Privacy in Location-based Services

Kandiraju Sai Ashritha and Prajwala T.M

Abstract With the growth of hand-held gadgets like mobile phones, tablets and smart phones, Location-Based Services have gained immense popularity. In order to cater to user's demands, a Location-Based Service server would need location information of the individual requesting the service. Providing geo-spatial information could pose serious threats to the Location-Based Service recipients in terms of their privacy. Privacy of the users could be of two types: Location privacy and Query privacy.

Location privacy refers to potential abuse and misuse of an individual's personal information. An individual's actions, whereabouts, movements, preferences, priorities and other such information can be inferred from the sensitive geo-spatial data provided. Query privacy refers to the compromise on the disclosure of query data. The anonymity of service recipients with regards to their query content is also extremely critical in preserving the privacy of the users. Thus the issue of privacy of user location has attracted immense attention from academica as well as industry. A survey on the state-of-techniques used for privacy protection in Location-Based Services has been presented in the paper. Open challenges, scope for research and future work have been elaborated towards the end.

1 Introduction

Location-Based Services (LBS) use real-time geo-information from a user's smartphone and provide various services based on the user location. These services encourage users to ”check in” to various malls, live events, restaurants, coffee shops

Kandiraju Sai Ashritha
NITK, Surathkal, e-mail: ashritha1615@gmail.com

Prajwala T.M
NITK, Surathkal e-mail: prajwala.tm@gmail.com

and other places. Location-Based Services are capable of finding the location of a person or an entity, like identifying the nearest medical store or a coffee shop, or tracking the location of a family member or a friend. Location based administrations utilize a smart phone's GPS technology to track an individual's area, if that individual has enabled the feature in his/her phone. After a phone client selects in, the administration can determine his or her area down to a road address without the requirement of any kind of manual information.

With the growth of hand-held gadgets like mobile phones, tablets and smart phones, Location-Based Services has gained immense popularity. The increase in position determination capabilities of a smart phone has facilitated in the extensive use of Location-Based Services (LBS) for various applications. These services have become an irreplaceable source of real-time information and guidance. As a result they are extensively used in an individual's day-to-day life in the following ways:

- **Locating Stores :** Finding the nearest medical, general and retail stores.
- **Real-time Information :** Obtaining real-time updates regarding traffic, weather etc.
- **Location based Marketing :** Marketing based on the proximity of potential customers.
- **Entertainment :** Suggesting places to visit, movies to watch etc based on the locality

In order to cater to user's demands, a Location-Based Service server would need location information of the individual requesting the service. Thus providing geospatial information could pose serious threats to the Location-Based Service users like compromise on privacy and security, pricing, accuracy in handling the data provided etc. Among the issues described, the most important issues that need to be addressed would be protecting the privacy of the users.

Privacy of the service recipients could be of two types[1] :

1. **Location Privacy :** Location privacy refers to potential abuse and misuse of an individual's personal information. For example, an adversary could associate a certain health condition to a user who queries a Location-Based Service from a hospital location.
2. **Query Privacy :** Query privacy refers to compromise of the service attribute. For example, with the increase in the number of queries obtained from a bar, the adversary could assume that the user is alcoholic.

1.1 Location Privacy

Having access to a user's location information provides an opportunity to trace down a user thus making this a serious privacy concern. An individual's actions, whereabouts, movements, preferences, priorities and other such information can be inferred from the sensitive geo-spatial data provided. To explain it in a precise manner, a user's identity, individuality and integrity could be jeopardized from his/her location information[2]. There is every possibility of a privacy breach, if an adversary could hack an LBS server to collect location information of its users which in turn would help deduce sensitive private information of the service recipients such as the locations of their home, office, their preferences, their official associations, their religious views, their lifestyles etc.

To state an example, Hoh et al.[3] and Krumm [4] mentioned that the location of a driver's house can be obtained from the GPS data collected on his/her vehicle even if the data regarding location were anonymized. In another example stated by Matsuo [5], an individuals indoor location data could be exploited to gather sensitive personal data, like age, if the user is a smoker or not, if the user drinks coffee often or not, position at work and other information. The consequences of a location leak can be embarrassing because people would not want to be seen at certain locations places like a gynaecologist clinic, rehabilitation center, AIDS clinic or any place related to a specific religious/political ideology[6] .Thus location privacy of the users is a serious concern and this problem has received considerable time and effort from various researchers, LBS providers, several business organizations and even governments.

