper can be summarized as follows. First, the proposed approach corporates privacy metrics of \(k\)-anonymity, \(l\)-diversity and cloaking granularity. Second, parameters for privacy metrics can be adjusted among a personalized range according to the specific contexts. Last but not least, the approach is TTP-free. The cloaking region is computed by collaborations of the users in the cloaking region.

The remainder of this paper is organized as follows. Section 2 surveys related work. Section 3 introduces the architecture of our method. Algorithm of our approach is detailed in Section 4. Section 5 concludes the paper.

2. Related Work

In order to protect relational data privacy, \(k\)-anonymity [2, 5] is introduced to prevent the adversary from distinguishing an individual record from at least \(k-1\) other tuples. However, mobile users may hold the same location coordinates in a populated area when the \(k\)-anonymity approach is used for location privacy preserving. So it may result in location privacy disclosure. \(L\)-diversity [3] was proposed to tackle this problem. It ensures an anonymity group contains at least other \(l-1\) diverse geographic location values. But the cloaking region will still be very small in a populated area. A cloaking granularity [4] based method ensures the cloaking region to be larger than a user defined threshold. But it can’t defend against attacks for user identifies in the case where user locations are publicly known and there is only one user in the cloaking region [11]. So incorporation of these approaches is necessary for location privacy preserving.

From the perspective of system architecture, location privacy preserving methods can be categorized into two groups. One group includes approaches based on trusted third party (TTP) where a user (requestor) sends a query to the LBS provider through a trusted anonymizer (cloaker). While the other group includes approaches without relying on a TTP. A TTP-free cloaking region is produced through collaboration among neighbors.

A lot of researchers payed attentions to centralized methods for location privacy preserving. \(K\)-anonymity was firstly used in special and temporal anonymization to protect location privacy by Gruteser et al. [5]. They assume that all the mobile users hold the same value of \(k\) which represents the privacy requirement. It can’t provide personalized privacy preserving. Gedik et al., [6] propose a location privacy preserving method named CliqueCloak. It supports personalized privacy requirement parameter, i.e., personalized \(k\). But it only supports small value of \(k\) for computing complex. They also propose an advanced version [12] of CliqueCloak. The advanced version allows a requestor to assign its longest anonymizing time delay, which the requestor can tolerant and maximum cloaking region. Bamba et al., [13] propose a location privacy preserving method named PrivacyGrid which corporates \(k\)-anonymity and \(l\)-diversity approaches for privacy protection. Kalnis et al., [9] propose a \(k\)-anonymity based approach named Hilbert Cloak (HC) using Hilbert order [20]. It pursues minimal cloaking region and satisfies reciprocity among users in the cloaking region. Though it is appropriate for preventing location based identity inference, the location value of a mobile user is likely to be revealed in a small cloaking region. Hilbert curve is also used for computing cloaking region in a road network [8]. Researches in [10, 11, 21] concentrate on location privacy of moving users. Pan et al., [11] adopts both the location \(k\)-anonymity and cloaking granularity as privacy metrics and propose a new incremental clique-based cloaking algorithm, called ICliqueCloak, to defend against location-dependent attacks. The main idea is to incrementally maintain maximal cliques needed for location cloaking in an undirected
graph that takes into consideration the effect of continuous location updates. The graph is maintained by a TTP.

