False Trajectory Privacy Protection Scheme Based on Location Service

Chen Yang¹²,*, Wei Yan³⁴

¹ School of Software, Dalian University of Foreign Languages, Dalian, China
² School of Computer Science and Technology, Harbin Engineering University, Harbin, China
³ Department of Information, Liaoning Police College, Dalian, China
⁴ School of Software, Dalian University of Technology, Dalian, China

*Corresponding author e-mail: yangchendl@126.com

Abstract. When the user applies for a location continuous query service, the user location privacy may be revealed due to different types of trajectory identification. A false trajectory privacy protection scheme is proposed, which uses the real-time generation of similar trajectories to protect the user's personal location privacy. The false location set is established by initial anonymous location selection, similar trajectory location calculation and generated location selection. The false location is used to form a false trajectory similar to the real location trajectory, which reduces the probability of attacker identification and achieves real-time trajectory anonymity based on location service. Finally, the effectiveness of the scheme is verified by performance analysis and experimental results.

1. Introduction

Location-based service (LBS) [1,2] refers to the information service closely related to the user's designated geographical location. With the continuous development and improvement of mobile communication and location technology, location-based services (LBS) are gradually penetrating into people's daily life, such as users can use Google Latitude and other applications to query the food, hotel and other information in specific locations. With the widespread application of LBS, users are no longer limited to enjoying real-time query services, but more widely use the trajectory publishing composed of location sequences [3,4], such as logistics companies to save their own trajectories, to analyze whether their transport routes are reasonable. However, because trajectory location data contains abundant spatial and temporal information, personal privacy information such as interest preferences, social status and health status of mobile users is often revealed. In addition, the user's real-time location information disclosure can also pose a serious security risk.

In LBS, privacy includes location privacy and query privacy. Location privacy refers to personal sensitive information that is directly related to the user's location or can be estimated based on the user's location. Query privacy refers to the personal sensitive information exposed by the user's query content. In order to solve the problem of privacy protection in LBS, many methods have been proposed in recent years. Location K-anonymity is the most widely used technology among them. K-Anonymous [5,6] technology was first proposed by Samarati and Sweeney to solve the privacy protection problem of data publishing in relational databases. The core idea is to generalize the attribute values of records so that each attribute tuple in the database has at least k records. Gruteser et al. [7] first introduced K-anonymity mechanism into the privacy protection of location-based services.
By reducing the spatial and temporal location accuracy of mobile users, the user's precise location was replaced by a location area containing k indistinguishable users. Although this method can protect the privacy of a single query, it cannot protect the trajectory privacy of users, nor can it defend against location attacks with background information, such as location link attacks [8], identity matching attacks [9] and so on. For that reason, the researchers proposed a method of trajectory privacy protection [10,11]. This type of method mainly protects the user's real trajectory from being recognized by generating (K-1) false trajectories. Compared with the previous technology, this method considers the correlation between locations and can defend against attacks against a single location. However, the positions on false trajectories are always generated randomly. There are often some positions that do not conform to reality [12]. This makes it easy for attackers to eliminate false trajectories by combining background information, which affects the effect of trajectory privacy protection. How to protect the generated and upcoming location trajectory information and reduce the risk of user privacy leakage has become an important issue for current researchers.

This paper considers the characteristics of trajectory anonymity and real-time continuous location service privacy protection, and proposes a false trajectory generation method based on location service. Based on the user's real trajectory, a similar trajectory constructed from real trajectories is built. The similar trajectory and the user's real trajectory have the same motion pattern, which can ensure the similarity with the user's real trajectory, and can also cope with background information attacks to achieve K-anonymity protection of the user trajectory.

2. Related Work

Location trajectory privacy protection originated from data publishing at the earliest time. There are three kinds of privacy protection methods for location trajectory publishing: Trajectory k-anonymity [13,14], trajectory suppression [15,16] and false trajectories [17-19]. The trajectory k-anonymity means that k-1 trajectories which cannot be distinguished from the real trajectory of users are selected by anonymous server before the trajectory is published. Trajectory privacy is protected by generalizing the position of corresponding time in K trajectories into corresponding anonymous regions. Xu et al. [13] proposed to use the historical footprint of other users to select k–1 tracks and form corresponding anonymous areas. Subsequently, Wang Chao et al [14] added the shape factor of the trajectory and proposed a privacy protection scheme based on the shape similarity of the trajectory position. However, the trajectory k-anonymous privacy protection method needs to introduce a trusted third party as an anonymous server, which reduces the practicability.