1.2 Query Privacy

Query privacy is somewhat closely related to Location privacy. Location privacy deals with preserving the location of the user whereas query privacy deals with preserving the query content itself and making sure that this sensitive information would not reach adversaries. For example, a user querying about casinos would imply that the user has a habit of gambling which he/she would not want to disclose it out to public. Thus apart from location privacy, the anonymity of service recipients with regards to their query content is also extremely critical in preserving the privacy of the users. In simple terms,query privacy is nothing but preventing adversaries from capturing information about the issuers of queries.

One way to achieve this would be to anonymise queries by eliminating issuers identities, so that even if the content is known the adversaries would not be able to figure out who made the query. But a problem with this method is that, since attackers can still obtain users locations in several of ways, e.g., triangulating smart phones signals and localizing users access points to Internet. Most of the times, publicly available information such as home address locations and yellow pages will also help significantly in acquiring issuers locations. Therefore, it is extremely important to achieve privacy with respect to user queries as well.

2 General Location-Based Service Architecture

A complete and common Location-Based Service architecture comprises of several units such as data providers, network operators, virtual operators, financial bodies, service providers and other parties. The major entities as illustrated in Fig. 1 are :

1. Mobile Devices that provide user location information
2. Positioning systems
3. Network Operators
4. Service administrators or Service providers

The users query the servers of the Location-Based Services using their smartphones. The locations embedded in the queries are acquired with the help of positioning systems such as the Global Positioning System (GPS). The queries, as well as their responses from the Location-Based Service servers, are transferred over the communication networks. The Location-Based Service providers are the servers which cater to the user queries by sending out well tailored responses by taking into account the location information embedded in the queries.

3 General Threat Model

The generally accepted threat model for LBS privacy protection treats LBS servers as the ones most prone to danger that is they are regarded as malicious observers. The service provider could be the adversary or an adversary could hack into the servers. In both the cases described, the adversary obtains access to the location information stored in the servers. Sensitive user data such as IP addresses and the location data embedded inside each query could be manipulated and misused by the adversary. Recent techniques such as object tracking algorithms can be used to map anonymous queries to the same user thus breaching privacy.

The assumption made at this point is that the threat model deals only with privacy in regards to location information contained in the queries. It doesn’t take into account threats posed by user smart phones or non-secure communication networks. Various

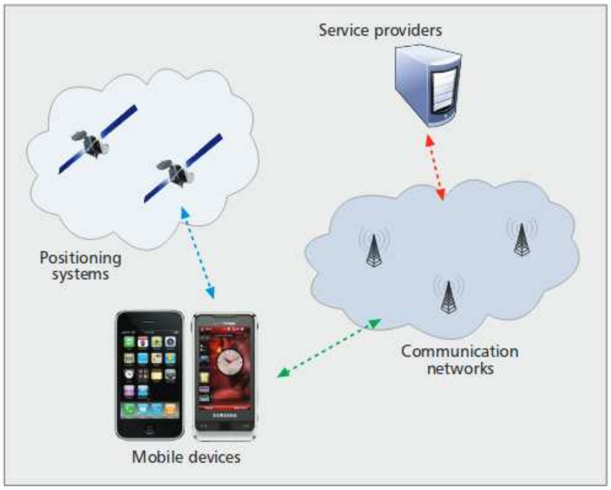


Fig. 1 Location-Based Service Architecture

researchers have spent considerable time in coming up with techniques to prevent data leakage and privacy risks associated with Location-Based Services. This paper aims at providing a comprehensive review of the state-of-art proposed that help preserve user identity.

4 Location Privacy Protection Strategies

Several researchers have put forward various theories for protecting the location privacy of a LBS recipient. The basic motive behind the working of the techniques is to prevent unnecessary breach of information and to somehow (explicitly or implicitly) achieve a level of control over what information is given to whom and when [7,8]. A trade-off exists between the utility of the Location-Based Service and the privacy of the location information that they are ready to share with the service providers. The following sub-sections elaborate in detail the different strategies that can be used for protecting user privacy.

