Location Positioning and Privacy Preservation Methods in Location-based Service

Xu Zhang\textsuperscript{1} and Hae Young Bae\textsuperscript{1,2}

\textsuperscript{1}Department of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China
\textsuperscript{2}Department of Computer and Information Engineering, Inha University, Incheon, South Korea
zhangx@cqupt.edu.cn, hybae@inha.ac.kr

Abstract

Location-based services (LBS) become more feature-rich and versatile due to the explosion of mobile devices and the advances of positioning technologies. However, revealing personal private locations to potentially untrusted LBS server and others may raise serious privacy concerns if these locations are not protected adequately. In this paper, we present a survey of existing methods dealing with outdoor and indoor localization techniques and location privacy protection techniques. We propose a taxonomy that summarizes the state-of-the-art. This survey is intended to help researchers in quickly understanding existing works and challenges, and possible improvements to bring.

Keywords: Localization Technology, Location Privacy, Location-based Service, Privacy Preservation

1. Introduction

Recently, various location-detection technologies have been implemented to record indoor and outdoor movement, e.g., Global Positioning System (GPS), cellular networks, Wi-Fi and radio frequency identification (RFID). Location-based services (LBS) become more feature-rich and versatile due to the explosion of mobile devices and the advances of positioning technologies that have been implemented to record indoor and outdoor movement. Explosion of mobile data with location information may reveal personal private locations to potentially untrusted LBS server and others may raise serious privacy concerns if these locations are not protected adequately. There is a potential conflict of interest in LBS: the high quality of service and utility needs a precise description of location, while the users want to enjoy the LBS by not disclosing sensitive movements with anonymized location.

Although many people clearly consider their privacy a fundamental right, comparatively few can give a precise definition of the term and cannot distinguish it from security. Privacy is defined as the right of individuals to determine for themselves when, how, and to what extent information about them is communicated to others. Privacy indicates a specific form of data protection requiring flexible control over the disclosure of personal information. Location-based service facilitates human daily life, however, recorded location data enables intrusive inferences that may reveal personal habits, social customs, and religious that can be used for unauthorized advertisement.
As it is shown in Fig. 1, we conclude privacy protection issue with the consideration of spatial and temporal information. Data privacy concerns about privacy preservation problem with independent data at the left bottom. When considering privacy preservation in spatial dimension, each data is attached with geographic location. Thus, how to prevent adversaries from obtaining geographic related activities is the main issue in location privacy. Then we consider temporal features between data from snapshot to continuous. Continuously processed data involved dependency from past to future, which make it possible to infer hidden information behind data. This is concerned as stream privacy. In location-based service, mobile users issue location-based queries to LBS service providers to obtain information based on their geographic location. This is a new challenge to traditional data privacy-preserving techniques due to both temporal and spatial information should be concerned. Furthermore, temporal and spatial information should not be considered separately. Spatio-temporal feature may indicate significant implicit information, which should be protected.

The remainder of this paper is organized as follows. First, we give an introduction of general positioning technologies in location-based service in Section 2, and discuss privacy categories and system architecture of LBS in Section 3. Section 4 addresses the privacy preservation algorithms. And, the some specialized area like indoor environment, wireless sensor network and cognitive radio network are discussed in Section 5 and draw a conclusion in Section 6.

2. Localization Techniques

Localization of an object has long been the subject of research within the signal processing community and industry area like outdoor/indoor location-based services. Classic localization is based on the cooperative or non-cooperative use of RF emissions by the object to be located or RF emissions made by a set of anchor nodes and processed by the radio to be located.

As it is shown in Table 1, there are several metrics to classify existing works.

**Table 1. Classification of Localization Algorithms**

<table>
<thead>
<tr>
<th>Metric</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance/Angle</td>
<td>Range-based (TOA, TDOA, AOA, RSS) [13][14]</td>
</tr>
<tr>
<td></td>
<td>Semi Range-based [15][16]</td>
</tr>
<tr>
<td></td>
<td>Range-free [16]</td>
</tr>
<tr>
<td>Topology of</td>
<td>Centralized</td>
</tr>
</tbody>
</table>
A. Distance based Localization

In range-based algorithms, necessary information to estimate the distance can be obtained by some estimation techniques, such as Time Of Arrival (TOA), Time Difference Of Arrival (TDOA), Angle Of Arrival (AOA) and Received Signal Strength (RSS). As it is shown in Figure 4, basic theory and computation method of range-based localization is illustrated in detail. In TOA-based (Time Of Arrival) trilateration, range measurements to at least three base stations make up a set of nonlinear equations that can be solved to estimate the position of a unit. The PU time-tag the transmitted signal and the SUs measure the exact TOA of that signal. In TDOA-based (Time Difference Of Arrival) method, the time difference of arrival approach requires the ability to measure the time difference between the receptions of different devices. In AOA-based (Angle Of Arrival) method, the combined angle information to compute an intersection point of the target device. On the other hand, there is not enough information can be exploited to estimate the exact distance in the range-free algorithms. The semi range-based localization algorithm is a compromise between range-based and range-free method.

