A Framework for detecting and avoiding location based queries to preserve content and user privacy in databases: A Review

[1]Ms Prutha Suresh Devikar
M.E. (IT) III Semester
Prof. Ram Meghe Institute of Technology and Research, Badnera.

[2]Dr. C. A. Dhote
Professor, Department of IT
Prof. Ram Meghe Institute of Technology and Research, Badnera.

***Abstract*:** Location based services (LBS) is a utility services accessible by various devices and are part that virtually controls systems which work in computer. Users are willingly performing queries from various places through different devices to learn any location. This searches by user are likely served through LBS. Protection is an important issue in this concern. A mechanism to protect users' location and query privacy various methods are experimented. These techniques are applied in order to detect and avoid the unwanted circumstances and violation. Probably, various methods minimize violating privacy of users in its different ways.

Keywords: databases, location based queries, privacy preserving.

1. Introduction

Location based queries are provided by location based service (LBS). These are generally based on a point of interest (POIs). By retrieving the Points Of Interest from the database server, user probably get answers to various location based queries, which are for example discovering the nearest hospital, ATM machine or police station, restaurant. In years there has been increase in the number of devices querying location servers for information about POIs. Queries are thus use for obtain required information from database [1].

Among many challenging barriers to the wide classification of such application, privacy assurance is a major issue. Say, users want to perform many queries from home. So users may feel unwilling to disclose their locations to the LBS, because it may be possible that a location server learns location of that user who is making a certain query. Detection thus implies location by context of queries. The more effectively detecting the locations where search queries are performed have made huge influence on increasing relevance. Probably, various methods must implement to avoiding such queries that are violating privacy of users. There have to be proper mechanism to tackle this problem by using encryption schemes to retrieve data without violating the privacy. The approach is to look upon LBSs have attracted the research and development community.

The information provided by LBS as a result of query can be manipulated by server itself. The problem arises when malicious attacker traces users' information. To protect it scheme must implemented, just authorized access should allowed. All this is used in techniques for privacy preserving giving different types of methods. The approach is to look upon unaware transfer and also retrieval of any Private Information in process of querying, in order to achieve a secure solution to user and server, both. Also improvement in time factor while querying to get the results even faster.

In this paper we have discussed about all techniques and approaches used in field of LB queries privacy. The paper is organised as follows. In section 1, we provide basic introduction and knowledge to the topic. In section 2, there are main purpose of privacy and challenges faced. In section 3, all different methods and working are being discussed. The final section ends with some conclusion of this discussion.

2. Objectives and Challenges

As per all terms taken in consideration working on quires, privacy is an important aspect. Main objective is protection of users' data while querying. This makes new thinking on LBSs providing result to query to reaching at proper user? So detecting this problem is essential to privacy policy.

ISSN: 2278 – 1323

However, LBSs suffer from a security pitfall in terms of violating users' privacy. There were various challenges faced during preserving privacy of user. As growth in internet applications, matter has become even worse in attacks of unauthorized access, resulting in violation of privacy. There raised number of approaches in order to protect data and provide privacy to user during querying. New methods introduced and more challenges were faced for improvement by reconsidering known privacy metrics. There all approaches tried solve problem of privacy protect in their own way. Methods including: Cloaking; Generation of dummies; Private information retrieval (PIR), etc.

3. Preserving Techniques & Methods

There was work proposed in year 1999 that enable a user to access k replicated copies of a database and privately retrieve information stored in the database [6]. This means that each individual server (holding a replicated copy of the database) gets no information on the identity of the item retrieved by the user. This schemes and similar works, proposed very earlier use the replication to gain full real saving with multi-server. In particular, it's presenting a two-server scheme with less communication complexity.

Then as work progress, there was application of the anonymity set technique to location data collected. The anonymity set measurements gives information that pseudonymity technique cannot give users adequate location privacy. However, positive thing made was if one is in the mix zone with 20 other people, it might consider better protected. But when one go in and out of the mix zone the observer will strongly suspect are those lonely pseudonym [3]. This motivated further work on it.

Anonymity had instinctively described as the property of being indistinguishable among a set of individuals. This provides privacy in more efficient way, safeguarding private information of the user. However, guaranteeing anonymous usage of location-based services requires transmission of location information by a user that cannot be easily used to re-identify the subject. The paper "Anonymous usage of location based services through Spatial and temporal cloaking" in year 2003 found middleware architecture and algorithms that can be used by a centralized location broker service. This adaptive algorithm adjusts the resolution of location information along spatial or temporal dimensions to meet specified anonymity constraints, which are based on the entities using location services within a given area [11]. The application of this technique in effective way, it requires location-based services that are used with precise position information by a large user base. It analyzed the technical feasibility of unknown usage of location-based services and properly studied the location data introduces new and potentially more severe privacy risks than in conventional services.

