A PRIVACY PRESERVING APPROACH TO PROTECT LOCATION IN LOCATION BASED SERVICES

DIVYESH BHANDARI [1], PRAMOD GOPHANE [2], PRADNYA GADE [3]

[1] Dept. of Computer Engineering, Sinhgad Institute of Technology, Lonavala , Savitribai Phule Pune University
[2] Dept. of Computer Engineering, Sinhgad Institute of Technology, Lonavala ,Savitribai Phule Pune University
[3] Dept. of Computer Engineering, Sinhgad Institute of Technology, Lonavala ,Savitribai Phule Pune University

ABSTRACT : *The increasing number of mobile device users indicates the expansion of personalized location-based services (LBS). Despite their proliferation, the risk of violating users' privacy by exposing user's location information remains. For this reason, many studies have researched to prevent privacy violation in LBS. However, previous researches only focused on protecting users' location information without considering semantic location privacy violation through contextual information. In this paper, we survey the process of inferring a user's behavior using semantic information which includes spatial and temporal information. We also suggest a privacy preserving technique to prevent exposure of sensitive behavior in semantic LBS. The proposed b-diversity technique is validated to prevent exposure of sensitive behavior and also minimizing data utilization degradation.*

Keywords - LBS, privacy protection, security, Smartphone, decision tree, b-diversity, k-anonymity, Cloaking Area.

I. INTRODUCTION

The convergence of technologies such as *Geographic Information System* (GIS), networks and mobile devices give birth to *location based services* (LBS), In these services, requests which contain user's locations are sent to *Third Service Provider* (TSP), and TSP then responds with some needed information for the users. Common examples are finding the nearest *Places of Interests* (POI) such as gas stations, hospitals, *etc*. In this process, user's privacy like locations is open and threatened. It has risk to expose a user's location history and contextual information to malicious adversary. For instance, user's political beliefs can be inferred by specific information such as duration of stay in a specific building occupied by political parties. Privacy is generally the information that you don't want others to know. There was news which reported tracing other people with GPS before. With the popularity of LBS, users' privacy information has aroused much concern. Therefore, in order to provide safe and reliable location-based services, user's privacy should be guaranteed.

Users' concern of location privacy is legitimate. Location data is sensitive since it can reveal where you live and work, where you go for movies and dinner and even if you stay at someone else's house. location privacy is *the ability to prevent other parties from learning one's current or past location*. In the scenario of LBS application, each user's LBS request will leave a footprint in LBS provider, and LBS provider, which is semi-honest, may sell these footprint records for commercial benefit. In this context, the above mentioned definition of location privacy contains two requirements:

• User's don't expect to report their real locations to LBS provider every time.
• It is difficult for other parties (include LBS provider) to get to know the link between user ID (which is used to communicate with LBS provider) and their real-world identities.

II. RELATED WORK

There are two classes of LBS: satellite based class and web based LBS. We continued to discuss security mechanisms from two aspects: policy-based and computational-based techniques. Policy-based mechanism can support more flexible protection, but the precondition is that LP is trusted. Otherwise, it will not work correctly. By this mean, computational-based technology is essential, which can guarantee the validity of privacy protection. They should become mutually complementary. As for the users, they should make a tradeoff between the benefit and privacy threat brought by LBS. Privacy preserving techniques in LBS are classified into location anonymization which hides a user's exact location information, and identity anonymization which prevents user identification. Location k-anonymity takes k-1 other user's locations to reduce the possibility of pinpointing a user's exact location. This method is the most popular technique for location anonymization.

Location anonymization techniques are further classified into spatial cloaking, space transformation, fake location method, k-anonymity, obfuscating method.

(I) Location Anonymization Techniques

K-anonymity: K-anonymity was proposed by Latanya Sweeney from Carnegie Mellon University, firstly used in privacy protection about data publication in relational database systems. K-anonymity requires that one piece of data cannot be distinguished from at least other k-1 pieces of data. That is, when one user's location cannot be distinguished from other k-1 users', it meets k-anonymity. Now location k-anonymity has been widely extended to protect users' privacy, and the most famous are strong k-anonymity, l-diversity, t-closeness. K-anonymity can not only protect users' location but the query privacy.

