Privacy Preserving in LBS resilient to Location Injection Attacks

M. Anuradha, R. Srikanth, Y. Rama Devi

Abstract: Location-based services have become indispensable in people's life with expeditious development of technology. Location-based services(LBS) refers to the services provided by the LBS servers with regards to area and point of interest. Alternatively, the LBS means getting the right information at the right place in time. Protecting user location privacy is the most challenging factor in LBS. This survey aims to present various mechanisms in preserving the user's location privacy and proposes a mechanism for preserving the privacy of user location and query against the location injection attacks. We will be discussing credibility based k- anonymity mechanism for preserving the location of the user and homomorphic encryption for preserving the query of the user resilient location injection attacks in this paper.

I. INTRODUCTION

With expeditious development of technology, users can get services on their fingertips. Among those services, location-based service is one which is extensively used. Location-based services provide additional value to a user of the device by integrating the device position or location with other information.

LBS use the geographic location of devices like smartphones, personal digital assistant (PDA), or navigation device to provide a series of services. Examples of LBS include finding restaurants, theaters, shopping malls, etc near your location. With the help of these LBS apps, users can send the queries which include locations, identities, point of interests, and other information to the LBS server. In return, users take the pleasure of the benefits provided by LBS such as searching for the Points of Interests (POI).

In most of the privacy preservation work, the probability of LBS Server becoming adversary is high, as it may sell user's the private data to the third party legitimately for maximizing its profits or it may act as an adversary to do business with privacy-based services. The remaining of the paper is organized as follows, general system model for preserving the privacy in location-based services and key attributes of LBS are presented in section 2, Section 3 discusses different types of attacks which leak the privacy of user and mechanism that protects the privacy of the user in LBS respectively. In Section 4 different Privacy Preserving Approaches in Location-based Services are discussed. In section 5 important Privacy Metrics of LBS are discussed. Section 6 discusses one of the important attacks of LBS i.e Location Injection attack which is discussed followed by preserving the privacy of the user and query content.

II. LBS SYSTEM MODEL

The conventional system model for preserving the privacy in location-based services is composed of three key entities - Users, Anonymizer and LBS server as shown in fig1

Users: The users send LBS query through their mobile device, the query is the point of interest(POI) in the nearby area, like requesting for restaurants, shopping malls, clinics, theaters in nearby locations of the user.

Anonymizer: Anonymizer is a third party server. Anonymizer entity is optional. Generally, the anonymizer receives the query from the user. The query is sent to the LBS server by anonymizing the user identity.

LBS server: It serves the user request or provides the information. The server receives LBS query either from users or anonymizer.

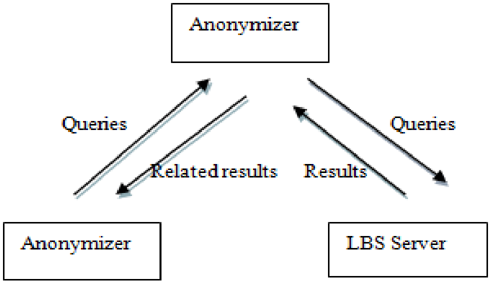

Fig 1:LBS System Model

Figure axis labels are often a source of confusion. Use words rather than symbols. As an example, write the quantity “Magnetization,” or “Magnetization *M*,” not just “*M*.” Put units in parentheses. Do not label axes only with units. As in Fig. 1, for example, write “Magnetization (A/m)” or “Magnetization ($\text{A}\cdot\text{m}^{-1}$),” not just “A/m.” Do not label axes with a ratio of quantities and units. For example, write “Temperature (K),” not “Temperature/K.” Multipliers can be especially confusing. Write **1 key attributes in LBS privacy**

When a user requests for the LBS, user identity, location or spatial information, temporal information and query contents are sent to the LBS server.

2.1.1 User Identity: It can be a user name or combination of properties which uniquely identify the user. User Identity is the sensitive information, and if the attacker gets access to the identity of the user, the adversary may deceive the user for personal, financial benefits.

Revised Version Manuscript Received on 16 September, 2019.

* Correspondence Author

M.Anuradha *, Pursuing M.Tech, Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana, anuradha.mudhigonda@gmail.com

Srikanth R, Assistant Professor, Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad, Telangana.