4.0.1 Cryptography-Based Techniques

Cryptography can come to the rescue when required to offer privacy and security to user information. In [9], the authors have proposed 2 cryptography based protocols in order to achieve the required information secrecy and accuracy. The protocols make use of certain regions known as the Minimal Uncertainty Region (MUR). According to the protocols, it is perfectly acceptable if the adversary knows the MUR X in which a user is present as long as no information regarding the position of the user within X is disclosed. In order to obtain the minimal uncertainty regions, the authors have made use of spatial granularity, which is a division of the spatial domain into a discrete number of non-overlapping regions, called granules. The protocols under discussion are C-Hide Seek and C-Hide Hash, which adopt symmetric encryption techniques where each user makes use of a unique key that is known only to his/her friends and also the other way round. The exchange of keys takes place through a secure communication channel before executing the protocols.

4.0.2 Private Information Retrieval

Private Information Retrieval is a popular technique in preserving privacy of the users based on nearest neighbour searches. It had formally been defined as ”Techniques that enable a user to access k replicated copies of a database (k¿2) and privately retrieve information stored in the database. This implies that each individual database gets no information on the identity of the item retrieved by the user” in [10]. This technique was mainly used for multiple databases and later based on this approach a computational PIR (cPIR) technique for single databases has been proposed in [11] and [12]. PIR techniques require additional computational and processing support on the server side of the Location-Based Service.

4.0.3 Location Obfuscation

Location Obfuscation is the process of degrading the quality of user location information to a certain extent, so as to protect the individuals location privacy. Location Obfuscation is the process through which the user location information is slightly altered or substituted or the location itself is generalized so as to avoid sharing the actual user location co-ordinates. Few of the popular techniques to perform obfuscation are pseudonyms,adding random noise, spatial cloaking, Redefinition of possible areas of location etc.

4.0.4 Pseudonyms

If Pseudonyms, are implemented and made use of properly, they can be an effective and easy way of offering location-based privacy. Pseudonyms have been made use of in [13] where secure authorization and access control had to be achieved. However, Pseudonyms alone don’t suffice in providing privacy in location-based applications since the pseudonyms that stay the same without changing as the time proceeds will eventually help in adversary in figuring out the locarion of the user [14]. Later techniques have made use of a concept called mixed zones [15]. This proposal can be mainly used for applications that cannot be used anonymously, but do not require the true details of the user as well. Thus an internal pseudonym can manage the service when third party applications provide users with pseudonyms and ensure that their identities would not be revealed to the LBS.

4.0.5 Spatial K-anonymity Paradigm

K-anonymity was one of the preliminary approaches proposed to achieve user location privacy, the basic idea of the approach is to make a particular user indistinguishable among other K-1 users [16, 17]. In the Spatial K-anonymity approach the user sends his query along with the location information to the middleware which constructs a spatial region (cloaking region) which contains the location of the user who queried the LBS along with K-1 other users. This spatial region and the user query are then sent to the Location-Based Service provider. LBS executes the query by taking the location of the spatial region into consideration and sends a set of results to the middleware. The anonymizer then filters out the results by eliminating false positives. This technique known as Spatial Cloaking has attracted considerable attention from the research community.

An improvement to the approach could be to replace rectangular cloaking regions with regions influenced by voronoi diagrams [18]. The improves security, performance and flexibility. [19] has modified the concept of k-anonymity by using fuzzy context parameters. The concept of k-anonymity has been extended by various researchers in order to obtain performance gain. The prominent ones are l-diversity, t-closeness, p-sensitivity, and historical k-anonymity. Authors in [20] have proposed a technique that provides privacy in P2P(Peer-to-Peer) networks, where members of the network exchange location data between them and calculate the cloaked regions which would be used by the LBS provider. This eradicates the need for a trusted third party middleware.

This approach is not the best one out there due to certain issues like Regions that have many users are smaller in size thus resulting in smaller cloaked areas. Another factor that turns out to be a hindrance is - cloaked anonymity approach often incurs the outlier problem as stated in [21]. Also, on the server side this approach will incur a significant computational overhead as processing a user query with a spatial region as input would require more spatial operations than a user query with a specified point location.

