Research on LBS Privacy Protection Technology in Mobile Social Networks

Wen He
College of Humanities and Information Management, Chengdu Medical College
Chengdu, China
hewen41731@163.com

***Abstract*—With the rapid development of mobile Internet and locating technology, location-based services (LBS) are widely used in mobile social networks, while the user obtained the LBS services, their location information is also provided to the service provider, this information may be obtained by the attacker and the user's other personal privacy may be inferred, threatening to the user's personal information security. Based on the location privacy protection of LBS services, this paper analyzes the method and model of user's location privacy protection, and introduces some location privacy protection technologies and trajectory privacy protection technologies in LBS.**

Keywords-LBS; Location Privacy Protection; Trajectory Privacy Protection

I. INTRODUCTION

With the rapid development of mobile Internet and the popularity of smart mobile terminals, more and more people use smart phones, tablet PCs and other mobile devices to access the Internet for social activities. In the mobile social network, users can publish their information on the network anywhere, share and display their lives; they can easily communicate with other users via text, voice or video to enhance their friends' connections and expand their circle of friends. Mobile social networking has become a state of life [1]. In mobile social networks, location-based services (LBS) is a widely used application, the user can through the locating technology of smart terminals to know their location, access to relevant services, such as location and sharing, travel services, catering Entertainment queries, find nearby friends and so on. However, location services brings a lot of conveniences to the daily life of the people, it also increases the risk of users' privacy disclosure, because users need to provide their location to service providers to obtain services, so the location information in the process of sending may be intercepted by the attacker, resulting in the user's location or identity information disclosure. If the user initiates the query continuously in a period of time, and the attacker continuous monitoring, it may be inferred that the user's trajectory or the location of the next moment and other important information, to the user's personal and property security poses a serious threat. Therefore, how to effectively protect the user's location privacy has become an important research topic.

In the location services of mobile social network, the user's location privacy can be roughly divided into two categories: one is snapshot query location privacy, that is, the single-point location privacy when the user initiates the location query service at a certain moment. Another is the trajectory privacy when the user initiates the query continuously in a period of time. At present, the research of privacy protection in LBS is mainly aimed at the above two types of location privacy.

II. LOCATION PRIVACY PROTECTIONS IN LBS

Location privacy information is composed of the user's identity information and location information, the traditional location privacy protection method is mainly based on these two types of information to be classified [2]. One method is to hide the user's identity information (such as anonymity, pseudonym), but to provide accurate location information to the service provider, in order to get high-quality services; the other method is to provide the user's identity information to the server and hide the user of the location information, in order to achieve the purpose of protecting the location privacy. This method to reduce the attacker's probability of getting the user's actual location by generalizing or changing the actual location of the user. At present, in the location privacy protections of LBS, the location fuzzy technology is widely used, mainly has the false dummy, location k-anonymity, spatial-temporal cloaking and other technologies.

A. False Dummy

False dummy means that when a user initiates a location service query, a false location is used to instead of the user's actual location, so that the attacker can only obtain a false location information even if the attacker intercept the user's query information. As shown in Fig.1, suppose that there is a user Bill within the rectangular area A, the real location of which is shown as P1 in the Fig.1. When Bill initiates a location service query, the client sends a set of false locations (such as P2, P3, etc.) to the location server with a certain policy, and the server according to the received location service request to the background database query out all the candidate results (as shown in the solid small squares in Fig.1), and then by the user to filter. Because the user is using a false location in the query, the quality of service and the degree of location privacy protection will be affected by the distance between the true and false location, the closer the distance between the false location and the real location, the more accurate the query results, the better the quality of service, but at the same time the real location of the user is more easily exposed, the lower the degree of privacy protection; the contrary, the farther the distance between true and false location, the degree of location privacy protection is higher, but the quality of service is worse. Therefore, the key of false dummy technology processing is how to balance the

978-1-4673-8979-2/17/$31.00 ©2017 IEEE

contradiction between service quality and location privacy protection, so that the users can obtain higher quality service under the premise of less risk.

