International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-09, Issue-04, April 2022

LBS Privacy Preservation Research on Location Semantic Based Fake Trajectory

Rongyu Wang, Yongyi Yang, Xingyu Wu, Kaihua Li

Abstract—**With the rapid development of location technologies, such as GPS, Wi-Fi and Radio Frequency Identification (RFID), Location-Based Service(LBS) are becoming increasingly popular in people's lives. While enjoying the great convenience brought by LBS to life, users' privacy is exposed to the risk of leakage. The contemporary traditional LBS privacy protection mainly considers spatio-temporal information but ignores inferring attacks based on user-related background information which leads to the attacker can use the user's background information to mine the user's identity information, personal preferences, behavioral habits, and other privacy information. According to the user's real trajectory influenced by the key map points in the area, this paper proposes an LBS privacy protection scheme incorporating map semantic information. The method combines the user's location trajectory and map influence to construct a privacy protection model with synergistic time, space, and map influence. And according to the model, the secure anonymization area that matches the user's motion logic and the semantic match between the user's corresponding map point is screened out, thus finally achieving the role of protecting the user's LBS privacy information.**

Index Terms—LBS privacy protection; map semantics; fake trajectory; location services

I. INTRODUCTION[1]

In recent years, mobile communication technology, positioning technology(GPS) and Internet technologies have fused to promote LBS(Location-Based Service) in the application industry market. While LBS providing people with efficient life applications, it has also raised serious issues of privacy leakage. Due to the characteristics of location-based services, users enjoy LBS services, users must first provide their location information and send the query request to the service provider LSP(Location Service Providers).According to the query results the LSP returned to the user, for example, driving navigation services, to find the surrounding restaurants and entertainment services, social services, travel services, etc. Although LBS can provide users with various quality services based on location information, the service provider is likely to leak users' private information to unscrupulous elements to gain profit. Once the geographical location of users is determined, malicious attackers can combine the location information with related service contents to obtain users' private information from it. The attacker can also use the trajectory formed by the user's location over time to deduce the user's behavior, home address, workplace and other private information. Therefore, while providing location services to users, the protection of their location privacy information is particularly important.

Currently, scholars at home and abroad have proposed various protection methods for trajectory privacy protection, such as trajectory generalization, trajectory suppression, adding fake trajectory methods, and differential privacy. Among them, the trajectory k-anonymity privacy protection method is one of the commonly used trajectory privacy and query protection methods, which means that the anonymous server selects k-1 trajectories that are indistinguishable from the user's real trajectory before the trajectory is published, and forms the path by generalizing the location of the corresponding moment in the k trajectories to the corresponding anonymous area to achieve the purpose of obfuscating the real data and ensuring that the probability of the user's real location being identified is 1/k. Although trajectory k-anonymity privacy protection has achieved a certain effect on trajectory privacy, in the continuous request service, on account of the user's historical trajectory has temporal and spatial continuity, and also has a certain correlation with background information, etc., it is still possible for an attacker to infer the user's real movement trajectory after observing and filtering the mastered information for several times, which makes the existing protection model ineffective against the inference attack.

For example, in the present time, an attacker can use the geographical attributes of map points (e.g., narrowing down the anonymous region to the only currently geographically passable road on the map, as in the left side of Figure .1), and the semantic information of map points (e.g., the attacker mines the trajectory information of a student and finds that the student frequently passes by the library instead of the trade center, as in the right side of Figure.1), and uses probabilistic statistics to speculate on excluding unreasonable anonymous regions, thus leading to the problem of exposing the user's location privacy information.