Dr. Y Rama Devi, Department of Computer Science and Engineering,

2.1.2 Location or Spatial information: Spatial information is the particular place of the user from where he/she is requesting the location-based services or locations of point of interest the user searches for. Spatial information of user needs to be preserved when requesting for LBS.

2.1.3 Temporal Information: It deals with time and spatial information. In some cases, the spatial information becomes critical if it is associated with the temporal information. The trajectory of the user can be estimated by the temporal information.

2.1.4.Query content: Query content refers to the type of Point of interest like a restaurant, shopping malls, universities ..etc. The query should be processed without the knowledge of the LBS server which prevents the leakage of privacy.

III. ATTACKS IN LBS

The adversary tries to get access to the private information of the user when he/she requests for LBS. There are different techniques and methods used by the attacker to disclose user information. Below are some of the attacks

3.1. Location Homogeneity Attack [3,23]

The location of all k- cluster members are analyzed by the attacker. When the users locations are almost identical, the location information of each member is disclosed. The location information is protected when the members are distributed over a larger area. An advanced location homogeneity attack utilizes map knowledge to reduce the effective area size where users can be found. As an example, all the k-cluster members' locations are limited in a restaurant or university landmarks.

3.2. Map Matching Attack [4]

To diminish the privacy of the user, the attacker uses external background information of the user. This attack is used to confine the obfuscation region to specific regions where the users are identified by expelling all the unrelated regions. For example, a map is used to shrink the size of the obfuscated region by eliminating the areas such as ponds from obfuscated region This way enables the attacker to define the user location.

3.3. Query Sampling Attack [5,22,24]

The attacker employs the unfair location distribution of the LBS users for his/her own malicious purpose. The adversary for his/her own vindictive cause utilizes the unfair location distribution of the LBS users. This type of inference attack targets isolated users in a sparse region. Isolated users in a sparse area are targeted by this type of inference attack. Therefore, it relies on the gathered traffic statistics of the environment where the users are located. The attacker tries to calculate a probability distribution function over a given area of the user location. The attacker can determine the areas where the user is located with a high probability when the probability is not uniformly distributed.

3.4. Semantic Location Attack[18,25]

The attacker can infer semantic meanings related to the behavior of the user by exploiting the amount of time a user spends in a location, for example, a hospital, restaurant, or a cafe. Consequently, the key idea is estimating the probability of the stay duration or usage time in a frequently visited Point of Interest by the user.

3.5. Location Injection Attack [1]

In location injection attacks, the untrusted users inject fake locations along with fake queries to the anonymizer, the probability of divulging user's location privacy could be great, yielding a much higher risk of privacy leakage. In location injection attacks, the possibility of divulging user's location privacy could be more as the untrusted users infuse fake locations with fake queries to the anonymizer. Mostly this type of attack occurs in continuous location-based services. For example, continuously report the restaurants near me. The attacker launches the location injection attack by getting access to the past locations of the user and attempts to manipulate the fake locations and by sending the fake locations along with the other users.

IV. LOCATION-BASED SERVICES PRIVACY PRESERVING APPROACHES

4.1. Spatial obfuscation [7]

This technique protects the privacy of user where it reduces the accuracy of spatial information sent by the LBS user to LBS server. It is an effective method used for the semantic attacks, where the probability of mapping the location of the user to the critical locations like hospitals, clinics is very high. In this obfuscation, the area is increased adaptively so that the possibility of the user in a particular semantic location is lower than the given threshold value.The location privacy is achieved without a trusted LBS server. It cannot secure the identity of the user.

4.2. Mix Zones

Mix zones proposed by Beresford et al. in [26], In this method the grouping of spatial regions where users are located in an area. The reallocation of the user is hidden by these regions, these regions will be mixed together. No updations are done to the location inside the mixed zone during the motion of the object. To ensure Mix zone approach more robust MobiMix [09] is developed. MobiMix approach considers the different context information which is used by the adversary to get detailed trajectories such as temporal and geometrical constraints.

4.3. Dummies

The key idea of Dummies[10] approach is to create false positions (dummies). In this method, the query is framed using the dummies and the true location for the same point of interest so that the adversary can be misleading. Dummy-Q [11]is another approach where the query context privacy is preserved. To hide the real query the dummy queries of different attributes form the same location. Easily integrable with the existing system.