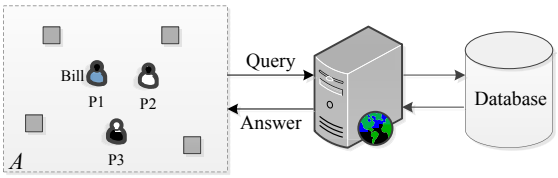


Fig.1. Sketch map of false dummy technology.

B. Location k-Anonymity

k-anonymity technology was first proposed by Sweeney in reference [3], at that time this method is mainly used in the relational database for generalizing the data which need to be published, so that the expression information of each released data and other k-1 data cannot be distinguished, so in a certain degree to protect the privacy of sensitive data. In reference [4], Gruteser first introduced k-anonymity into location privacy protection, and proposed the concept of location k-anonymity. The main idea is to generalize the user's actual location and then get a cloaked region which contains k-locations such that the location of the user cannot be distinguished from the location of other k-1 users, at this time the user satisfies location k-anonymity. Table 1 is an example of the location anonymity of k=4, suppose there are four users in an area whose IDs are U_1, U_2, U_3, U_4, and their locations are (1,3), (3,6), (2,5), (4,9), the location of each user is expanded to a fuzzy location range [(1,4)-(3,9)] after k-anonymity ,so when the user initiates the query, they sent to the server is a region containing four users, even if the attacker intercepted the query content, they cannot correspond with specific users, so to achieve the protection of the user's location privacy purpose.

TABLE I. LOCATION K-ANONYMITY EXAMPLE

User	Actual Location	Anonymous Location	Query
U_1	（1,3）	[(1,4)-(3,9)]	Q_1
U_2	（3,6）	[(1,4)-(3,9)]	Q_2
U_3	（2,5）	[(1,4)-(3,9)]	Q_3
U_4	（4,9）	[(1,4)-(3,9)]	Q_4

In general, the larger the value of k, the higher the user's location privacy protection degree, but the worse the quality of service, users can set the k-value according to their own privacy and quality of service requirements. At the same time, location k-anonymity technology will be affected by the population density of the region where the user is located. Suppose that the users in more densely populated areas, such as gymnasium, shopping malls, a small area can meet the demand of the larger k-value; and in the sparsely populated places, such as desert, grassland, the same k-value may return very large anonymous area, even may not meet the anonymous requirements, resulting in the failure of anonymity. Therefore, k- anonymous technology is largely dependent on the user's surrounding environment.

K-anonymity technology usually takes the trusted third-party architecture, between the client and LBS server set up a third-party, trusted server, called "anonymizer", the "anonymizer" is used to generalize the location of the user and to refine the candidate set of queries. As shown in Fig. 2, the location k-anonymity service process is as follows:

(1) The user sends his own actual location and query content to the anonymizer through encryption.

(2) The anonymizer generalizes the user's location into a k-anonymous spatial region and sends it to the LBS server. Then, the LBS server queries the back-end database for the fuzzy candidate sets to be sent back to the anonymizer.

(3) The anonymizer performs the refinement treatment on the fuzzy candidate sets according to the user's actual position and query content, and then sends the result of refinement to the user through encryption.

(4) The user client decrypts the received result and obtains the final service query information.

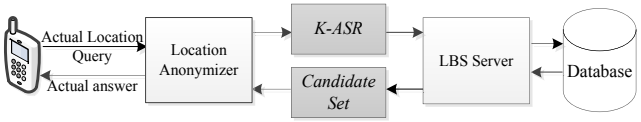


Fig. 2. k-Anonymity processing flow in trusted third party architecture.