The trajectory suppression refers to the removal of sensitive or frequently accessed locations before the publishing of trajectories. Terrovitis et al. [15] assume that the attacker has part of the real trajectory. It is proposed to suppress the location in the sensitive area when the trajectory is published, so that the risk of trajectory leakage is not higher than the privacy protection threshold set by the user. Among them, the sensitivity of the region is calculated according to the ratio of the number of users in the region to the total number of users. Zhao Jing et al. [16] Based on the trajectory frequency, two trajectory suppression privacy protection schemes are proposed by adding false data to the problematic trajectory and using the relationship between privacy relevance and data utility to process the trajectory data locally. However, how to find the location information that is reasonably suppressed is still the key problem to be solved in the privacy protection method of trajectory suppression.

The false trajectory method refers to the fact that the user himself generates k-1 false trajectories similar to the real trajectory, and publishes the set with k trajectories with the real trajectory. The basic idea is first proposed by Kido et al. [20, 21]. In their scheme, users can use the false positions [21-23] generated by their two requests to form false trajectories to protect their real trajectories. WU et al. [24] proposed a false trajectory generation method based on mobile terminal. According to the order of the positions in the real track of the user, the method sequentially selects a point in a direction at a certain angle from the corresponding position to generate a false position until a false track is generated. The false trajectory generated by the method avoids the coincidence with the real trajectory and plays the role of protecting the trajectory privacy. However, it is not possible to defend against location privacy attacks with background information. HARA T et al. [25] proposed a false trajectory generation method based on vehicle trajectory. When generating a false trajectory, the method comprehensively considers the trajectory motion pattern of the vehicle movement, and thus can generate a false trajectory similar to the real trajectory, and can defend the location privacy attack with the background information to a certain extent. However, the generated trajectory will have some locations that cannot
be reached in reality, which affects the trajectory privacy protection effect. At the same time, although these methods can protect the trajectory privacy, they are easily identified by the attacker using specific attack means because they are false locations on the false trajectory. PHAN T N et al. [26] proposed constructing false trajectories using the user's historical trajectory. BINDSCHAEDLER V et al. [27] proposed a method for constructing false trajectories using real user trajectories. These methods are constructed based on real user trajectories when constructing false trajectories. However, it is difficult to find user trajectories with high similarity in specific regions, and it is difficult to achieve K-anonymity.

3. System Architecture

The system architecture based on location-based service privacy protection is mainly divided into two types: two-tier architecture and three-tier architecture. The two-tier architecture is composed of users and location service providers (LSP). The structure is simple and does not need to depend on third parties, but it has the disadvantages of overweight computing load of users themselves and unreliable or low collaboration degree of collaborative users. The three-tier architecture is to add a trusted third-party anonymous server (AS) between users and LSP. This architecture can provide better privacy protection effect, while reducing the user's own load. Due to the limited computing power of mobile communication devices, this paper adopts a three-tier central server architecture. As shown in Figure 1, the architecture consists of three entities: mobile user (U), anonymous servers (AS), and location service provider (LSP). U carries positioning devices (such as smartphones, GPS positioning module), which can accurately locate their location and interact with LSP, and then enjoy location-dependent services. However, U's communication and computing capabilities are limited. AS is responsible for calculating the location of similar trajectories and has strong computing and data processing capabilities. LSP can provide services for User or AS, and has strong data analysis and processing capabilities. As can be seen from Figure 1, the user first sends the location query to AS. Anonymous locations which can generate similar trajectories are generated by the location calculation module in AS, and then sent to LSP after completing the anonymity by the location anonymity module. After the LSP uses the query module to query, the Query result is saved to the historical data, and the result is returned to the AS. The AS filters the obtained anonymous query results and sends them to the user to complete the entire query process.