4.4. Cache-based approach

The key idea in Cache based approach[12] is to avoid the dealing with LBS server by exploiting the collaboration among the LBS users. The result of the LBS query is stored in the device cache and passed to the other users seeking that information. The privacy of the user is protected from the server until the collaborative peers lack the sought information.

4.5. Cryptography

Cryptographic approach[13,27] uses encryption to protect user location. Mascetti et al., propose an approach to notify users, without revealing the current location to the LBS server, when the friends are there within the proximity. In this, the assumption is that the user shares the secret with each of his/her friends and use symmetric encryption techniques.

4.6. Position Sharing

Position sharing approach divides the obfuscated location data into position shares, and share is restricted to strict precision. The shares are distributed among a set of LBS servers such that each LBS server only has a position of limited precision. To prevent an adversary from increasing the precision of locations this work is extended by considering map knowledge[14].

4.7. K-Anonymity

K- Anonymity[6] introduced by Gruteser and Grunwald is an effective approach where the location perturbation of a certain user is achieved with other k – 1 users. The key concept of this k-anonymity is that the user sends his perturbed area instead of the actual location. K-anonymity is a commonly used approach for protecting the user location privacy, this approach uses Anonymizer(trusted third party). In this mechanism the user(say u1) sends LBS query q1 to anonymizer, the anonymizer aggregates the queries of (k - 1) users around u1, and u1's query qu1 within a so-called cloaked region (CR). This cloaked query q_{ui} (i ∈ (1, 2, 3)) is sent to the LBS server. The user's location cannot be distinguished from (k−1) locations of other users by the LBS server.

Method	Spatial Info	Temporal Info	User Identity	Anonymizer	location privacy	Query Privacy
Spatial obfuscation [7]	✓	X	X	Not Required	✓	X
Mix Zones[26]	✓	✓	✓	Required	✓	X
Dummies[10]	✓	X	X	Not Required	✓	✓
Cache based[12]	✓	✓	✓	Not Required	✓	X
Cryptography[13, 27]	✓	X	X	Not Required	✓	X
Position Sharing[14]	✓	X	X	Not required	✓	X
K-Anonymity[6]	✓	✓	✓	Required	✓	X

V. PRIVACY METRICS OF LOCATION

5.1. Location Entropy(LE)[15]**:** By quantifying the diversity of LBS user's location the privacy is measured.

LE = H(l) = $\sum_{u \epsilon Ul} Pl, u \, log Pl, u$

U_l : the set of distinct users that visited l , $p_{l,u}$, the fraction of visits to l that belongs to user u, $p_{l,u}$

5.2. K-anonymity: K-anonymity is also used as privacy metric of location privacy. The position of the user is indistinguishable among at least k-1 users' position, not the query issuer.

5.3. Expected distance error: [16] The Expected distance error measures how accurately an adversary can estimate a user's location, considering the differences among the locations observed by the adversary.

5.4. Incorrectness: [17] In this the distance is defined as the difference between the adversary's estimation and the true value related to the exact position of the user.

VI. THE VULNERABILITY OF USER'S PRIVACY BY LOCATION INJECTION ATTACKS

Location Injection attack is a concealed attack by an adversary. This attack can reveal the user's private information like the places user visited which reveals the user's behavior by knowing the places he /she visited and the time spent, the medical records of user can be inferred by knowledge of the which clinic the user visits. By this attack, the personal data of the user is at stake. Location and query privacy are two sorts of security issues in LBSs. Location privacy is about private information on the user's current position or trajectory within the duration. Query privacy is sensitive information about query content.

The pivot of this section is on an approach which is against location injection attacks in location-based services(LBS). The LBS server may disclose the user information in queries, for its own profits. Continuous LBS queries by the user are more vulnerable to location injection attacks. The user's location and query need to be protected.

The attacker first gets the information of the user's(u) cloaked region(CR) at time t_i to disclose the location of the user at t_{i+1}. By analyzing user CR and maximum moving speed obtained from the speed limit of the road[19] or statistical data[20], the attacker can infer the user's area at time $t_{i+1.}$ When the attacker knows that the user is in a particular region, consequently generates fake users in the user's region and injects fake locations.