4.0.6 Introducing Noise for location obfuscation

Introducing noise to the user locations results in losing out on the original location permanently and obtaining an obfuscated location that is close enough to be given as an input to LBS such that decent enough performance could be achieved. The induced noise however, cannot be too large. The advantage of this method is that no additional computational overhead is required at the server side. Authors in [22] have proposed a multiple point-based noise induction technique to generate an imprecise obfuscated location. N random points are generated with a particular radius with the center as the original location and the farthest point is chosen as the obfuscated location. An Extension of this approach is 0-Rand [23] which generates random points with a radius based on a defined angle 0 and the original circle radius. The farthest among points thus obtained is chosen as the obfuscated location.

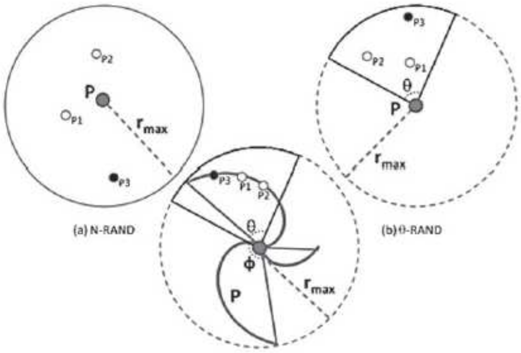


Fig. 2 Example of generation of random point in N-Rand and 0-Rand

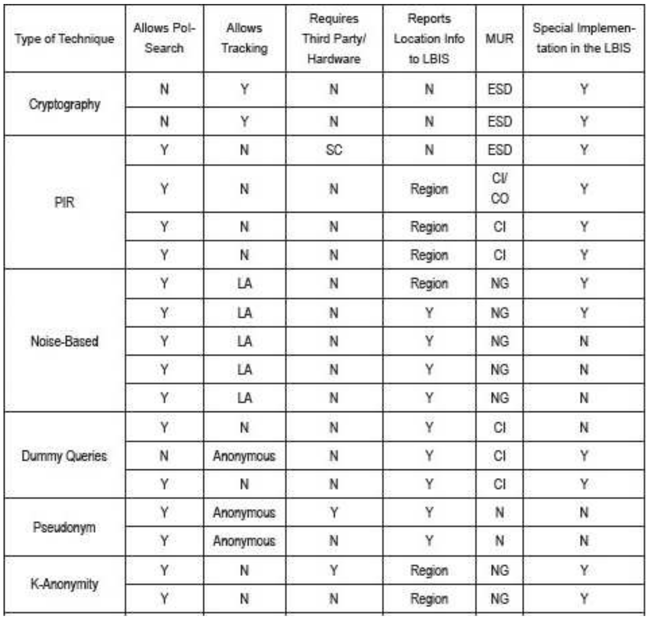
4.0.7 Dummy Queries

Usage of dummy queries is another approach to prevent privacy breach of the user. In the approach, N fake(dummy) queries are sent along with the real user query so as to disguise and hide the user's true geo-location [24, 25]. The obfuscation region thus consists of a set of discrete locations to be sent to the LBS for query execution. The drawback of this method is the unnecessary additional cost incurred by processing N queries for one relevant query.

A comparative study of the location privacy protection strategies discussed above has been outlined in the tabular form presented as Fig. 3.

5 Query Privacy Protection Strategies

The paper[26] discusses novel approaches to preserve the privacy of service attributes such as (bar, hospital, church, temple etc) when a user continuously queries the Location-Based Service as his/her position keeps changing. This paper is the first of its kind in initiating an investigation on privacy preservation for continuous queries to the Location-Based Services server. DUMMY-Q, a dummy query generation scheme which takes into account the users motion model and the query context