The continuous query service means that the User initiates multiple requests to the LSP, and the request can represent $q = (id, loc, t, p)$. Among them, $id$ represents the user's identity or pseudonym; $loc$ represents the location of the current query; $t$ represents the time interval; $p$ represents the point of interest of the query, such as restaurants, gas stations, etc. In this paper, User first sends $q$ to the AS. After the AS completes the anonymous location calculation, it generates a false user that satisfies the trajectory anonymity. The User establishes an anonymous group with the false user. The AS sends the anonymous group as the requesting user to the LSP. LSP returns the query results to AS after completing the query. After filtering the false query result, the AS returns the real query result to the User to complete the current query.
4. False Trajectory Generation Method

4.1. Initial Anonymous Location Selection

In order to obtain the best results of trajectory anonymity, try to find a continuous locating position that can be equal in length and in the same direction. However, due to the influence of specific geographical conditions such as actual road sections, there may be a small interval between the trajectories in the generated trajectory, which is called the accompanying trajectory. This kind of trajectory can achieve better trajectory anonymity in the published privacy protection, but in the location service, especially the continuous location query service, this accompanying trajectory may have common query locations and points of interest, so that the attacker can any other trajectory obtains the smallest area where the base user is located, thereby obtaining the privacy of its location trajectory. As shown in Figure 2, T is a real trajectory composed of four locations A, B, C and D, and T’is a false trajectory composed of four locations A’, B’, C’, D’. Both trajectories pass through similar or identical query locations and can be viewed according to the same trajectory. At this time, if an attacker knows any query location of any user, the location privacy of the base user can be known. In the method proposed in this paper, the minimum threshold of trajectory direction is used in the following stage. In order to eliminate the hidden danger caused by the accompanying trajectory, the distance between the initial anonymous false user and the base user should be enlarged in the initial anonymous stage. Therefore, in the initial location anonymity phase, the distance between the user locations needs to be defined to prevent the trajectory accompanying phenomenon after the subsequent anonymous operations are completed, which affects the user's location privacy.

Figure 2. Accompanying trajectory

4.2. False Trajectory Generation Algorithm

By calculating the location of similar trajectories, the set of positions satisfying the generation of false trajectories is obtained, and the similar trajectories are established through position selection. Assuming that user U and false user V establish anonymous groups with loc_{u} and loc_{v} as the initial query locations respectively, in order to find the follow-up query locations of loc_{u+1} and loc_{v+1}, and connect the adjacent query locations of user U and V to generate similar trajectories, the following query locations of loc_{u+1}, loc_{v} and loc_{v+1} are needed. The equation is established by using parallel distance and angle distance, and then the position coordinates of loc_{v+1} are solved. Let r_{u} and r_{v} be trajectory sequences at both ends with loc_{u}, loc_{u+1}, loc_{v}, loc_{v+1}. θ denotes the angle threshold between two trajectories, and d denotes the parallel distance required for calculating the location region of similar trajectories.

\[
\cos \theta = \frac{|loc_{u}loc_{u+1}| - |loc_{v}loc_{v+1}|}{|loc_{u}loc_{v+1}|}
\]

\[
d = |loc_{u}loc_{v+1}| - |loc_{v}loc_{v+1}|
\]

Dist denotes the trajectory distance between two trajectory segments, where \(\omega_1\) and \(\omega_2\) denote the weight of the parallel distance and the angle distance respectively, and the default values are set to 1.

\[
dist(r_{u}, r_{v}) = \omega_1 d + \omega_2 \sin \theta
\]

The similarity between the two trajectories is \(\text{sim}(U, V)\), and \(0 < \text{sim}(U, V) \leq 1\).

\[
\text{sim}(U, V) = 1 - \frac{\text{dist}(r_{u}, r_{v})}{\text{dist}(r_{u}, r_{v}) + \text{dist}(r_{u}, r_{v})}
\]

In order to reduce the probability of attackers identifying false locations, it is necessary to screen the similar locations of the generated false trajectories. For inaccessible location areas, such as building block location areas, should be deleted. Combine the idea of initial anonymity, similar
trajectory position calculation and filtering to generate position, and design false trajectory. The specific algorithm is shown in Table 1.

| Table1. False Trajectory Generation Algorithms |
|-----------------------------------------------|
| Input: \( loc_u, \ loc_{u+1}, \ loc_v \)      |
| Output: \( loc_{v+1} \) set                  |
| 1: Received request q                         |
| 2: Set similarity and angle thresholds, \( \text{sim}(U,V) \) and \( \cos \theta \) |
| 3: Set the weight of parallel distance and angle distance, \( \omega_1 \) and \( \omega_2 \), and set the default value to 1 |
| 4: set S=null;                                |
| 5: for every V                                |
| 6: {                                           |
| 7: Calculate the angular distance \( \sin \theta \) between two tracks |
| 8: Calculate the parallel distance \( d \) of the two tracks |
| 9: Calculate the distance \( \text{dist}(r_u, r_v) \) between two track segments |
| 10: Calculate dot(U,V)                        |
| 11: Solving the binary quadratic equation about \( \text{loc}_{v+1} \), \( \text{loc}_{v+1} = \text{solve}(d, \sin \theta) \) |
| 12: Use \( \text{loc}_{v+1} \) to create a similar location |
| 13: If the current area can be reached from \( \text{loc}_v \) |
| 14: Randomly select any location to join the set S |
| 15: else                                      |
| 16: Break;// anonymous failure                |
| 17: end if;                                   |
| 18:}                                          |