Spatial cloaking: It has been widely used to tackle privacy issues in LBS. The basic idea of spatial cloaking techniques is to blur or generalize a user's exact location into a cloaked area to satisfy location k-anonymity. Spatial cloaking is a simple technique; however it deteriorates the quality of service and require additional computation cost.

Space transformation: It is a method that transforms a user's location into another space to hide his/her location, while maintaining the spatial relationship. This method is more efficient than spatial cloaking; however the calculations involved are complex, and are unable to support various query types, such as range query.

Fake Data: Fake data methods protect users' privacy by sending fake data instead of real data to service providers, such as using pseudonym to protect the user's ID. By sending false locations which are called dummies, users' real location can be protected. This method can be easily implemented, because users themselves can make dummies. The degree of privacy protection depends on the distance between false and real location. The fake location does not deteriorate the quality of service, but disclosing dummy information is possible. Moreover, the creation and management of dummies is expensive.

Obfuscation: Obfuscating method protect users' definite location by reducing the location precision, mainly includes spatial obfuscation and spatio-temporal obfuscation. Spatial obfuscation extends users' definite location to a larger area which containsusers' location, such as a circle. Spatio-temporal obfuscation adds time obfuscating on the basis of spatial obfuscation.

1. Spatial Obfuscation**:** One example is using a circle area instead of a point to represent users' location. Since the user can be randomly located in any point in the area, attacker only knows the user is in this area but the accurate location is obscure. The advantage of spatial obfuscation is simple implementation;users themselves can specify obfuscation area. But if the users' definite location is replaced with an area, the quality of service will decrease. Therefore, how to make a balance between privacy protection and quality of service is worth studying.

2. Spatio-Temporal Obfuscation: Spatio-temporal obfuscation reduces the precision of not only location but the time-related information so as to satisfy the predefined k-anonymity standard.

(II) Identity Anonymization Techniques

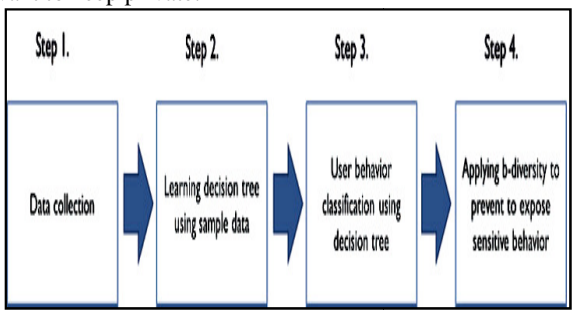
Identity anonymization uses pseudonyms to prevent user identification by hiding information, such as name and Social Security Number (SSN). However, if a certain pseudonym is used for a long time, it can eventually be inferred. A previous work has introduced the concept of mix zones. A mix zone is an area that assigns new pseudonyms without exposing the relationship between the user's old and new pseudonyms. Pseudonymizers scheme add a trusted third party (a pseudonymizer) to the basic model. The pseudonymizer mediates between users and providers. Users send their queries to the pseudonymizer, which replaces the real identity
of the users (e.g. their IP addresses) by a pseudonym. This way, providers cannot identify users because they become hidden behind the pseudonymizer. Notwithstanding, users *must trust* pseudonymizers because *they have full access to their real locations and identities*. Also, if users send several queries from the same location (e.g. from their residence), providers can determine their real identities by using e.g. a public telephone directory. These attacks are known as Restricted Space Identification (RSI) and Observation Identification

III. PROPOSED SYSTEM

In general, each location has a unique behaviour based on location's type. For instance, specific behaviour that related with health care occurs in hospitals. In government offices, public service activities are performed. Each unique behaviour has a pattern, such as residence time, frequency, in/out time, and unique behaviours, which can be distinguished using these patterns. For instance, if a certain user's residence time pattern is usually 6 hours and his in/out time pattern is 09:00 to 18:00, we can infer that the user is someone who performs medical related work. Hence, the unique behaviour is determined by the location type, and such behaviours can be inferred based on their own pattern..