Anonymity and pseudonymity again are not a complete answer to privacy. Moreover, it presents a barrier to authentication and personalization and also vulnerable to data mining, since there is possibility of getting location hence revealing identity often.

Thus, there was the argument that obfuscation is complementary to that times existing privacy protection strategies and demands further investigation. The paper proposed on "A formal model of obfuscation and negotiation" [8]. In this Obfuscation is defined as the means degrading in carefully way the quality of information about an individual's location in order to protect that individual's location privacy. Key assumptions made by architecture are: A client device uses some combination of location-sensing techniques to provide accurate information about the client's location. That client device can be able to communicate with a third-party location based service provider (TPLBSP) via a wireless network. So some information service based on the client's current location is received. The information chosen to reveal about clients' location contains only source of information available to the TPLBSP about that location.

Obfuscation thus became an essential component of an overall approach to location privacy then. It provides a framework for the provision of high quality location-based services based on low quality location information. This model includes algorithms able to achieve a proper balance between privacy and location-based service utility properly.

Along with this, methods such as path confusion or using dummies were implements. Dummy variable thus help in hiding location thus providing users' privacy [14]. Further discussion on devices like computers and other devices are not

ISSN: 2278 – 1323

only used for requesting location. Moving devices like mobile also came in frame working on topic. Mobile devices equipped with positioning capabilities (e.g. GPS) can ask location-dependent queries to Location Based Services (LBS) [10]. To protect privacy, the user location must not be disclosed. Existing solutions before this has utilize a trusted anonymizer between the users and the LBS.

The approach has several drawbacks:

(i)All users must trust the third party anonymizer, which is a single point of attack.

(ii) A large number of cooperating, trustworthy users is needed.

(iii) Privacy is guaranteed only for a single snapshot of user locations; users are not protected against correlation attacks.

So again overcome above problem, new propose of novel framework to support private location dependent queries, based on the theoretical work on Private Information Retrieval (PIR) [9]. This framework does not require a trusted third party; instead it made use cryptographic techniques for privacy. Compared to existing work, this approach achieves stronger privacy for snapshots of user locations; moreover, provide provable privacy guarantees against correlation attacks. In this work implement involved approximate and exact algorithms for nearest-neighbor search. Also optimize query execution performed in by employing data mining techniques. The experimental results suggest that PIR approaches incur reasonable overhead and are applicable in practice. This work was first to provide a practical PIR implementation with optimizations that achieve communication and CPU cost as well as to protect against correlation attacks compared to previous work. In the future, plans to investigate the extension of this framework to different types of queries.

Also other contributions in privacy preserving by defining location-based quasi-identifiers and by introducing the notion of Historical k-anonymity [4], some of the research work provide a formal framework to evaluate the risk of revealing personal sensitive location information. It proposes a technique to preserve a specified level of anonymity, and identify several evolving research directions on this topic.

Existing privacy-enhancing techniques protect user identities. Nevertheless, the query contents may disclose the physical location of the user. Hence, new presentation of a framework for preventing location-based identity inference of users who issue spatial queries to location-based services [19]. It proposes transformations based on the well-established K-anonymity concept to compute exact answers for range and nearest-neighbor search, without revealing the query source. The methods optimize the entire process of anonymizing the requests and processing the transformed spatial queries.

There were again implementations of two other approaches in working on it. The first one, referred as Naive [2], assumes the location updates made a service user are independent to each other. For each location update, Naive just finds a cloaking box and reports it as the service user's location in her service request. The second approach is referred to as Plain. This scheme determines the cloaking set for the service users by finding the footprints closest to users' start position. After fixing the cloaking set, algorithm is applied to compute the cloaking boxes for the user to know during entire service session.

In middle there is theory that present a single-database private information retrieval (PIR) scheme with communication complexity O(k+d), where $k \geq \log n$ is a security parameter that depends on the database size n and d is the bit-length of the retrieved database block [20]. This communication complexity is better one than earlier single-database PIR schemes. The scheme also provided improved performance for practical parameter settings whether whatever is size of block.

Previously work focused on finding good trade-offs between privacy and performance of user protection techniques. The approaches to protect based on hiding locations inside cloaking regions (CRs) and encrypting location data via PIR protocols. Further, contribution of this work has the approach that proposed (i) a cryptographic protocol which allows private evaluation. They use this protocol as a building block in determining the nearest POI to a given user location. It could adapt to other types of spatial queries easily. (ii) Development of a hybrid approach that efficiently supports PIR processing with respect to a user-generated cloaked region Q. The proposed method controls CRs as well as disclosed POI information. Furthermore, it proved more efficient than its PIR-only [12].