5. Experimental Results and Analysis

5.1. Performance Analysis

In the running process of the similar trajectory real-time generating method, since the anonymous group is established by using the method of generating a false user, the user pseudonym, the query content and the query time interval are all consistent with the real user. Therefore, on the basis of satisfying trajectory anonymity and effectively resisting mobile type association attacks, it also has good resistance to pseudonym association, content association and time interval Association attacks.

To achieve trajectory \( r \)-anonymity, \( r \) similar trajectories need to be generated in each continuous query interval. In the best case, the time complexity is \( O(r^{-1}) \). If each generated position is unreachable, the similar parameters need to be adjusted. At this time, the execution time complexity of the algorithm is \( 2O(r^{-1}) \). Therefore, in general, the time complexity of the algorithm is \( O(r) \).

5.2. Result Analysis

This experiment uses MATLAB R2016A software to simulate on a PC (Intel i7 CPU, 2.8GHz, 8GB memory). The experimental data is based on the real location data provided by the BerlinMOD Data Set, and the city center location is randomly selected for experimental verification. At the same time, this paper adds LTPPM [28] method for comparison. LTPPM mainly targets trajectory difference recognition attack.

In the comparison experiment of execution efficiency, the difference in execution time of each algorithm is checked according to the change of the anonymous value. At the same time, view the continuous execution time under the same anonymous value limit during consecutive multiple executions. Let's take 3 consecutive queries as an example to see the execution time changes of each algorithm as the anonymous value increases.
As can be seen from Figure 3, LTPPM resists trajectory discrimination attacks by building similar trajectories through user collaboration. This method involves not only searching for collaborative users, but also screening anonymous trajectories. Therefore, the time efficiency of LTPPM is most affected by the change of anonymous values. This method only deals with similar trajectory positions, and does not need to find collaborative users, and has better execution efficiency than LTPPM.

From Figure 4, we can see the execution time of each algorithm in the continuous running process with the same parameters. Among these algorithms, the algorithm proposed in this paper needs to filter the location of similar trajectories after calculating the anonymous location, and its execution time is lower than LTPPM, which is a privacy preservation algorithm for searching cooperative users to establish similar trajectories.

6. Conclusion

In this paper, an attacker can obtain user privacy through location trajectory difference in the process of LBS continuous query. The idea of trajectory similarity is introduced, and a real-time generation method of false trajectory is proposed. The location privacy protection in the continuous location service process is realized by initial anonymous location selection, similar trajectory location calculation and generation location screening. By comparing the performance analysis with the experimental results, it further shows that the method has better privacy protection effect.

7. Acknowledgments

This work was financially supported by the Basic Scientific Research Projects of Liaoning Province (No.2017JYT09).
References

[1] Krumm J. A survey of computational location privacy[J]. Personal and Ubiquitous Computing, 2009, 13(6): 391-399.

[2] Wang S, Hu Q, Sun Y, et al. Privacy Preservation in Location-Based Services [J]. IEEE Communications Magazine, 2018, 56(3): 134-140.

[3] Zheng X, Cai Z, Li J, et al. Location-privacy-aware review publication mechanism for local business service systems[C]//International Conference on Computer Communications. IEEE, 2017: 1-9.

[4] Zhang Z, Sun Y, Xie X, et al. An efficient method on trajectory privacy preservation[C]//International Conference on Big Data Computing and Communications. 2015: 231-240.

[5] Samarati P, Sweeney L. Generalizing data to provide anonymity when disclosing information[C]/Proceedings of the 17th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 1998.

[6] Sweeney L. k-anonymity: a model for protecting privacy[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(5):557-570.

[7] Gruteser M, Grunwald D. Anonymous usage of location based services through spatial and temporal cloaking[C] //International Conference on Mobile Systems, 2003:31-42.