We focus on a privacy protection technique to prevent the exposure of the sensitive behavioural information in semantic LBS. This technique uses the spatial and contextual information of the user to provide a personalized service. We propose a b-diversity method to reduce the probability of sensitive behaviour inference to 1/b by

maintaining b or more candidate behaviour that can be inferred. The proposed semantic location privacy preserving technique that generates cloaking area using user's context including spatial and temporal information. This technique is used to prevent inferences with respect to user's sensitive behaviour that user want to keep private.



Overview of the proposed technique

Algorithms:

ID3: uses information entropy to calculate information gain, because this is typically used to generate the decision tree. Clearly, other decision tree algorithms can be used.

B-Diversity: The proposed b-diversity is used to protect a user's semantic location privacy by adding noise or generating cloaking area to infer b or more behaviours, which reduces the probability that user's behaviour can be inferred to less than 1/b.

Modules:-

Geofence Marking: Users will have facility in their android application to mark the circular area.Where if user goes it will trigger an alarm.

Semantic behaviour pattern Extraction: Suppose if user is visiting his office daily from 7 to 9. Then we will consider it as a pattern and we will tag it as "Office Working". Similarly with other scenarios. We will show these matched patterns to user on his application

Behaviour Hiding: Will show fake location of user on Google maps in this case.

IV. CONCLUSION

With the development of powerful smart phones, LBS will become more and more popular. When enjoying these convenient services, users need to provide their privacy information such as location which is likely abused. Many kinds of existing privacy preserving technologies are reviewed in this paper. Then we continued to discuss security mechanisms from two aspects: policy-based and computational-based techniques. Policy-based mechanism can support more flexible protection, but the precondition is that LP is trusted. Otherwise, it will not work correctly. By this mean, computational-based technology is essential, which can guarantee the validity of privacy protection. They should become mutually complementary. As for the users, they should make a tradeoff between the benefit and privacy threat brought by LBS.

Privacy protection is an important research challenge. Although progresses have been made in recent years, lots of problems are still unresolved and more solutions are needed to put forward.

V. REFERENCE

[1] Julong Pan [1],Zhengwei Zuo [1], Zhanyi Xu [1]andQun Jin [1, 2], "Privacy Protection forLBS in Mobile Environments: Progresses, Issues and Challenges," in proc International Journal of Security and Its Applications Vol.9, No.1, 2015, pp. 249-258.

[2] Yuna Oh*, Kangsoo Jung, Seog Park, "A privacy preserving technique to prevent sensitive behavior exposure in semantic location-based service," in proc 18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems - KES2014, 2014, pp. 318 – 327.

[3] Zheng Jiangyu[1,2], Tan Xiaobin[1,2], Cliff Zou [3], Niu Yukun[1,2] and Zhu Jin [1,2], "A Cloaking-Based Approach to Protect Location Privacy in Location-Based Services," in proc Proceedings of the 33rd Chinese Control Conference July 28-30, 2014, Nanjing, China, 2014, pp. 5459 – 5464.

[4] Fizza Abbas, Rasheed Hussain, Junggab Son and Heekuck Oh, "Privacy Preserving Cloud-based Computing Platform (PPCCP) for using Location Based Services," in proc 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing, (2013), pp. 60-66.

[5] Agusti Solanas, Josep Domingo-Ferrer, and Antoni Mart´ınez-Ballest´, "Location Privacy in Location-Based Services: Beyond TTP-based Schemes", pp. 1-12.

[6] Pablo A. P´erez-Mart´ınez, Agusti Solanas, Antoni Mart´ınez-Ballest´, "Location Privacy Through Users' Collaboration: A Distributed Pseudonymizer," in proc Third International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, 2009, pp. 338-341.