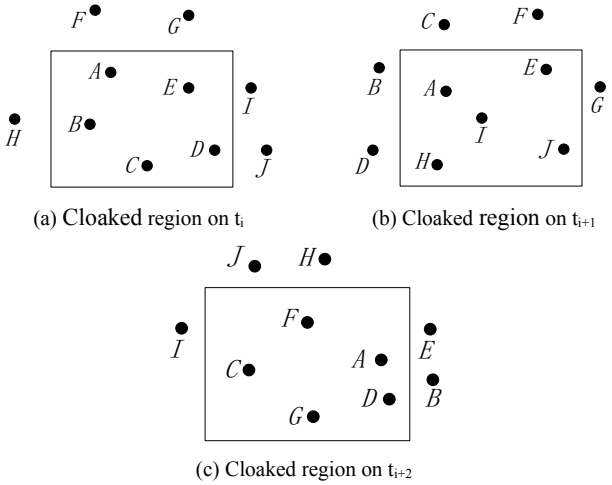
III. TRAJECTORY PRIVACY PROTECTION IN LBS

The k-anonymity algorithm has a better location privacy protection effect for the query initiated by the user at a certain moment, but if the user sends the location query continuously for a period of time, it is possible for the attacker to obtain its moving trajectory, for example, the attacker can take an intersection of anonymous sets of each query according to the time sequence, then the users who send messages can be inferred, and the trajectory of users can be obtained. If the attacker has a certain background knowledge to know the maximum movement speed of the user, they can according to the user's cloaked region at the previous moment to infer the region which user appears at the next moment, which threat to the users' personal and property security. So, how to protect the user's trajectory privacy is also an important part of location privacy protection. In general, the trajectory privacy attack model mainly has continuous query attack and maximum speed attack two.

A. Continuous Query Attack

When a user initiates a location service query at different times and places, if an attacker obtain the anonymous sets generated by the user at every times by attacking the anonymizer, and then take the intersection of these anonymous sets, it is easy to find the user who sent the message, this attack is called continuous query attack [5]. As shown in Fig.3, suppose that user A initiates the location query continuously at $\{t_i,t_{i+1},t_{i+2}\}$, the anonymity degree k = 5, then the anonymizer generates a cloaked region containing five users at time t_i. As shown in Fig.3(a), suppose the anonymous set contains five users {A, B, C, D, E}. At t_{i+1}, the location of all users changes, the users in the anonymous set becomes {A, E, H, I, J}, as shown in Fig.3(b). The attacker will seek the intersection of anonymous sets at time t_i and time t_{i+1} to get {A, E}. It can be

concluded that the continuous location query may be issued by A or E. At time t_{i+2},the user in anonymous set becomes{A,C,D,F,G},and then take the intersection with{A,E},get{A},it can be judged the continuous query is issued by the user A, at this point the user A's trajectory data is exposed.



(a) Cloaked region on t_i (b) Cloaked region on t_{i+1}

(c) Cloaked region on t_{i+2}

Fig. 3. Continuous query attack model

In this case, we can use the method of finding common anonymous set, so that the subsequent anonymous set always contains the anonymous set generated by the user in the initial query, as shown in Fig.3, the anonymous sets {A, B, C, D, E} generated at time t_i will be included in t_{i+1} and t_{i+2} moments so that the attacker cannot through the intersection operation of anonymous sets to find the user who sent the message. However, this method will make the anonymous area increased sharply with the user's location continuously update, to cause a great burden to the anonymizer. Reference [6] proposed a continuous query location anonymity algorithm based on location distribution probability. This method uses the distribution of users' historical location in the active area to find the k-1 users with the highest number of occurrences and the most densely distributed of location to form the common anonymous set, which can effectively reduce the area of the common anonymous region. However, the algorithm is based on the user's historical location distribution to find the common anonymous set, which is equivalent to predict the user's route, once the user deviated from the expected route, it will make anonymous area of the area surge.

Bereford and Stajano proposed a continuous query location privacy protection algorithm based on mix-zone [7], which defines the user access region as an application region and a mixed region. In the application region, the user can use pseudonym communication to propose location service requests and receive service information; in the mixed region, the user prohibits communication, cannot send and receive any location service information, and replaces their pseudonym after the user leaves the mixing region, thus making it difficult for an attacker to connect two consecutive location service requests, thereby protecting the user's location privacy. As shown in Fig.4, suppose there are three users U_1, U_2, and U_3, and the pseudonyms A, B, and C are used in the application region, and when they leave the mixing region, the system change their pseudonyms to X, Y, Z, because in the mixed region is prohibited to communication, so the attacker cannot correlate the pseudonyms used before and after the user entered the mixing region, it increases the difficulty of obtaining user identity information, so as to achieve the purpose of protecting user privacy.