B. Centralized VS Distributed Localization

As it is shown in Fig. 3, in centralized localization, there is one central base station for computation. Thus, it suffers from overhead and cost increases. In distributed localization, computation is done by distributed server or nodes communication between each other to get their position in the network.
C. Cooperative VS Non-Cooperative Localization

Cooperative localization was first proposed in Japan to acquire real-time positioning information on mobile robots [17]. When mobile unit cannot independently determine its own position based on distance estimates with respect to the anchors (base stations), they can cooperatively find their positions. Generally, cooperative localization can dramatically increase localization performance in terms of both accuracy and coverage [18]. As it is shown in Fig. 4, base station (anchor) is needed for cooperative localization.

<table>
<thead>
<tr>
<th>(A) Non-Cooperative</th>
<th>(B) Cooperative</th>
</tr>
</thead>
<tbody>
<tr>
<td>(anchor)(Base Station)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Non-Cooperative & Cooperative Localization

3. Privacy Issue and Architecture in LBS

Location privacy and trajectory privacy stand for two aspects in location-based service. As it is shown in Table 2, location privacy generally considers independent locations, while trajectory privacy concerns more about continuous query and dependent locations.

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Location Privacy</th>
<th>Trajectory Privacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query</td>
<td>Explicit/Implicit query</td>
<td>Explicit/Implicit query</td>
</tr>
<tr>
<td></td>
<td>Snapshot query</td>
<td>Continuous query</td>
</tr>
</tbody>
</table>

Table 2. Privacy Issue in Location-based Service
Existing privacy preservation issues considered in LBS can be separated in three categories from the architecture perspective: Client-Server, Third Trusted Party and Peer-to-Peer Cooperative.

- **Client-Server Architecture**

  As it is shown in Fig. 5, all the users request for location-based service directly communicated with a single server to query information and perform an anonymization. It is easy to be implemented and integrated with existing technologies, however, the quality of service is low. Privacy issues concerned in this area mainly try to cheat the server with either fake locations or fake spaces. Featured privacy protection algorithms including CAP [42], PIR [43] and Dummy-Q [44].

- **Third Trusted Party Architecture**

  A trusted party server is deployed between users and service provider that take the response for gathering information, positioning and providing required privacy. As it is shown in Fig. 6, the trusted anonymizer receives exact locations from users, blurs locations and sends the blurred locations to service provider. This architecture can provide powerful privacy guarantees with high quality of services. However, it suffers same system bottleneck problems as client-server architecture, and it need sophisticated implementations. The common purpose of proposed methods with this architecture is to minimum anonymized queries and thus answer set to reduce both computation and communication overhead. Featured algorithms including Casper[45], DGCC [46], L2P2 [47] and IClque [48].

- **Peer-to-Peer Cooperative Architecture**

  In a peer-to-peer environment, users can collaborate with each other without the interleaving of a centralized server to reduce communication cost. As it is shown in Fig. 7, a trusted anonymizer is not necessary. However, a certificate could be applied to approve trustworthy users. There are two assumption scenarios: (1) both the peer users and service providers are malicious adversaries, (2) only service providers are malicious adversaries. Featured algorithm is P2P-IS-HL-CA[49].

---

**Figure 5. Client-Server Privacy Protection LBS Architecture**

**Figure 6. Third Trusted Party Privacy Protection LBS Architecture**

**Figure 7. Peer-to-Peer Cooperative Privacy Protection LBS Architecture**
4. Privacy Preservation Algorithms

Recently, various privacy protection techniques in LBS have been widely studied based on several concepts: privacy policies, false locations, space transformation and spatial cloaking.

Several methods have been proposed to relax trusted anonymization server and guarantee privacy in LBS. Private information retrieval (PIR) [43] can prevent any type of location-based attack, however, it incurs significant computational overhead on the server side and imposes stringent requirements on LBS server deployment. Dummy-Q method [44] is proposed to generate dummy queries with different attributes from the same location, however, generated dummy queries may be irrational and thus easily be identified by adversaries.