In a centralized architecture, the TTP becomes a performance bottleneck and may also be vulnerable to malicious attack [22]. Much researches have paid attention to TTP-free architecture which was realized through peer-to-peer communication. A TTP-free based cloaking region is formed through collaborations among neighbor users. Chow et al., [15] propose a peer-to-peer location cloaking algorithm. The main idea is that before requesting any location-based service, the mobile user will form a group from her neighbors via single-hop communication and/or multi-hop routing. Then, the cloaking region is computed as the minimum region that covers the entire group members. The original requestor will select a neighbor randomly from the group as the proxy which will substitute the requestor send the request to the LBS provider. Once receiving the querying results for the cloaking region from the LBS provider, the proxy will transmit them to the original requestor. And the original requestor will refine the results according to its geographical position. Ghinita et al., [23] propose a peer-to-peer cloaking method, called Prive, based on Hilbert curve. Solanas et al., [18] propose a TTP-free protocol for location privacy based on a public-key privacy homomorphism [24, 25]. Each user’s masking location value, got by adding Gaussian noise on the exact location value, is added into the encrypted data along a random-order chain. Thus, any user does not know the location of its neighbors even in the same group. The LBS provider can only know the masking locations in the group through decrypted data. Though the approach provides high location privacy, it relies on a public-key infrastructure (PKI) and does not support personalized privacy requirement. Hashem et al., [7] present a decentralized approach. Each mobile user holds a locally cloaked area (LCA). If a user requires a service from an LBS provider, the user’s current position is replaced by the globally cloaked area (GCA). GCA is the minimum bounding box of the union of the user’s own locally cloaked area (LCA) and the LCAs of k-1 other users. After the GCA is grouped, the user randomly selects one of its neighbors to forward the service request. The approach supports personalized privacy requirement. However, it does not require the group to meet the metrics of l-diversity and cloaking granularity.

Gedik et al., [6, 12] provide a context-sensitive personalized location privacy preserving framework. It enables each mobile user to specify the minimum level of anonymity as well as the maximum temporal and spatial resolutions it is willing to tolerate when requesting for a k-anonymity LBS. However, the framework relies on a centralized anonymity server.

3. The CSPC Architecture

With the development of mobile devices, computing and storage capabilities have been significantly improved. They can process complicated computing. So it is feasible to carry out location cloaking and query processing. CSPC is decentralized and the location anonymization is realized through collaboration among neighbor users. In addition, each user can assign a personalized minimum level as well as an anticipated level parameters for k-anonymity requirement. Similarly, it can assign a personalized minimum level as well as an anticipated level parameters for l-diversity requirement. Furthermore, the mobile user can assign a personalized minimum level and anticipated level of s for the side length of the cloaking region. Each user can also assign a maximum temporal value it is willing to tolerate when waiting for forming an anonymizing area. In a general area, the cloaking region should satisfy the requirements of the anticipated levels for k-anonymity, l-diversity and the side length of the cloaking region. If these requirements can’t be satisfied within the assigned temporal limitation, the requirements can be relaxed to the minimum levels. If these
requirements can’t be satisfied within the assigned temporal limitation, the user’s location can’t be normally anonymized.

![System Architecture for CSPC](image)

Figure 2. The System Architecture for CSPC

Figure 2 depicts the system architecture of CSPC. It contains two main components: mobile clients and LBS providers. Each mobile client holds a privacy profile that specifies its desired privacy requirements. A privacy profile includes seven parameters, \(k_{\text{min}}, k_{\text{nor}}, l_{\text{min}}, l_{\text{nor}}, s_{\text{min}}, s_{\text{nor}}\), and \(T\). \(k_{\text{min}}\) and \(k_{\text{nor}}\) represent the minimum level and anticipated level for \(k\)-anonymity requirement respectively. \(l_{\text{min}}\) and \(l_{\text{nor}}\) represent the minimum level and anticipated level for \(l\)-diversity requirement respectively. \(s_{\text{min}}\) and \(s_{\text{nor}}\) represent the minimum level and anticipated level for the side length of the cloaking region respectively. \(T\) represents the temporal limitation the user can tolerate for computing the cloaking region. The larger the value of \(k_{\text{min}}, k_{\text{nor}}, l_{\text{min}}, l_{\text{nor}}, s_{\text{min}}, s_{\text{nor}}\), and \(T\), the more strict privacy requirements a user needs. A mobile user can change their privacy profile according to the context at any time.

In this architecture, each mobile client equipped with a global position system (GPS) and two wireless network interface cards. The GPS device is used to censor its current geographic location. One of the network interface cards is dedicated to communicate with other mobile users by ad hoc networks, while the other is dedicated to communicate with base stations through the 3G telecommunication networks or with access points (APs) of the Internet. So it can communicate with an LBS provider.