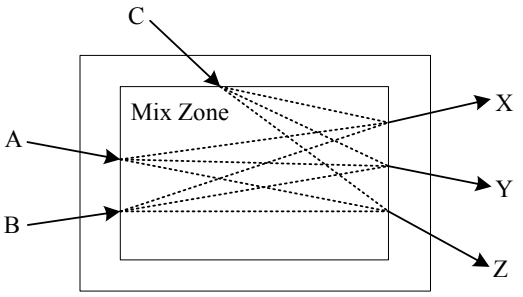


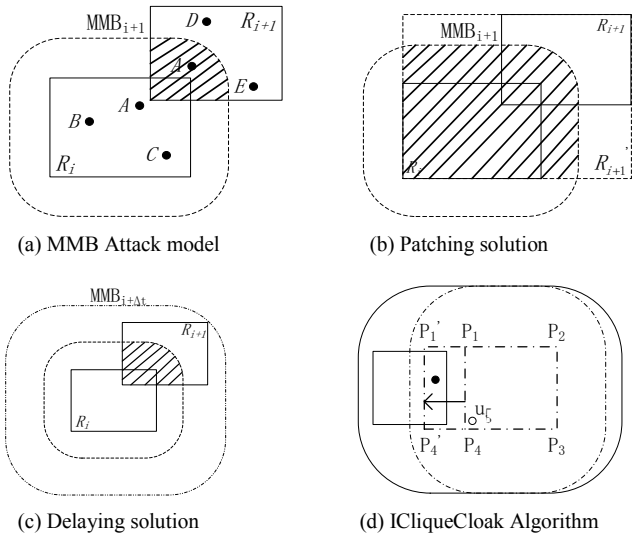
Fig. 4. Mix-zone privacy protection model

B. Maximum Movement Boundary Attacks

Maximum Movement Boundary attacks is the attacker uses the user's maximum movement speed and the cloaked region of multiple consecutive time to infer the location of the user initiated the query [8], it takes place when: (1) Location of the object continuously updated or continuous queries are issued); (2) The same user uses the same pseudonym in two consecutive location query; (3) The maximum possible speed of the user is known. As shown in Fig.5(a), user A sends a query with k = 3 at time t_i, and the cloaked region is R_i. Assuming the attacker knows that user A is in the region and knows its maximum movement speed, the maximum boundary MMB_{i+1} that user A may reach at time t_{i+1} can be calculate (As shown in the rounded rectangle with dotted line in Fig. 5(a)). When user A issues a query again at time t_{i+1}, the cloaked region becomes R_{i+1}, and the attacker can infer that the region where user A is located is only possible in the shaded area where MMB_{i+1} and R_{i+1} intersect, as shown in Fig.5(a), the cloaked region where user A is located is reduced, in extreme cases, if there is only one user in this shaded area, the location information of user A is exposed.

In order to defense maximum movement boundary attacks, Reynold Cheng et al. proposed a method of Patching and Delaying [9], The 'Pathing' method by expanding the current cloaked region to expand the overlapping area, as shown in Fig.5(b), this method combines the cloaked region R_{i+1} at the current time and the cloaked region R_i at the previous time, a combined cloaked region R_{i+1}' (As shown in the rectangle with dotted line which contains R_i and R_{i+1} in Fig.5(b)) is sent to the server. In this way, if the attacker takes the intersection of MMB_{i+1} and CR (Cloaked Region), the shaded portion in Fig.5(b) is obtained, which significantly increases the user's cloaked region so that the user's location is not easily traced. But as time goes by, the cloaked region will become increasingly large, bring a heavy burden to the server, while service quality also will decline. The 'Delaying' method extends the MMB to protect a user's location privacy by suspending the user request for Δt time, as shown in Fig.5(c), the cloaked region where the user is located at time t_i is R_i. By delaying the user's location request by Δt until the $MMB_{i+\Delta t}$ can completely cover the current user cloaked region R_{i+1}, so