The paper then proposed a hybrid technique for private location-based queries which provides protection for both the users and the service provider. It was then first work to

ISSN: 2278 – 1323

consider the protection of the POI database. Even further working carried on with protocol [13].

An alternative and complementary approach to spatial cloaking based location privacy protection is to break the continuity of location exposure by introducing techniques, such as mix-zones. The aim of the mix zone model was to prevent tracking of long-term user movements, but still permit the operation of many short-term location-aware applications. Mix-zones anonymize user identity by restricting the positions where users can be located [3]. Presenting on this a model MobiMix is framework for building mix-zones on road networks. It is used for protecting the location privacy of mobile clients [23].

One of the methods is new metrics to measure users' query privacy considering user profiles. It computes regions expressed in terms of metrics using design spatial generalisation algorithm. The extend k-anonymity that it propose new metrics to correctly measure users' query privacy in the context of LBSs [5], which enable users to specify their query privacy requirements in different ways. The main idea of k-anonymity is to guarantee that a database entry's identifier is indistinguishable from other k−1 entries. Further by knowing concept deeper, k-anonymity reveals its drawbacks in preserving. Based on the analysis, they conclude that cloaking (e.g., k-anonymity) is effective for protecting query privacy but not location privacy focus on protecting query privacy using cloaking with the assumption that the opponent learns users' real-time locations. The illustration with the features of different metrics gives a better protection than k-anonymity to users. We consider a powerful attacker who can obtain user profiles and has access to users' real-time positions in the context of LBSs.

The concept of answering location-related information for encrypted positions became better and promises to improve security needs. Indeed, such a mechanism can strongly attract the attention of researchers as it supports the preservation of the users' privacy. Further working developed a novel fully secure location-based mechanism based on a homomorphic encryption scheme [24]. It described the circuits that allow a LBS server to process encrypted inputs to retrieve targeted records that match the user's request.

This work model presents and describes protocol. It analyses the security performance and efficiency of the protocol with working using two platforms: a desktop and a mobile. The ultimate goal here was to obtain records from the LS and maintain privacy at users' and server. It is by applying an approach of oblivious transfer. It uses public grid to obtain record. The analysis is concern with the security of the client and the server both.

In recent year, paper proposed a novel protocol for location based queries that have major performance improvements with respect to the approach by Ghinita at el. [12] and [13]. This protocol is organized according to two stages. In the first stage, the user privately gets its location within a public grid, using oblivious transfer. This data contains both the ID and associated symmetric key for the block of data in the private grid. And then in second stage, the user executes a communicational efficient PIR, to retrieve the appropriate block in the private grid. This block is decrypted using the symmetric key obtained in the previous stage. This protocol thus provides protection in both ways. The user is protected because the server is unable to determine the location. Similarly, the server's data is protected since a malicious user can only decrypt the block of data obtained by PIR with the encryption key. The working implementation demonstrates the efficiency and practicality of the new approach.

4. Conclusion

In today's world, privacy has proved to be major concern. Sensitive information is preserve by people and there is always worry about not allowing it to be share in process of querying. This paper thus put forth survey on existing literature and techniques used in field of privacy for protection of data and other content. Working with privacy preserving, various different techniques used are studied in paper along with their pros and cons. All methods implemented new approach of working in order to satisfy objective is reviewed. The proper maintenance of privacy and the detection of the query that violate privacy is the aim to look upon in process of transfer and retrieval of data between user and server. Working on PIR and related work proved adaptive method among them. Based

ISSN: 2278 – 1323

on this future work could be done in efficient way and faster in much more real time. This could be contribution to the system further.

REFERENCES

[1] Russell Paulet, Md. Golam Kaosar, Xun Yi, and Elisa Bertino," Privacy-Preserving and Content-Protecting Location Based Queries", IEEE Transactions on knowledge and data engineering, VOL. 26, NO. 5, MAY 2014.

[2] T. Xu and Y. Cai, "Feeling-based location privacy protection for location-based services," in Proc. 16th ACM CCS, Chicago, IL, USA, 2009, pp. 348–357.

[3] A. Beresford and F. Stajano, "Location privacy in pervasive computing," IEEE Pervasive Comput., vol. 2, no. 1, pp. 46–55, Jan.–Mar. 2003.