Manuscript received April 22, 2022

Rongyu Wang, School of Computer Science and Technology,School of Software,Tiangong University,Tianjin,300380,China
Yongyi Yang, School of Computer Science and Technology,School of Software,Tiangong University,Tianjin, 300380,China
Xingyu Wu, School of Computer Science and Technology,School of Software,Tiangong University,Tianjin,300380,China
Kaihua Li, School of Computer Science and Technology,School of Software,Tiangong University,Tianjin,300380,China

[1] Fund projects：Funded by the National Innovation and Entrepreneurship Training Program for College Students-NO.202110058013

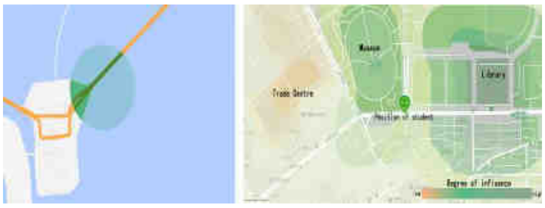


Fig.1 Location point semantic map

For conventional privacy protection schemes, it is difficult to defend against inference attacks that use map semantic information combined with trajectory spatio-temporal models. So this paper constructs a space-time-map influence privacy protection scheme based on the probability graph model based on the map semantic information of mobile users. The author's contributions are:

(1) The constraints and influence of map points on user behavior trajectories are considered comprehensively, and a measurement model of the influence of map semantic information on the probability distribution of user trajectories is proposed.

(2) For the current the influence of the map points on the distribution of the trajectory of users with different social identities to infer the trajectory of the user LBS Privacy attacks lack effective countermeasures. This paper proposes an LBS privacy protection model of space-time-map influence privacy protection scheme based on the traditional spatio-temporal correlation attack for speculative attacks in this area.

II. Related Work

At present, numerous privacy protection techniques have been proposed by domestic and foreign scholars in the field of privacy protection of location services. The first application of k-anonymity to LBS privacy protection was proposed in the literature[1], where the main idea is to mix anonymity in location requests against anonymous users with no less than k - 1 other user requests, such that the probability that an attacker can infer the true unknown of the user does not exceed 1 / k. The k-anonymity privacy-preserving model was later enriched by related researchers.

The literature[2]proposes the l-diversity model, which targets location homogeneity attacks and requires k anonymous locations to be l-diversity to resist location homogeneity attacks. Since k-anonymity is protected against snapshot locations, it does not provide effective anonymity protection under continuous location queries.

In the literature[3], based on the k-anonymity privacy-preserving techniques, is proposed for the maximum movement boundary and maximum arrival boundary constraints of spatio-temporal relations. The privacy-preserving technique RobLop for spatio-temporal constraint attacks (a Robust Location preserving algorithm). To address the limitations of snapshot location protection techniques in continuous location query attacks, researchers have proposed techniques such as trajectory generalization, trajectory anonymization, virtual trajectories , and differential trajectories for trajectory privacy based on the k-anonymity protection model[4-7]. Among them, virtual trajectory not only can save users' real trajectory but also can effectively protect users' trajectory privacy, making them widely used in trajectory publishing.

Xu et al[4]not only proposed to use the historical footprints of other users to select k-1 trajectories to form the corresponding fuzzy regions, and also proposed an entropy measure mechanism based on the size of the fuzzy regions, which not only considers the number of entities in the fuzzy regions, but also their anonymity probability distribution to quantify the anonymity of the system.

Zhong et al[5]divided the user activity area into sensitive levels and set the life cycle of the pseudonym, when the life cycle of the pseudonym ends or the user enters or leaves the sensitive area, the pseudonym will be changed and the anonymous area will be formed with the aid of footprints to protect the privacy of user's trajectories.

Huang et al[6]in Silent Period[7,8] based on the silent period method to protect the privacy of user's trajectories, which divides the spatial area of user's motion into mixed area and application area, and the user's location in the mixed area neither will neither send request information to LBS server nor receive any service information.The user neither sends request information to the LBS server nor receives any service information until the mobile user changes his pseudonym and exits from the hybrid area. Mobile users will use different pseudonyms when leaving or entering the hybrid zone, thus protecting the user's trajectory information, but the lack of communication between the hybrid zones will lead to degradation of service quality.