Type of Technique	Allows PoI-Search	Allows Tracking	Requires Third Party/ Hardware	Reports Location Info to LBIS	MUR	Special Implementation in the LBIS
Cryptography	N	Y	N	N	ESD	Y
	N	Y	N	N	ESD	Y
PIR	Y	N	SC	N	ESD	Y
	Y	N	N	Region	CI/ CO	Y
	Y	N	N	Region	CI	Y
	Y	N	N	Region	CI	Y
Noise-Based	Y	LA	N	Region	NG	Y
	Y	LA	N	Y	NG	Y
	Y	LA	N	Y	NG	N
	Y	LA	N	Y	NG	N
	Y	LA	N	Y	NG	N
Dummy Queries	Y	N	N	Y	CI	N
	N	Anonymous	N	Y	CI	Y
	Y	N	N	Y	CI	Y
Pseudonym	Y	Anonymous	Y	Y	N	N
	Y	Anonymous	N	Y	N	N
K-Anonymity	Y	N	Y	Region	NG	Y
	Y	N	N	Region	NG	Y

Fig. 3 Comparison table of various location privacy protection strategies

information as identified from the historical query logs has been proposed. In order to reduce the storage and computational overhead incurred, a quad-tree based storage/retrieval scheme has been developed. Thus, the proposed technique is computationally efficient as well as communication-wise cost efficient. They make use of minimal storage without affecting the accuracy significantly. The most important note-worthy contribution is the elimination of a trusted third party as well as a pre-known user trajectory of movements.

6 Open Challenges and Future Work

Researchers are trying their best to come up with better techniques in order to achieve complete privacy and security of the service recipients of an LBS. This area of research still has certain open issues to be addressed. They are as follows :

1. **Semantics :** State-of-art techniques do not consider the semantics of the user location and the user query. Taking the semantics into account could help in improving the techniques used for privacy preservation. There is a lot of scope for future research in this direction.

2. **Privacy during Data Collection :** The location data that needs to be sent to an LBS when a user seeks a particular service is first captured in the hand-held device of the user. The device manufacturers make use of fraud practices and collect user data which can be misused when shared with others. State-of-art techniques assume that data collection is done in a secure manner but it is clearly not that case in the practical scenario. Relevant work could be done in this direction.

3. **Ensuring Security in Communication Networks :** Earlier research approached do not consider security aspects of the networks that transmit the queries from the user to LBS and vice-versa. Thus it becomes increasingly important for researches to take into account the threats posed when security is compromised at the network level.

4. **Compatibility :** An issue that needs to be addressed is how to develop an approach that is compatible with the existing hardware infrastructure and popular Location-Based Services available in the industry. Future work should focus on the technicalities while integrating the proposed techniques with the services available. Any issues that creep up should be dealt with accordingly.

7 Conclusion

The threat posed as a result of the increasing number of Location-Based Services to user's privacy has been successful in attracting significant attention from the research community. As privacy breach is a serious concern that needs to be immediately looked into, the issue has drawn interest not only from academia but also from industry. The general Location-Based Service architecture model and the general threat model has been briefly discussed in the paper as an introduction. The paper then, elaborated upon the state-of-art techniques used in achieving location privacy and query privacy. The research work that has been accomplished in this regard has been elaborated and various extension to the proposed approaches their drawbacks and limitations have been clearly discussed. A comparative study of the location privacy protection strategies has been outlined in a tabular format. Despite all the measures taken to prevent user data from slipping into abusive hands, there are still several attacks that take place on a daily basis which pose serious threats to a user's privacy, identity and integrity. There are still several open challenges that need to be resolved. Directions to be focused on while carrying out future work have also been briefed upon.