6.1. Preserving the privacy of user Location

The most common and efficient mechanism used for preserving the privacy of the user is K-anonymity. The probability of discovering the user location is not more than 1/k as the trusted third-party anonymizer aggregates the queries of k nearby users within a cloaked region. The drawback of K-anonymity is that it assumes that all the users involved are trusted users. When there are location injection attacks the probability of privacy leakage of the user is more than 1/k. A credibility based[1] k - anonymity mechanism can be used for location privacy in case of location injection attacks.

Credibility based k - anonymity[1] considers location injection attacks on high-risk users. Based on cloak regions the user's mobility trajectories are modeled instead of real locations. Next, it computes the mobility similarity by correlating high-risk users mobility and fake users. Each user is assigned a credibility value based on the mobility similarity. The credibility based k-cloaking algorithm ensures the location privacy of all the users by cloaking the high-risk users with credible users and clocks other users scoring approximation credibility in the same cloak region.

In this mechanism, the probability distribution and Euclidean distance are used to find the mobility similarity. The probability distribution indicates the similarity in possibilities of visiting locations at a different time and attacker's motivation that they try to enable the fake locations cloaked with high-risk users is revealed by Euclidean. When LBS query is received from the high-risk user, the fake trajectories around the user incorporating location inference attack are filtered out[21] and these users filtered out are treated as fake users and assigned credibility zero. The privacy of the high-risk users is protected by cloaking the users who meet the credibility constraints with them.

6.2. Preserving the privacy of user query content

To preserve the privacy of query content, the results of the LBS queries need to be obtained with zero LBS server knowledge. This can be accomplished by using Homomorphic Encryption(HE) as stated by Shengling Wang, Qin Hu, Yunchuan Sun, and Jianhui Huang [2]. In homomorphic encryption, certain operations on ciphertext can obtain desired operations on plain text to achieve secure computing. In this method, the server obtains the obfuscated location, which is obtained by implementing the location preserving scheme. The server extracts local database which contains all the POIs say $\{P_1,P_2,P_3..,P_M\}$ in a particular area around the user submitted location, where Pi is set of POIs with the type of $a_{i.}$

The attribute vector in the server is represented as type $a_i = \{a_{i,1},a_{i,2},a_{i,3},....a_{i,j}\}$ for each POI, where $a_{i,j}$ is the jth attribute of a_i, such as "restaurant" ,"hospital", and "shopping malls". Thus, the local database contains two matrices, that is, the attribute matrix and the POI matrix. Each row of the former has a series of attributes denoting the POI type and the attribute matrix is an index of the POI matrix. Key pair generated by the server with above notation, the encrypted Euclidean distance between the user's POI type $a_i = \{a_{i1}, a_{i2}, …,a_{iJ}\}$ and $a_{i..}$

In the next stage, OT protocol is used (OT is a protocol in which a sender delivers a piece of information to a receiver while resting oblivious to what piece has been transmitted), which facilitates to search the exact match of the requested POI type of the user, cryptographic computation between two mistrusting parts.

To secretly obtain the POI set with the known i* from the previous stage, computational private information retrieval (cPIR) based on quadratic residuosity assumption (QRA) is used. Thus the user query results are retrieved with zero server knowledge. Thus the user location and query content are protected against the location injection attacks by credibility based k-anonymity and homomorphic encryption.

VII. RESULTS

We have validated the proposed mechanism by extensive simulations on real world dataset loc-Gowalla[28]. For query encryption we have used the partial homomorphic encryption RSA method. The data set loc-Gowalla collects check-ins from Feb. 2009 to Oct. 2010 , 6.5 million. It consists of 196,591 nodes and 950,327 edges. The K value is taken in the range of [4,12]. We got the mobility similarity between high-risk user and trusted user is not less than 0.009. The mobility similarity between

high-risk user and the attacker is much similar. The credibility of the high risk user is considered as one and the credibility of user is decreased as the mobility similarity between high risk user and the user increases. The credibility of the attacker is considered as zero. This approach focuses more on cloaking credible users with high risk users to protect the privacy of the high risk users. For random 1000 attackers the proposed method is able to detect 838 attackers. The proposed method fails when there are no previous cloak regions stored in the database and suffers from the location injection attacks when the difference between the is the total number of users in the Cloak region of user u, and the number of fake locations in that Cloak region, is less than the k users.