Kido et al[9-10]were the first to propose a method to protect user's trajectory privacy by constructing fake trajectories in consecutive requests, and You et al[11]proposed two fake trajectory generation schemes for the number of locations corresponding to a moment in time and the total number of trajectories in the trajectory collection. In the first scheme, the user first decides the starting point and the ending point of the fake trajectory, and then randomly generates fake trajectories with similar running patterns to the real trajectory between the starting point and the ending point; The second scheme takes the user's real trajectory as the base and rotates the real trajectory to generate the fake trajectory by selecting the position point on the real trajectory as the rotation axis point; Scholars Lei et al[12]proposed the method of adding intersection points on the trajectory obtained after rotation to increase the number of fake trajectories, thus improving the privacy protection level of the user's real trajectory. Scholars Wu et al[13]considered not only the distance between real and fake trajectories but also the distance between fake trajectories. They perturb the generated fake trajectories so that the final set of trajectories can satisfy the privacy protection needs of users.

Feng Hua Li and other scholars[14]studied the trajectory privacy protection scheme under the scenario that user action patterns and trajectory similarity and other features may be obtained by adversaries, based on the user's real trajectory, the similarity of the user's action trajectory is ensured by trajectory rotation, finally, the points on this trajectory are offset to the nearby location closest to the real service request probability to generate a fake trajectory.

Kaiyue Lei[15]proposed a pseudotrajectory generation scheme based on spatio-temporal correlation by analyzing the overall direction of trajectories, spatio-temporal correlation, and trajectory similarity. The scheme fits the overall direction of the real trajectory, generates the starting and ending positions and center points of the fake trajectory, ensures that the fake trajectory has the same direction and slope as the real trajectory, checks the moving distance and reachability of each segment of the path in the generated fake trajectory, makes it difficult for the attacker to identify the fake trajectory, and protects the privacy of the user's trajectory.

Although the above methods have achieved certain achievements in location privacy and trajectory privacy protection, due to the correlation between the user's historical trajectory and map semantic information, and the drawback that the fake trajectory does not match the user's behavior pattern, the attacker can perform speculative attack through background knowledge to screen the fake trajectory that does not match the user's behavior pattern, to find the user's real moving trajectory and make the fake trajectory protection method invalid. To address the above issues, this paper constructs a location privacy anonymization region based on a fake trajectory privacy protection method by measuring the influence of key map points and the relationship between map influence and users, and generates fake trajectories using the k-anonymity algorithm, and uses the generated fake trajectories to replace the service query requests sent by users at their real locations in continuous time. By using the PrefixSpan algorithm to mine the dataset to get frequent sequences, delete the fake trajectories that do not match the frequent sequences, and finally get the secure anonymous region that matches the user's movement pattern and the corresponding map semantics, to achieve the purpose of user location privacy and trajectory privacy protection and resist the attacker's speculative attacks combining map influence and background knowledge.

III. Preparing Knowledge

3.1 Related Definitions

3.1.1 Check-in trajectory

The check-in trajectory consists of check-in points with continuity in time and space for a certain user. The check-in points count the check-in information of the user at a certain location, denoted as:

$$\text{point} = \{uid,\ timestamp,\ lat,\ lng,\ poi\},$$

$$\text{traj(n)} = \{point_{1},\ point_{2},\ \ldots\ldots,\ point_{m}\}$$

Table 1. Item symbol and meaning

Symbols	Meaning
point.uid	Check-in user's identification code
point.timestamp	Record sign-in time
point.lat	Latitude of the location
point.lng	Longitude of the location
point.poi	Semantic information about the location
traj(n)	The nth sign-in trajectory obtained after data set pre-processing
point $_{i}$	The i-th check-in point in the trajectory (0 <= i <= m) and point $_{i-1}$. timestamp < point $_{i}$. timestamp

3.1.2 User request

Simulate a user initiating a request with <id, t, (x, y), k, Q>, where id is the request identifier, t is the time information, (x, y) is the user plane location information, k is the anonymous request parameter, and Q is the query content.