References

1. Gruteser M, Grunwald D (2003) Anonymous usage of location-based services through spatial and temporal cloaking. In: Proc. 1st international conference on Mobile Systems, applications, and services (MobiSys). USENIX Association
2. Jagwani, Priti and Kaushik, Saroj, ”Privacy in Location Based Services: Protection Strategies, Attack Models and Open Challenges,” 2017
3. B. Hoh et al., Enhancing Security and Privacy in Traffic-Monitoring Systems, IEEE Pervasive Computing, vol. 5,no. 4, 2006, pp.3846.
4. J. Krumm, Inference Attacks on Location Tracks, PERVASIVE07, Proc. 5th Intl. Conf. Pervasive Computing, Springer-Verlag, 2007, pp. 30109.
5. Y. Matsuo et al., Inferring Long-Term User Properties Based on Users Location History, IJCAI 07 Proc. 20th Intl. Joint Conf. Artificial intelligence, Morgan Kaufmann Publishers Inc., 2007
6. Krumm, J.: Inference attacks on location tracks. In: LaMarca, A., Langheinrich, M., Truong, Khai, N. (eds.) Pervasive 2007. LNCS, vol. 4480, pp. 127143. Springer, Heidelberg (2007).doi:10.1007/978-3-540-72037-9_8
7. Liu, L.: Privacy and location anonymization in location-based services. SIGSPATIAL Spec. 1(2), 1522 (2009)
8. Liu, L.: From data privacy to location privacy. In: VLDB 2007, pp. 14291430 (2007)
9. D. Riboni, L. Pareschi, and C. Bettini, ”Privacy in georeferenced context-aware services: A survey”, in Privacy in Location-Based Applications. Springer, 2009, pp. 151-172.
10. B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, ”Private information retrieval”, Journal of the ACM, vol. 45, n 6, pp. 965-981, Nov. 1998.
11. A. Khoshgozaran, H. Shirani-Mehr, and C. Shahabi, ”SPIRAL: A Scalable Private Information Retrieval Approach to Location Privacy”, 2008, pp. 5562.
12. B. Chor and N. Gilboa, ”Computationally private information retrieval”, 1997, pp. 304-313.
13. Kaushik, S., Tiwari, S., Goplani, P.: Reducing dependency on middleware for pull based active services in LBS systems. In: Snac, P., Ott, M., Seneviratne, A. (eds.) ICWCA 2011. LNICSSITE, vol. 72, pp. 90106. Springer, Heidelberg (2012). doi:10.1007/978-3-642-29157-9_9
14. M. F. Mokbel, C.-Y. Chow, and W. G. Aref, ”The new casper: A privacy-aware location-based database server”, in Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on, 2007, pp. 1499-1500.
15. A. R. Beresford and F. Stajano, ”Mix zones: User privacy in location-aware services”, in Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second IEEE Annual Conference on, 2004, pp. 127-131.
16. B. Gedik and L. Liu, ”Location Privacy in Mobile Systems: A Personalized Anonymization Model”, in 25th IEEE International Conference on Distributed Computing Systems, 2005. ICDCS 2005. Proceedings, 2005, pp. 620-629.
17. Liu, L.: From data privacy to location privacy. In: VLDB 2007, pp. 14291430 (2007)
18. Khuong, V., Zheng, R.: Efficient algorithms for K-anonymous location privacy in participatory sensing. In: IEEE Infocom Proceedings 2012 (2012)
19. Jagwani, P., Kaushik, S.: Defending location privacy using zero knowledge proof concept in location based services. In: Proceedings of MDM 2012, Bangluru, India (2012)
20. Y. Che, Q. Yang, and X. Hong, ”A dual-active spatial cloaking algorithm for location privacy preserving in mobile peer-to-peer networks”, in Wireless Communications and Networking Conference (WCNC), 2012 IEEE, 2012, pp. 209-2102.
21. D. Quercia, I. Leontiadis, L. McNamara, C. Mascolo, and J. Crowcroft, ”Spotme if you can: Randomized responses for location obfuscation on mobile phones”, in Distributed Computing Systems (ICDCS), 2011 31st International Conference on, 2011, pp. 363-372.
22. S. Lederer, A. K. Dey, and J. Mankoff, ”Everyday privacy in ubiquitous computing environments”, in Ubicomp 2002 Workshop on Socially-informed Design of Privacy-enhancing Solutions in Ubiquitous Computing, 2002.

23. P. Wightman, M. Zurbaran, E. Zurek, A. Salazar, D. Jabba, and M. Jimeno, ” 0-Rand: Random Noise-based Location Obfuscation Based on Circle Sectors”, in IEEE International Symposium on Industrial Electronics and Applications (ISIEA) on, 2013.
24. You, T.H., Peng, W.-C., Lee, W.C.: Protecting moving trajectories with dummies. In: Proceedings of PALMS 2007 (2007)
25. Duckham, M., Kulik, L.: A formal model of obfuscation and negotiation for location privacy.In: PERVASIVE 2005 (2005)