[8] Ding Yuan, Peddinti S T, Ross K W. Stalking Beijing from Timbuktu: a generic measurement approach for exploiting location-based social discovery[C]//Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices. New York, USA: ACM Press, 2014:75-80.

[9] Hochreiner C, Huber M, Merzdovik G, et al. Towards practical methods to protect the privacy of location information with mobile devices[C]//Proceedings of International Conference on Security of Information and Networks, 2014:17-24.

[10] Mano K, Minami K, Maruyama H. Pseudonym exchange for privacy-preserving publishing of trajectory data set[C]//Proceedings of Consumer Electronics. Washington D.C. USA: IEEE Press, 2015:691-695.

[11] Sheng Hang, BAI Guangwei, Yang Mei, et al. Protecting trajectory privacy: a user-centric analysis [J]. Journal of Network & Computer Applications, 2017, 8(2):128-139.

[12] Chen J, He K, Yuan Q, et al. Blind Filtering at Third Parties: An Efficient Privacy-Preserving Framework for Location-Based Services [J]. IEEE Transactions on Mobile Computing, 2018, 17(11):2524-2535.

[13] Xu T, Cai Y. Exploring historical location data for anonymity preservation in location-based services[C]//The 27th Conference on Computer Communications. IEEE, 2008.

[14] Wang C, Yang J, Zhang J P. Privacy preserving algorithm based on trajectory location and shape similarity [J]. Journal on Communications, 2015, 36(2): 144-157.

[15] Terrovitis M, Mamoulis N. Privacy preservation in the publication of trajectories[C]//9th International Conference on Mobile Data Management. IEEE, 2008: 65-72.

[16] Zhao J, Zhang Y, Li X H, et al. A trajectory privacy protection approach via trajectory frequency suppression [J]. Chinese Journal of Computers, 2014, 37(10): 2096-2106.

[17] You T H, Peng W C, Lee W C. Protecting moving trajectories with dummies[C]//2007 International Conference on Mobile Data Management. IEEE, 2007: 278-282.

[18] Lei P R, Peng W C, Su I J, et al. Dummy-based schemes for protecting movement trajectories [J]. Journal of Information Science and Engineering, 2012, 28(2): 335-350.

[19] Wu X, Sun G. A novel dummy-based mechanism to protect privacy on trajectories[C]//2014 IEEE International Conference on Data Mining Workshop (ICDMW). IEEE, 2014: 1120-1125.

[20] Kido H, Yanagisawa Y, Sato T. An anonymous communication technique using dummies for location-based services[C]//International Conference on Pervasive Services. IEEE, 2005: 88-97.

[21] Kido H, Yanagisawa Y, Sato T. Protection of location privacy using dummies for location-based services[C]//21st International Conference on Data Engineering Workshops. IEEE, 2005: 1248-1248.
[22] LU H, JENSEN C S, YIU M L. Pad: privacy-area aware, dummy based location privacy in mobile services[C]//The Seventh ACM International Workshop on Data Engineering for Wireless and Mobile Access. ACM, 2008: 16-23.

[23] NIU B, LI Q, ZHU X, et al. Achieving k-anonymity in privacy-aware location-based services[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. IEEE, 2014: 754-762.

[24] WU Xichen, SUN Guangzhong. A novel dummy-based mechanism to protect privacy on trajectories[C]// Proceedings of IEEE International Conference on Data Mining Workshop. Washington D. C., USA: IEEE Press, 2015:1120-1125.

[25] HARA T, ARASE Y, YAMAMOTO A, et al. Location anonymization using real car trace data for location based services[C]//Proceedings of the 8th International Conference on Ubiquitous Information Management and Communication. New York, USA: ACM Press, 2014:34.

[26] PHAN T N, KUNG J, DANG T K. K UR-algorithm: from position to trajectory privacy protection in location-based applications[C]//Proceedings of International Conference on Database and Expert Systems Applications. Berlin, Germany: Springer-Verlag, 2015:82-89.

[27] BINDSCHAEDLER V, SHOKRI R. Synthesizing plausible privacy-preserving location traces[C]// Proceedings of IEEE Symposium on Security and Privacy. Washington D. C., USA: IEEE Press, 2016:546-563.

[28] GAO S, MA J F, SHI W S, et al. LTPPM: a location and trajectory privacy protection mechanism in participatory sensing [J]. Wireless communications & mobile computing, 2015, 15(1): 155—169.