3.2. System structure

The validation system of the LBS privacy protection algorithm model is constructed using a highly stable, feasible centralized model, as shown in Figure.2, consisting of three main components: several mobile communication devices requesting LBS services, a trusted third-party LBS privacy protection server (both referred to as the protection server hereafter) and an untrusted server that provides LBS services. The protection server is responsible for accepting anonymous user requests, which are transformed into anonymized requests <i, R, Q> based on anonymous user requests, where i is this anonymous request identifier, R is the coarse-grained planar geolocation information after K anonymization, and Q is the query content.

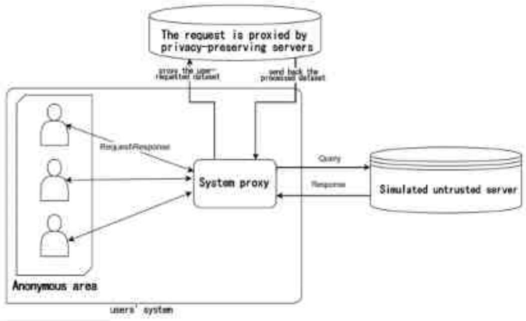

The request is proxied by privacy-preserving servers
send back the processed dataset
Request/Response
System proxy
Query
Response
Simulated untrusted server
Anonymous area
users' system

Fig.2 Algorithm model validation system

The workflow of the LBS service process is as follows.

(1) The protection server has been initialized and completed the historical trajectory database in advance,and the user trajectory behavior model is constructed by using the established location semantic database while relying on the above database and the relevant algorithms proposed in this thesis.

(2) When a regional user (the corresponding mobile communication device in practice) requests LBS services from an untrusted server, the request is first proxied to a trusted third-party protection server.

(3) When the protection server obtains a legitimate LBS service proxy request that meets the requirements of the algorithm, it generates the corresponding data and sends the data proxy to the untrustworthy server after generalizing and desensitizing the user's LBS privacy sensitive information by the model algorithm.

(4) The untrusted server returns the real requested LBS service.

3.3 Temporal correlation attack

The traditional k-anonymity protection model is based on the minimum anonymity region of the target user, the privacy level, and the maximum tolerable latency ($A_{\min}$, *k,T* $_{\max}$). The three parameters establish the anonymous region, but when the user continuously requests, the attacker can obtain the maximum movement boundary that the user can move in the corresponding time by contacting the movement speed of the target user (e.g., road speed limit and usual speed, which can be obtained using IoT sensors, etc.), and the information of the corresponding interval of the server. The motion boundary obtained by combining spatio-temporal relations will overlap with the anonymized region of the user at the next moment, and the attacker can exclude the non-overlapping part of the anonymized points by this means. This enables probabilistic prediction of the correlation between target users and sensitive source data. This type of attack is called MMB attack[3]. And using the reverse thinking, by being able to reach the maximum reach boundary at that point, as well as obtaining the anonymous region at the previous moment. The intersection of the two is the same as the area where the target user is likely to appear, so the attack of excluding anonymous points to increase the prediction possibility is called MAB attack.

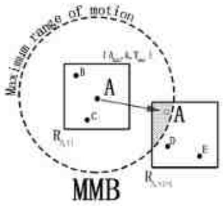


Fig.3 MMB Attack

To the overlapping region of the maximum motion boundary of t_{i+1} where the MMB attack is the anonymous region before and after the anonymous user at the time of t_{i+1} with the anonymous user presence region at the time of t_i , i.e., the shaded part is shown in Fig.3, and according to its presence region, the anonymous user is locked by the background knowledge attacks.

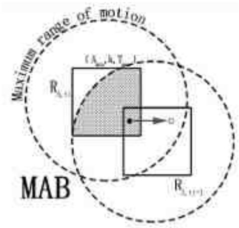


Fig.4 MAB Attack

The MAB attack, on the other hand, is similar to the MMB attack by using the overlap region between the anonymity region at the t_i time and the maximum arrival boundary from the t_{i+1} time to the t_i time as the presence region of the anonymous user, i.e., the shaded part shown in Figure .4, according to which the attack is carried out.