To relax the trusted third party assumption and lower the communication costs, Mokbel et al. in [45,49] proposed a new scheme that leverages the peer-to-peer concept. However, peers are assumed to be trusted, which made the management of trust relationships among autonomous peers in LBS remains an open issue.

A lot of research work has focused on anonymity and obfuscation-based techniques for privacy preservation. The k-anonymity and l-diversity are generally used. The k-anonymity based location privacy protection methods[50][51] propose to extend a cloaking region until k-1 other users are included, while l-diversity based location privacy protection methods take l-1 different locations into consideration.

The state-of-the-art privacy techniques can be categorized into four classes: anonymization[19][20][21], perturbation[22], differential privacy[23], and cryptographic[24] techniques. In location privacy protection, anonymization can be further divided into privacy policies, false locations, space transformation and spatial cloaking method [25].
The $k$-anonymity\cite{20} is the first and the most fundamental anonymization privacy model. The goal of $k$-anonymity is to ensure that each individual’s location is indistinguishable from at least $k-1$ other individuals’ location. Based on the notion of $k$-anonymity, many other anonymization models have been proposed including $l$-diversity\cite{19}, $(\alpha, k)$-anonymity\cite{21}. A large number of spatial cloaking algorithms have been proposed for protecting the location privacy of mobile users. Spatial cloaking techniques rely on $k$-anonymity concept and cloaking granularity, which blurs a user’s location into a cloaked spatial area that satisfies the user’s specified privacy requirements. Existing works on spatial cloaking follow the same idea to blur a user’s location into a cloaking region.

The basic of random perturbation is to replace the original data values with some synthetic data values so that the statistical information remains relatively the same while the original values never get disclosed. It has been adopted in many privacy preservation applications such as data mining\cite{26}, collaborative sensing\cite{27} and collaborative spectrum sensing\cite{28}. Differential privacy uses priori and posterior beliefs to guarantee the data privacy. For the location privacy, location data and sensing data from a user should be considered as a tuple in histogram data or contingency data table, and then it can be processed with differential privacy model. There is seldom cryptographic based study on location privacy due to the computation overhead.

5. Specialized Area

5.1 Indoor Environment

As we know, GPS is not available in indoor environment, most localization methods rely on signal strength, which have driven much more attention in research and development during the last decade. For the localization issue in indoor environment, angle and distance methods can be used to some extent. However, due to the complex environment, localization accuracy is easy to be influenced. For example, human movement, arbitrary walls and settings may influence the signal strength and thus output worse location estimation. WiFi fingerprint-based localization is regarded as one of the most promising techniques for indoor localization\cite{52}. Other methods proposed to use RFID or extra sensor nodes to improve localization accuracy. However, this is expensive and need pre-deployment.

5.2 Wireless Sensor Network

In WSN, localization is one of the most active areas of research in recent years because the location information is typically useful for coverage, deployment, routing, location
service, target tracking, and rescue. For a large number of sensor nodes, straightforward solution of adding GPS to all nodes in the network is not feasible. There are several challenges for locating sensor nodes needed to be solved. The first challenge is the energy consumption and localization accuracy problem. The second challenge is the NLOS ranging error problem. The third challenge is localization in low beacon density. Besides general adopted methods discussed in previous sections, there are some existing techniques which use two localization techniques such as multidimensional scaling (MDS) and proximity based map (PDM) [53] or MDS and Ad-hoc Positioning System (APS) [54].

5.3 Cognitive Radio Network

Localization problem in CRNs is in general different from localization in other applications such as Wireless Sensor networks (WSN) and Global Positioning System (GPS), in which the target to be localized cooperates with the localization devices. In contrast, a PU does not communicate directly with the CRs during the localization process [3,41]. Localization of primary user is crucial in enabling several key capabilities in Cognitive Radio Networks. There are various security concerns specific to CRNs, one of the most dominant threats among these is the Primary User Location Attack.

The SUs at different locations in the CRNs with a given interference ranges of the PUs, may perceive different profiles of spectrum holes due to different distances from the PUs. In order to reuse the unoccupied spectrum in an opportunistic fashion, it is important for the SUs to know the position information of the PUs. Localization is a methodology that can be adopted to obtain such kind of position information. Existing localization can be categorized into self-positioning, remote positioning and in terms of different localization objectives, where PU position estimation performed by the SUs belongs to remote positioning [4].

Privacy preservation has become a major issue across different applications, from information sharing to data publishing, from wireless communication to location-based services [2]. Location privacy was first introduced in mobile network, and then it arises with the open nature of wireless communication as well as software defined radio platforms in CRNs. Two types of location privacy issues in CRNs should be considered, namely, collaborative spectrum sensing location privacy and database query privacy [2].