4. The CSPC Algorithm

This section describes the data structure and the CSPC spatial cloaking algorithm.

4.1. Data Structure

In order to get a cloaking region for anonymity LBS query, the mobile user and its neighbors communicate with each other to discover at least other \(k-1\) neighbors until the cloaking region satisfies \(k\)-anonymity, \(l\)-diversity and cloaking granularity metrics. In addition, these requirements should be satisfied in \(T\) units of time.

Messages disseminated in the network can be divided into several types, \(i.e., type=\{FG, ACK\}\). A message belongs to \(FG\) is used to find nearest neighbors for forming an anonymity group. Likewise, a message belongs to \(ACK\) is used to reply the \(FG\) message.

Structures of a message broadcasted by the LBS requestor for searching nearest neighbors and the corresponding reply message are described in Def. 1 and Def. 2. The structure of a cloaking region is described in Def. 3.
Definition 1 (Message for searching nearest neighbors) A query message for searching the nearest neighbors is defined as \( Q = (M_{u}, h, \text{type}, S_{u}) \) where \( M_{u} \) represents the message sequence, \( h \) represents the hop distance propagated in the network, \( \text{type} = FG \) and \( S_{u} \) is the pseudonym of the message sender.

\[
Q = (M_{u}, h, \text{type}, S_{u})
\]

Algorithm 1 Spatial Cloaking

Input: \( R_{u} : \) LBS requestor

\begin{align*}
\text{Personalized request parameters } & k_{\text{min}}, k_{\text{sim}}, l_{\text{min}}, l_{\text{sim}}, s_{\text{min}}, s_{\text{sim}} \text{ and } T \\
\text{(}x, y\text{): } x \text{ and } y \text{ represent longitude and latitude of the current position}
\end{align*}

Output: A cloaking region for anonymity requesting

1. Let the hop distance \( h = 0 \)
2. Let the discovered neighbors set \( P = \emptyset \) and the number of discovered neighbors \( n = |P| = 0 \)
3. Let the diversity of locations \( d = 0 \)
4. Computing the LCA of the requestor \( L_{r} \) using the Alg. 2
5. Let the discovered LCAs set \( L = L_{r} \)
6. While \((n < k_{\text{min}} - 1 \text{ or } d < l_{\text{sim}} - 1 \text{ or one of the side length of the cloaking region is less than } s_{\text{sim}})\) and the expended time not exceeds \( T \) do
   7. \( h = h + 1 \)
8. Broadcast a message \( (M_{u}, h, \text{type}, S_{u}) \) with \( \text{type} = FG \)
9. \( A = \{(LCA_{i}, R_{j})\} \) is the set of neighbors that respond back to the requestor by executing Alg. 3
10. \( L = L \cup \{LCA_{i}\} \)
11. \( P = P \cup \{R_{j}\} \)
12. \( n = |P| \)
13. Set \( d \) with the number of different geographic locations of the discovered neighbors
14. Set \( x_{1} \) with the minimum longitude of the LCAs in \( L \)
15. Set \( x_{2} \) with the minimum latitude of the LCAs in \( L \)
16. Set \( y_{1} \) with the maximum longitude of the LCAs in \( L \)
17. Set \( y_{2} \) with the maximum latitude of the LCAs in \( L \)
18. \( CR = ((x_{1}, y_{1}), (x_{2}, y_{2})) \)
19. End while
20. If \( n \geq k_{\text{min}} - 1 \text{ and } d \geq l_{\text{sim}} - 1 \text{ and any side length of the cloaking region is not less than } s_{\text{sim}} \) Then
   21. Propagating \( CR \) to the neighbors within the cloaking region with \( h \) hops.
   22. Else
   23. cloaking failed
   24. End if

Definition 2 (Reply message for group forming) A reply message for group forming is defined as \( R = (M_{u}, A, \text{type}) \) where \( M_{u} \) represents the message sequence given by the requestor, \( \text{type} = \text{ACK} \), \( A = \{(LCA_{i}, R_{j})\} \) is the set of the tuples, each of which consists of locally cloaked areas (LCA) [7] and corresponding \( R_{j} \) which represents the pseudonym of the replier. An LCA is used to obfuscate the precise location of the mobile user. \( LCA = \)
where \( x_1 \) and \( y_1 \) are the longitude and latitude of the left button corner of the LCA respectively, while \( x_2 \) and \( y_2 \) are the longitude and latitude of the top right corner of the LCA respectively.