3.4 Sequence pattern mining algorithm

3.4.1 PrefixSpan Research Background

Sequential pattern mining is to mine patterns with high frequency relative to time or other patterns from sequential databases, and in practice, sequential pattern mining is widely used in various sequential datasets, proposed by Agrawal and Srikantin 1995[16], followed by their generalized sequential pattern (GSP) mining algorithm[17], and then Han et al. proposed the pattern growth method PrefixSpan for sequential pattern mining[18], and its precursor algorithm FreeSpan[19]in 2001. Among them,PrefixSpan, i.e., pattern mining by prefix projection, is an association rule mining algorithm based on frequent pattern growth, which is more efficient than other pattern mining algorithms due to the advantage of using database projection techniques for depth-first search without generating candidate sequence patterns, which not only reduces the storage space needed to retrieve for the operation of the algorithm but also makes the algorithm have the ability to handle very large sequence databases, so it is more efficient than other pattern mining algorithms.

3.4.2 Related Concepts

Prefix: The sequence A={a_1,a_2...$_a$n} and the sequence B={b_1.b_2,..b_m}, n≤m, satisfying $a_{i\,=}\,b_i$ (1 ≤ i ≤ n) and $a_n \subseteq b_n$, call A a prefix of B.

Suffix: That is, the prefix projection, for a certain prefix A, the sequence B prefix after the remaining sub sequence B´ that is, the suffix.

Support: The number of sequences A contained in database D, denoted as support(A), where the minimum support of sequence A contained in database D is denoted as min_support.

Frequent sequence: If the support of sequence A is not less than the minimum support min_support in database D, then A is considered as frequent sequence.

IV. Time-Space-Map Influence LBS Privacy Protection Model

4.1 Fake trajectory to resist spatio-temporal attacks

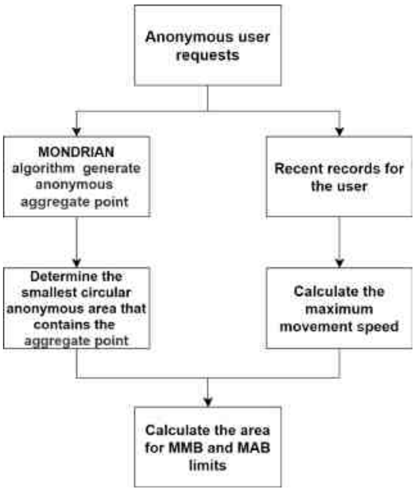
Based on the privacy protection for users with continuous requests, a dummy trajectory satisfying the resistance to MMB and MAB attacks is generated using k-anonymity,in the following steps.

Simulate a user initiating a request with <id, t, (x, y), k, Q>, where id is the request identifier, t is the time information, (x, y) is the user plane location information, k is the anonymous request parameter, and Q is the query content.

step1: Using MonDrian algorithm to generate anonymous point set in each successive request of user's location point information, the generation rule is to find k points in the data set with similar location point lat and lng and similar timestamp, to reach k-anonymous generalization, and plan the minimum circular area of each anonymous point set, if the request time is exceeded, the request will be returned to the user, and declare the request failed.

step2: Calculate the maximum velocity v of the requested user based on the timestamp, lat, lng of the data set. From t_i to t_{i+1} The maximum velocity v of the anonymous user at the time of the request.

step3: Based on the user's planar position information (x, y), the distance *d* between two consecutive requests is obtained, and the maximum motion distance d_max of the user is calculated based on the maximum motion speed v possessed by the anonymous user object. The radius difference between the front and back anonymous regions of the user is r, then d < min{d_max - r, d_max+ r} to get the corresponding motion boundary range and screen the false trajectories that satisfy the spatio-temporal association constraint in the anonymous region.