The open nature of wireless communication as well as software defined radio platforms makes CRNs face many new challenges in the aspects of location privacy. Nowadays, the growing privacy threats of sharing location information in wireless sensor networks and cognitive radio networks have been widely concerned. The fine-grained location data may indicates user’s beliefs, regular activity and behavior. It may raise serious privacy concerns if these locations are not protected adequately. Being aware of such potential privacy risks, SUs may not want to share its data with fusion center or database. This safety consideration of SU may disable itself enjoying the benefit from collaborative spectrum sensing and database-driven CRNs if their privacy is not guaranteed. Therefore, it is essential to enable SUs to enjoy services provided by CRNs with privacy preserving approaches.

Location privacy issues in collaborative spectrum sensing are divided into two contexts: single-service-provider context and multi-service-provider context [2].

As it is shown in Fig. 9, six SUs are served by one FC, and sense three channels. Each SU sends sensing reports containing RSS values to FC, and FC combines the sensing reports to learn the spectrum. Since the sensing results are highly correlated to user’s physical location, which can be exploited by adversaries to launch location privacy attacks including:
As it is shown in Fig. 10, SU1–SU6 are served by three SPs, and sense three channels. Each SU sends sensing reports containing RSS values to its own SP. The three SPs exchange information with each other to collaboratively learn the spectrum.

In [28], a framework with two protocols to cope with privacy threats in the single-SP collaborative spectrum sensing was proposed. To prevent RLC attacks, it adopts secret sharing technique to enable the FC to obtain the aggregated results without knowing each individual sensing report. Privacy preservation is also studied in distributed settings, in which the aggregated results are derived from multiple partitions of data held by different entities. This is privacy preserving spectrum sensing in multi-SP scenario. The distribution setting are classified into vertical partitioning[35,36] and horizontal partitioning[37]. However, all these methods fall short under the collusion attacks in multi-SP context. Thus, more strict privacy preservation schemes that are specially designed for multi-SP collaborative spectrum sensing are required, which currently is still an open issue.

Besides spectrum sensing, geo-location database query approach is another typical approach to obtain spectrum availabilities at SU’s location. The database query approach is enforced by the latest FCC’s rule released in May 2012[38,2]. In[39], it discusses the impersonation attacks towards master device, database and man-in-the-middle-attack between SUs and DB. The database is assumed to be semi-honest or an easy-to-be-attacked, that is, the database exactly follows the protocol but tries to infer SU’s locations. Potential privacy threats can come both from database and secondary users. In [40], the knowledge of database is assumed to include the complete communication content between SU and the database, and the spectrum utilization information of SUs. Instead of directly learning the SUs’ locations from their queries, some attacks can infer an SU’s
location through his used channels. They show a new kind of location privacy attack, Spectrum Utilization based Location Inferring (SULI) attack. They propose a novel Private Spectrum Availability Information Retrieval (PSAIR) scheme that utilizes a blind factor to hide the location of the SU. To defend against the discovered attack, a novel prediction based Private Channel Utilization (PCU) protocol is proposed, which reduces the possibilities of location privacy leaking by choosing the most stable channels.

6. Conclusion

In this paper, we study the fundamental techniques for user localization and privacy preservation mainly in location-based service, and also discussed in specialized areas such as indoor environment, wireless sensor network and cognitive radio network. The objective is to provide a comprehensive analysis and guide of existing efforts. This survey is intended to help researchers in quickly understanding existing works and challenges, and possible improvements to bring.

Acknowledgement

This work was financially supported by the Natural Science Foundation of China (41201378), Natural Science Foundation Project of Chongqing CSTC (cstc2012jjA40014), Young Scientist Foundation of Chongqing University of Posts and Telecommunications (A2013-36), Talent Foundation of Chongqing University of Posts and Telecommunications (A2014-06), and Creative Foundation of WenFeng (A2014-12).

This paper is a revised and expanded version of a paper entitled “Privacy-preserving Localization in Cognitive Radio Networks” presented at FGCN, Hainan, China on 22th DEC.

References


Authors

Xu Zhang, he is an assistant professor of Chongqing University of Posts and Telecommunications, received Ph.D degree from Inha University, South Korea. His research area mainly includes ubiquitous computing(sensor network, localization), large scale data processing, database, etc.
Email: zhangx@cqupt.edu.cn

HaeYoung Bae, he is a tenured full professor of Inha University of Korea, and he is an honorary professor of the Chongqing University of Posts and Telecommunications of China. His research area mainly includes database and spatial information processing.
E-mail: hybae@inha.ac.kr