VIII. CONCLUSION

Preserving the privacy of the user is a serious issue in the Location-based services. This paper presented different kinds of attacks and mechanisms that are presently being used to deal with this issue of privacy. Metrics and the attributes of location privacy which needs to be protected are mentioned. A mobile user is open to Location Injection Attack, in this paper we have discussed a system to preserve the privacy of user location and query content against location injection attacks. The implementation fully homomorphic encryption of the proposed system can be considered as future work.

REFERENCES

[1] Ping Zhao, Jie Li, FanziZeng, Fu Xiao , Chen Wang , Member, IEEE,andHongbo Jiang , "ILLIA: Enabling k-Anonymity-Based Privacy Preserving Against Location Injection Attacks in Continuous LBS Queries", IEEE INTERNET OF THINGS JOURNAL, VOL. 5, NO. 2, APRIL 2018

[2] Shengling Wang, Qin Hu, Yunchuan Sun, and Jianhui Huang, "Privacy Preservation in Location-Based Services", IEEE Communications Magazine ,March 2018

[3] Pan, Xiao, et al. "Protecting personalizedprivacy against sensitivity homogeneity attacks over road networks in mobile services."Frontiers of Computer Science 10.2 (2016):370-386.

[4] Krumm J,"Inference attacks on location tracks", In: Proceedings of the 5th international conference on pervasive computing (Pervasive07). Springer, Toronto, pp 127–143(2007).

[5] Saravanan, Shanthi, and Balasundaram Sadhu Ramakrishnan. "Preserving privacy in the context of location based services through location hider in mobile-tourism." ,Information Technology & Tourism 16.2 (2016): 229-248.

[6] Gruteser, M., & Grunwald, D. (2003). Anonymous usage of location-based services through spatial andtemporal cloaking. Proceedings of the 1st International Conference on Mobile Systems, Applications and Services (pp. 31–42).

[7] Ardagna C, Cremonini M, Damiani E, DeCapitani di Vimercati S, Samarati P ,"Location privacy protection through obfuscation-based techniques". In: Proceedingsof the 21st annual IFIP WG 11.3 working conference on data and applications security, Redondo Beach, CA, USA, pp 47–60(2007).

[8] Duckham M, Kulik L ,"A formal model of obfuscation and negotiation for location privacy", In: Proceedings of the third international conference on pervasive computing (Pervasive '05), Munich, Germany, pp 152–170(2005).

[9] Palanisamy B, Liu L, "Mobimix: protecting location privacy with mix-zones over road networks", In: Proceedings of the 27th IEEE international conference on data engineering (ICDE '11), pp 494–505(2011).

[10] H. Kido, Y. Yanagisawa, and T. Satoh, "An Anonymous Communication Technique Using Dummies for Location- based Services", IEEE Proc. Int'l. Conf. Pervasive Services, ICPS '05, July 2005.

[11] A. Pingley, et,al., "Protection of Query Privacy for Continuous Location Based Services", IEEE INFOCOM'11, Apr. 2011.

[12] Shokri, Reza, et al. "Hiding in the mobile crowd: Location privacy through collaboration." Dependable and Secure Computing, IEEE Transactions on 11.3 (2014): 266-279.

[13] Mascetti S, Freni D, Bettini C, Wang XS, Jajodia S, "Privacy in geo-social networks: proximity notification with untrusted service providers and curious buddies", VLDB J 20(4):541–566.

[14] Skvortsov, P, Durr, F., &Rothermel, K., "Map-aware position sharing for location privacy in nontrusted systems", Pervasive Computing, 388–405, (2012)..

[15] Voulodimos, Athanasios S., and Charalampos Z. Patrikakis, "Quantifying privacy in terms of entropy for context aware services", Identity in the Information Society 2.2 (2009): 155-169.