Anonymous user requests
MONDRIAN algorithm generate anonymous aggregate point
Recent records for the user
Determine the smallest circular anonymous area that contains the aggregate point
Calculate the maximum movement speed
Calculate the area for MMB and MAB limits

Pig.5 Flowchart to defend against temporal correlation attacks

4.2 PrefixSpan mining frequent sequence patterns

The PrefixSpan algorithm is used to mine the data set to get the frequent sequences of the user satisfying the minimum support, and the false trajectories that do not meet the frequent sequences are deleted to get the safe anonymous regions that satisfy the user's motion logic and the semantic match between the user and the corresponding map points, and the specific steps are as follows.

step1: Read in the sequence database D and the minimum support threshold min_support.

step2: Set the length of the first sequence L=1, and find the frequent sequence set List={A1,A2...An} with length L occurring no less than min_support from the database.

step3: Divide the original sequence database into n subspaces, and mine the frequent sequences with length L+1 respectively,until the result is empty, then stop.

step4: L+1, set the found frequent sequence in the List of step3, then go to step3.

step5: Output all frequent sequences.

step6: Delete the fake trajectories in the anonymous region that do not satisfy the mining result.

Among them,step6 uses three metrics in data mining association analysis to measure frequent series X and pseudo trajectories Y. These three metrics are Support, Confidence and Lift.

Support = P(X &Y), if Support is very low, which means that such a fake trajectory has a low probability of being true for the user and should be deleted.

Confidence = P(X&Y)/ P(X) indicates the probability that the association rule "X→Y" will lead to Y if the precondition X occurs. Indicate how likely the false trajectory Y is to occur in the case of a frequent sequence X that conforms to the user's motion logic. If the confidence level is too small, the probability of the true occurrence of the false trajectory Y is small, it should be deleted.

Lift = P(X&Y) /P(X)/ P(Y). Lift reflects the correlation between X and Y. Lift > 1 and higher indicates higher positive correlation, Lift < 1 and lower indicates higher negative correlation, Lift = 1 means X and Y are independent of each other, i.e. no correlation, so the fake trajectory with Lift <= 1 does not satisfy the user motion logic and should be deleted.

V. Experimental Analysis

5.1 Experimental setup

The experimental data set was collected from the Gowalla data set[20], the Gowalla data set collected 36001959 logins from 319063 users at 2844076 locations. It has the advantages of authentic data, semantic integrity and sufficient data volume.

In order to verify the effectiveness of the algorithm in this paper, after the fake trajectory constructed according to k-anonymity satisfies the spatio-temporal association constraint, this study mines the preprocessed data set by PrefixSpan algorithm to get the frequent trajectory of users, and converts the sequence number pid of user check-in locations in the frequent trajectory into poi with map semantic information, and finally obtains the protection model that can resist the map point background knowledge attack by comparing the fake trajectory with the real frequent trajectory of users.

The above algorithm is programmed in Python language and runs on a computer with Intel(R) Core(TM) i7-9750H CPU@2.60GHz and 8GB main memory on Microsoft Windows 10 Home.

5.2 Analysis of experimental results

5.2.1 Data pre-processing

The data set is divided into 6 files containing location type numbers and semantics, about 3600w user logins, about 442w user relationship mappings, about 272w location information, about 12w supplemental locations information and about 21w users' account information, where the time span between user logins and next logins is shown in Figure.6. The difference in the frequency of login records of different users is shown in the Figure.7 of the figure, and the number of people who have logged in 20 times or less is about 15.6w, accounting for about 49.02%.

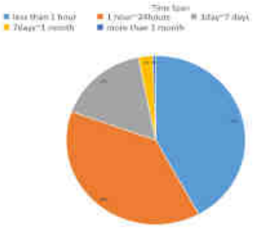


Fig.6 The time span between user login and next login;

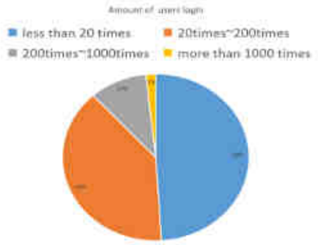
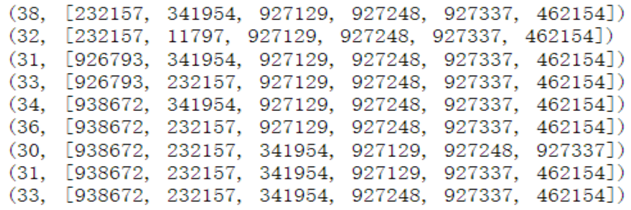


Fig.7 The login frequency of different users.