[4] C. Bettini, X. Wang, and S. Jajodia, "Protecting privacy against location-based personal identification," in Proc. 2nd VDLB Int. Conf. SDM, W. Jonker and M. Petkovic, Eds., Trondheim, Norway, 2005, pp. 185–199, LNCS 3674.

[5] X. Chen and J. Pang, "Measuring query privacy in location-based services," in Proc. 2nd ACM CODASPY, San Antonio, TX, USA, 2012, pp. 49–60.

[6] B. Chor, E. Kushilevitz, O. Goldreich, and M. Sudan, "Private information retrieval," J. ACM, vol. 45, no. 6, pp. 965–981, 1998.

[7] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in Proc. EUROCRYPT, vol. 1592, Prague, Czech Republic, 1999, pp. 223–238.

[8] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in Proc. 3rd Int. Conf. Pervasive Comput., H. Gellersen, R. Want, and A. Schmidt, Eds., 2005, pp. 243–251, LNCS 3468.

[9] G. Ghinita, P. Kalnis, A. Khoshgozaran, C. Shahabi, and K.-L. Tan, "Private queries in location based services: Anonymizers are not necessary," in Proc. ACM SIGMOD, Vancouver, BC, Canada, 2008, pp. 121–132.

[10] B. Gedik and L. Liu, "Location privacy in mobile systems: A personalized anonymization model," in Proc. ICDCS, Columbus, OH, USA, 2005, pp. 620–629.

[11] M. Gruteser and D. Grunwald, "Anonymous usage of location based services through Spatial and temporal cloaking," in Proc. 1st Int. Conf. MobiSys, 2003, pp. 31–42.

[12] G. Ghinita, P. Kalnis, M. Kantarcioglu, and E. Bertino, "A hybrid technique for private location-based queries with database protection," in Proc. Adv. Spatial Temporal Databases, N. Mamoulis, T. Seidl, T. Pedersen, K. Torp, and I. Assent, Eds., Aalborg, Denmark, 2009, pp. 98–116, LNCS 5644.

[13] G. Ghinita, P. Kalnis, M. Kantarcioglu, and E. Bertino, "Approximate and exact hybrid algorithms for private nearestneighbor queries with database protection," GeoInformatica,
vol. 15, no. 14, pp. 1–28, 2010.

[14] H. Kido, Y. Yanagisawa, and T. Satoh, "An anonymous communication technique using dummies for location-based services," in Proc. Int. Conf. ICPS, 2005, pp. 88–97.

[15] J. Krumm, "A survey of computational location privacy," Pers. Ubiquitous Comput., vol. 13, no. 6, pp. 391–399, Aug. 2009.

[16] L. Marconi, R. Pietro, B. Crispo, and M. Conti, "Time warp: How time affects privacy in LBSs," in Proc. ICICS, Barcelona, Spain, 2010, pp. 325–339.

[17] T. Hashem and L. Kulik, "Safeguarding location privacy in wireless ad-hoc networks," in Proc. 9th Int. Conf. UbiComp, Innsbruck, Austria, 2007, pp. 372–390.

[18] B. Hoh and M. Gruteser, "Protecting location privacy through path confusion," in Proc. 1st Int. Conf. SecureComm, 2005, pp. 194–205.

[19] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias, "Preventing location-based identity inference in anonymous spatial queries," IEEE Trans. Knowl. Data Eng., vol. 19, no. 12, pp. 1719–1733, Dec. 2007.

[20] C. Gentry and Z. Ramzan, "Single-database private information retrieval with constant communication rate," in Proc. ICALP, L. Caires, G. Italiano, L. Monteiro, C. Palamidessi, and M. Yung, Eds., Lisbon, Portugal, 2005, pp. 803–815, LNCS 3580.

[21] M. F. Mokbel, C.-Y. Chow, and W. G. Aref, "The new casper: Query processing for location services without compromising privacy," in Proc. VLDB, Seoul, Korea, 2006, pp. 763–774.

[22] S. Mascetti and C. Bettini, "A comparison of spatial generalization algorithms for lbs privacy preservation," in Proc. Int. Mobile Data Manage., Mannheim, Germany, 2007, pp. 258–262.

[23] B. Palanisamy and L. Liu, "MobiMix: Protecting location privacy with mix-zones over road networks," in Proc. ICDE, Hannover, Germany, 2011, pp. 494–505.

[24] Youssef Gahi, Mouhcine Guennoun, Zouhair Guennoun1, Khalil El-Khatib "Privacy Preserving Scheme for Location-Based Services", Journal of Information Security, 2012, 3, 105-112

ISSN: 2278 – 1323