Definition 3 (Cloaking region) A cloaking region is defined as \( CR = ((x_1, y_1), (x_2, y_2)) \), where \( x_1 \) and \( y_1 \) are the longitude and latitude of the left button corner of the cloaking region respectively, while \( x_2 \) and \( y_2 \) are the longitude and latitude of the top right corner of the cloaking region.

4.2. Cloaking Algorithm

A CSPC area cloaking algorithm consists of following steps. Firstly, a mobile requestor broadcasts searching messages to find nearest neighbors. After finding enough neighbors to form a cloaking region which satisfies \( k \)-anonymity, \( l \)-diversity and cloaking granularity requirements, the mobile requestor computing the cloaking region and sends it to the members among the cloaking region.

In order to prevent a neighbor to get the exact location information, each neighbor response its LCA substituting the actual location to the message forwarder. Then, the original requestor computes a globally cloaked area (GCA) and broadcasts it to the neighbors in the cloaking group. The idea to use LCA for protecting location privacy from neighbors is firstly proposed by Hashem [7].

4.2.1. Searching Neighbors

Alg. 1 is the pseudo code for searching nearest neighbors. The original requestor \( r_o \) wants to get an LBS from the server. It firstly sets the broadcasting hop number \( h = 0 \), the set of discovered neighbors \( P \) to be null, diversity of \( d \) to be zero. The original LCA set \( L \) is set to be the LCA of \( r_o \), which is computed using Alg. 2. Then, \( r_o \) generates a union message sequence number and broadcasts a message for group forming with broadcast hops \( h = 1 \).

Algorithm 2 LCA Generating

Input: The position \((x, y)\) of the user
Personalized side length \( c \) of the LCA

Output: A rectangle defined by \((x_{left\ button}, y_{left\ button}), (x_{right\ button}, y_{right\ button})\)

1. Generate two random numbers, \( m \) and \( n \), where \( 0 \leq m, n \leq c \)
2. \( x_{left\ button} = x - m \)
3. \( y_{left\ button} = y - n \)
4. \( x_{right\ button} = x + c - m \)
5. \( y_{right\ button} = y + c - n \)

After sending the group forming message, \( r_o \) waits for replies from neighbors. Sec. 4.2.3 details the response done by the message receiver. After receiving the response message \( R = (M_{ack}, A.C.K) \), new discovered neighbors and their LCAs are added into corresponding sets. Meanwhile, cloaking region (CR) is computed based on the received LCAs. The most left button coordinate among LCAs is regarded as the left button coordinate of CR, and the most top right coordinate among LCAs is regarded as the top right coordinate of CR.
While the number of response neighbors is less than $k_{\text{min}} - 1$, or the number of different locations is less than $l_{\text{min}} - 1$, or the CR is not large enough, $r_c$ will broadcast the neighbor searching message again with $h = h + 1$ before timeout. The loop will terminate until the CR is created successfully or timeout.

If the loop is terminated for timeout, and the current CR satisfies $k_{\text{min}}$, $l_{\text{min}}$ and $s_{\text{min}}$, CR will be accepted, or the requestor can’t found appropriate CR for LBS requesting. It may try to requiring an appropriate CR later.

### 4.2.2. LCA Generating

LCA is firstly used by Hashem et al. [7] to protect location privacy among neighbors. The idea of Alg. 2 is modified from the Alg. 1 in [7]. A mobile user assign a number, $c$, which represents the side length of LCA. $m$ and $n$ are generated randomly with $0 \leq m < c$ and $0 \leq n < c$. And the coordinates of left button and top right are computed from the line 2 to 5 in Alg. 2.