After desensitizing the dataset, users and their check-in points are obtained, and at least more than three check-in points of a user are merged to form a trajectory, resulting in a usable data set of about 105.5w items without redundant information. And this data set was further divided according to map semantics and latitude and longitude to obtain a total of 30707 trajectories located in New York State.

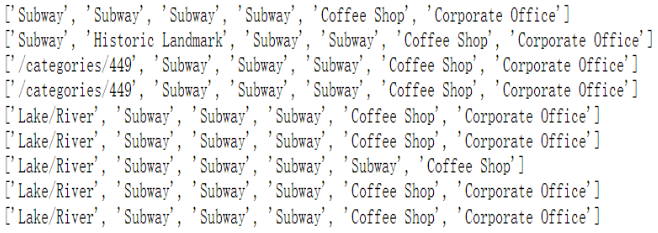
5.2.2 Semantic information conversion of map points

Since the user's check-in point in the data set is a pid representing the map, the pid needs to be converted to a poi with map semantic information. The gowalla_category_structure.json file containing the location type number and semantics is reorganized and written to get the simplified information of location id and semantics, and the pid and semantics are combined to get map.csv of about 272w items, so the pid can be converted to the corresponding poi.



```
(38, [232157, 341954, 927129, 927248, 927337, 462154])
(32, [232157, 11797, 927129, 927248, 927337, 462154])
(31, [926793, 341954, 927129, 927248, 927337, 462154])
(33, [926793, 232157, 927129, 927248, 927337, 462154])
(34, [938672, 341954, 927129, 927248, 927337, 462154])
(36, [938672, 232157, 927129, 927248, 927337, 462154])
(30, [938672, 232157, 341954, 927129, 927248, 927337])
(31, [938672, 232157, 341954, 927129, 927337, 462154])
(33, [938672, 232157, 341954, 927248, 927337, 462154])
```

Fig.8 Pid of user check-in location;



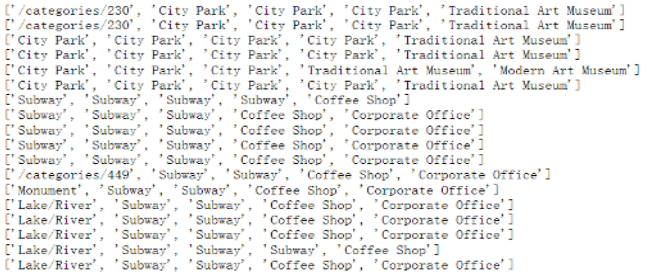
```
['Subway', 'Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Subway', 'Historic Landmark', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['/categories/449', 'Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['/categories/449', 'Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Subway', 'Subway', 'Coffee Shop']
['Lake/River', 'Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
```

Fig.9 Conversion of pid to poi with map semantic information;

5.2.3 Sequence pattern mining analysis

By mining the preprocessed Gowalla data set through the PrefixSpan algorithm,setting the minimum support threshold min_support=40 and each trajectory with no less than five points, the frequent trajectories of users with map point semantics are obtained as Figure x. Each list represents a frequent trajectory of the user, and the result can be observed as "subway" → "subway" → "subway" → "Coffee Shop" → "Corporate Office".

After converting the latitude and longitude of each trajectory point of the frequent sequence to x-axis y-axis coordinates, the trajectory can be reproduced in the two-dimensional plane as shown in Figure.11.



```
['/categories/230', 'City Park', 'City Park', 'City Park', 'Traditional Art Museum']
['/categories/230', 'City Park', 'City Park', 'City Park', 'Traditional Art Museum']
['City Park', 'City Park', 'City Park', 'City Park', 'Traditional Art Museum']
['City Park', 'City Park', 'City Park', 'City Park', 'Traditional Art Museum']
['City Park', 'City Park', 'City Park', 'Traditional Art Museum', 'Modern Art Museum']
['City Park', 'City Park', 'City Park', 'City Park', 'Traditional Art Museum']
['Subway', 'Subway', 'Subway', 'Subway', 'Coffee Shop']
['Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Subway', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['/categories/449', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Monument', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
['Lake/River', 'Subway', 'Subway', 'Subway', 'Coffee Shop']
['Lake/River', 'Subway', 'Subway', 'Coffee Shop', 'Corporate Office']
```