[16] Hoh, Baik, and Marco Gruteser, "Protecting location privacy through path confusion", First International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05). IEEE, 2005

[17] Humbert, Mathias, et al. "Reconciling utility with privacy in genomics", Proceedings of the 13th Workshop on Privacy in the Electronic Society. ACM, 2014.

[18] Li, Yanhui, et al. ,"Semantic-Aware Location Privacy Preservation on Road Networks", International Conference on Database Systems for Advanced Applications. Springer International Publishing, 2016.

[19] G. Ghinita, M. L. Damiani, C. Silvestri, and E. Bertino, "Preventing velocity-based linkage attacks in location-aware applications", in Proc. ACM SIGSPATIAL GIS, Seattle, WA, USA, 2009, pp. 246–255.

[20] G. Ghinita, M. L. Damiani, C. Silvestri, and E. Bertino, "Protecting against velocity-based, proximity-based, and external event attacks in location-centric social networks", ACM Trans. Spatial Algorithms Syst., vol. 2, no. 2, pp. 1–36, 2016.

[21] R. Shokri, G. Theodorakopoulos, G. Danezis, J.-P. Hubaux, and J.-Y. Le Boudec, "Quantifying location privacy: The case of sporadic location exposure", in Proc. Int. Symp. Privacy Enhanc. Technol. Symp., Waterloo, ON, Canada, 2011, pp. 57–76.

[22] Shokri R, Theodorakopoulos G, Le Boudec J, Hubaux J (2011) Quantifying location privacy. In: Proceedings of the 31st IEEE symposium on security and privacy (SP '11), Berleley/Oakland, California, USA, pp 247–262.

[23] Machanavajjhala, Ashwin, et al. "L-diversity: Privacy beyond k-anonymity." ACM Transactions on Knowledge Discovery from Data (TKDD) 1.1 (2007): 3.

[24] Lin, Chi, Guowei Wu, and Chang Wu Yu. "Protecting location privacy and query privacy: a combined clustering approach." Concurrency and Computation: Practice and Experience 27.12 (2015): 3021-3043.

[25] Lee, Byoungyoung, et al. "Protecting location privacy using location semantics." Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011.

[26] Beresford AR, Stajano F (2004) Mix zones: user privacy in location-aware services. In: Proceedings of the second IEEE annual conference on pervasive computing and communications workshops (PerCom '04 Workshops), pp 127–131.

[27] Marias G, Delakouridis C, Kazatzopoulos L, Georgiadis P (2005) Location privacy through secret sharing techniques. In: Proceedings of the 1st international IEEE WoWMoM workshop on trust, security and privacy forubiquitous computing (WOWMOM '05), pp 614–620.

[28] SNAP Datasets:Stanford Large Network Dataset Collection.[Online].Available:http://snap.stanford.edu/data

M.Anuradha, currently pursuing M.Tech at Chaitanya Bharathi Institute of Technology in the Department of Computer Science and Engineering, she has 6 years of Industry experience in various WebTechnologies. Her current research interests include Mobile ,Privacy Preserving, Machine Learning, Blockchain Technology.

AUTHORS PROFILE

M.Anuradha, currently pursuing M.Tech at Chaitanya Bharathi Institute of Technology in the Department of Computer Science and Engineering, she has 6 years of Industry experience in various WebTechnologies. Her current research interests include Mobile ,Privacy Preserving, Machine Learning, Blockchain Technology.

Srikanth R is Assistant Professor in the Department of Computer Science and Engineering at Chaitanya Bharathi Institute of Technology. He has taught subjects like Machine Learning, Streaming Technologies and Open Source software courses. He did his master from National Institute of Technology,Hamirpur(H.P) and currently pursuing his Ph.D from Osmania University. His research interests include Spatial Informatics, Security, User Privacy, Social Networking and Machine Learning.

Dr. Y Rama Devi has done her Ph.D. from Hyderabad Central University. Her area of specialization are Data Mining, Bio-Informatics, Artificial Intelligence, Algorithms, Soft Computing, Machine learnign She has published around 100 papers in International Journals/Conferences. She is recipient of Best Teacher Award and Distinguish women in Engineering. She has about 28 years of teaching experience and 15 years of research Experience. She has successfully guided five Ph.D. She has one patent. She is member of CSI,IEEE, IETE,ISTE . She has delivered keynote lecture in various conferences.