**Algorithm 3 Receiver Response**

Input: an FG message $Q_s = (M_{id}, h, FG, S_{id})$

Output: $A : \{(LCA_{id}, R_{id})\}$

1. Let $r$ be the responser
2. If the $M_{id}$ is duplicate then
3. Reply the request forwarder with an ACK message
4. Return
5. End if
6. Computing the LCA of the requestor $L_r$ using the Alg. 2, and the result is represented by
   $((x_{\text{left button}}, y_{\text{left button}}), (x_{\text{top right}}, y_{\text{top right}}))$
7. If $h=1$ then
8. Send the tuple $((M_{id}, h, FG, S_{id}), (x_{\text{top right}}, y_{\text{top right}}), A \text{CK})$ back to the $S_{id}$
9. Else
10. $h = h - 1$
11. Broadcast a message $M_{id}, h, type, r)$ with type=FG
12. $A = \{(LCA_{id}, R_{id})\}$ is union of the response set to $r$
13. $A = A \cup \{(LCA_{id}, R_{id})\}$
14. Return $A \cup \{(LCA_{id}, R_{id})\}$
15. End if

### 4.2.3. Receiver Response

Alg. 3 describes the response of a neighbor when it receives a group forming requiring message. If the message ID is duplicate, it just reply with an ACK message. Otherwise, it will act according to the value of $h$. If $h=1$, it will just reply an ACK message with its LCA. If $h>1$, then $h$ is set to be $h-1$, and the message is rebroadcasted. The sender’s pseudonym of the message is replaced with the pseudonym of the forwarder itself. After receiving the responses from its neighbors, the forwarder replies the prior forwarder with its collected LCAs and its own LCA.
5. Conclusion and Future Work

This paper proposes a decentralized approach to protecting mobile location privacy. The cloaking region is formed through collaboration among neighbors to hide the precise position. A mobile user can set personalized privacy requirements at different contexts. The system will pursue the privacy requirement on an anticipated level before time out. Otherwise, the system will pursue to satisfy the minimum level of privacy requirement. The approach can satisfy $k$-anonymity, $l$-diversity and cloaking granularity simultaneously for privacy protecting.

In the future, experiments will be carried out based on the collected real GPS data of taxies in Shanghai to verify the effectiveness and performance of the approach.

Acknowledgments

This work was supported by Shanghai 085 Project for Municipal Universities, the Research Innovation Program of Shanghai Municipal Education Commission under grant No. 14YZ134 and Shanghai special scientific research funds for selecting and cultivating excellent young teachers in colleges and universities under grant in 2012.

References


Authors

Nianhua Yang received his BSc in Management Information Systems from Beijing Information Technology Institute, China, in 2001, MSc and PhD in Computer Application Technology from East China University of Science and Technology in 2004 and 2011 respectively. He is now a postdoctor in department of Computer Science & Engineering at Shanghai Jiaotong University. His current research interests include wireless ad hoc networks, location privacy preserving, reputation management, VANETs et al.

Yuru Cao received his BSc in Applied Mathematics from Anqing Normal University, China, in 2000, MSc in Applied Mathematics from Anhui Normal University, China, in 2003 and PhD in Computer Application Technology from Anhui University, China, in 2006. She is now an associate professor at Shanghai University of International Business and Economics. Her current research interests include data management, privacy preserving et al.
Qing Liu received his BSc and MSc in Computer Application Technology from Central China Normal University, in 1995 and 2001 respectively. She is now an associate professor at Shanghai University of International Business and Economics. Her current research interests include data management, Software engineering, Petri nets et al.

Jiming Zheng received his BSc in Foundry from Shanghai University of Engineering Science, China, in 1991, MSc in Management Engineering from Shanghai University, China, in 1998. He is now an associate professor at Shanghai University of International Business and Economics. His current research interests include data management and data mining.