that the probability of an attacker inferred the user location will be greatly reduced. However, since the user's request is delayed, the system response time increases, and at $t_{i+\Delta t}$, the user's location may have changed, thus affecting the quality of service. Pan Xiao et al. proposed the Incremental Clique-based Cloak algorithm, called ICliqueCloak [10], the main idea is that when a new query arrives, it first finds a cloaking region that satisfies the anonymous request, and then appropriately extend the cloaking region, to ensure that each user within the region will not suffer the maximum movement boundary attack [8]. As shown in Fig.5(d), the ICliqueCloak algorithm extends the anonymous region P1P2P3P4 to P1'P2P3P4'. This ensures that the MAB (Maximum Arrival Boundary) at the current moment can completely contain the cloaked region R_i at the previous time, so even if the attacker knows the maximum movement speed of the user, but cannot find out the location of the user in the anonymous region.



(a) MMB Attack model (b) Patching solution

(c) Delaying solution (d) ICliqueCloak Algorithm

Fig. 5. Maximum Movement Boundary Attack model and solution

IV. SUMMARY

With the rapid development of mobile Internet technology, location-based services (LBS) has been widely used in mobile social networks, but users access to services at the same time there is the risk of location privacy disclosure. For the location privacy protection, scholars have put forward many effective solutions, and achieved good results, but there are also some problems, such as: Many privacy protection technology is to reduce the quality of service as the premise, so how to ensure user privacy security as much as possible to improve the quality of service will be an important research topic. At the same time, the user's privacy needs are personalized, different users have different privacy protection needs, and how to meet the needs of different users of personalized privacy protection will also be one of the hot spots for future research. In addition, in the mobile social network, the attacker can obtain the user's personal profile through some means, may have complex background knowledge, therefore, how to achieve a variety of background knowledge of location privacy protection will be a challenge for future research.

ACKNOWLEDGEMENT

This work was supported by Web Culture Project Sponsored by the Humanities and Social Science Research Base of the Sichuan Provincial Education Department (No.WLWH15-26).

REFERENCES

[1] X.L. Xu, "The Research on the Problem and Protection of Mobile Social Network User's Privacy Security", (MS., Chongqing University, China 2014), p.6-11. (in Chinese)

[2] S. Qiu, "Location Privacy Protection in LBS", China Computer&Communication, 2015, (1):50-52. (in Chinese)

[3] L. Sweeney, "k-anonymity: A model for protecting privacy", International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(05):557-570.

[4] M. Gruteser, D. Grunwald, "Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking", Proceedings of the lst International Conference on Mobile Systems, Applications and Services. San Francisco, USA, 2003:163-168.

[5] J.Y. Jia and F.L. Zhang, "Overview of location privacy protection technology", Application Research of Computers, 2013, 30(3):641-646. (in Chinese)

[6] Y.N. Wu and Z.M. Zhao, "Research on Location Anonymity Method Based on Continuous Location Services Requests", Netinfo Security, 2015, (1):39-44. (in Chinese)

[7] A R. Beresford and F. Stajano , "Mix Zones:User Privacy in Location-aware Services", Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications, Workshops, p 127-131,2004.

[8] J.M. Han, Y. Lin, J. Yu, J. Jia, and L.Q. Zheng, "LBS Privacy Preservation Method Based on Location k-Anonymity", Journal of Chinese Computer Systems, 2014, 35(9):2088-2093. (in Chinese)

[9] R. Cheng, Y. Zhang, E. Bertino, and S. Prabhakar, "Preserving User Location Privacy in Mobile Data Management Infrastructures", Proceedings of 6th Workshop Privacy Enhancing Technologies,2006:393-412.

[10] X. Pan, J.L. Xu, and X.F. Meng, "Protecting Location Privacy against Location-Dependent Attacks in Mobile Services", IEEE Transactions on Knowledge and Data Engineering, 2012, 24(8):1506-1519.