Fig.10 PrefixSpan mining results graph

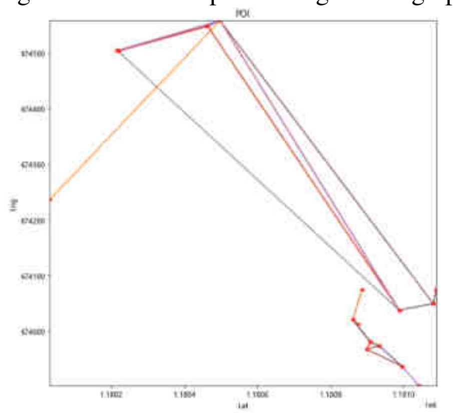


Fig.11 Frequent sequence trajectory recurrence diagram

Where minimum support threshold. The higher the min_support, the more frequent the trajectory is obtained. Min_support and the number of points n of each trajectory can determine the number of frequent trajectories, and the relationship between min_support and the number of trajectories when n=5 is shown in Fig.12.

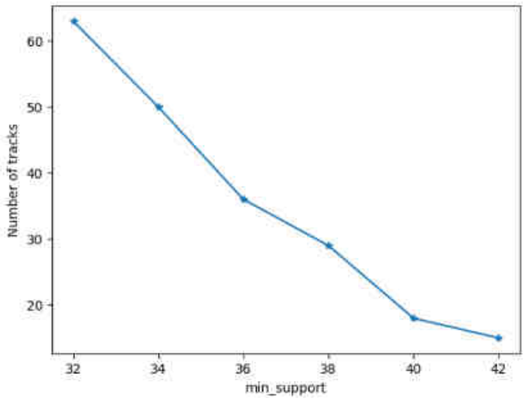


Fig12 The relationship between min_support and the number of trajectories when n=5

VI. Conclusion

The study of LBS privacy protection is of great importance for location-based services to achieve universal, secure, and large-scale adoption. In this paper, we propose a spatio-temporally correlated trajectory privacy protection scheme that using map semantic information combined with spatio-temporal models of trajectories to defend against speculative attacks. The method uses generated fake trajectories to replace service query requests sent from users' real locations in continuous time, and mines users' frequent trajectories by the PrefixSpan algorithm to remove unreasonable fake trajectories, thus reducing the risk of speculative attacks combining map influence and background knowledge and increasing the protection of users' real trajectories. Finally, the effectiveness of the algorithm is verified by simulation experiments.

REFERENCES

[1] Gruteser M , Grunwald D . Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking[C]// Proceedings of the First International Conference on Mobile Systems, Applications, and Services (MobiSys 2003), San Francisco, CA, USA. 2003.

[2] Machanavajjhala A,Gehrke J,Kifer D.L-diversity:privacy beyond k-anonymity[C].Proceedings of the^{22nd} International Conference on Data Engineering(ICDE),2006.24-24.

[3] Zhao P. Key technology research on privacy protection in location services [D].

[4] Xu T, Cai Y. Location anonymity in continuous location-based services. in: Samet H, ed. Proc. of the 15th Annual ACM Int'l Symp. on Advances in Geographic Information Systems. Seattle: ACM, 2007. 1-8. [doi: 10.1145/1341012.1341062]

[5] Montgomery P L. Modular multiplication without trial division [J].Mathematics of Computation,1985,44(170);519-521.

[6] Huang L, Yamane L, Mastsuura K, Sezaki K. Silent Cascade: Enhancing location privacy without communication QoS degradation.
In: Clark JA, ed. Proc. of the 3rd Int'l Conf. on Security in Pervasive Computing. lNCS 3934, Heidelberg: Springer-Verlag, 2006.
165−180. [doi: 10.1007/11734666_13]

[7] Huang LP, Matsuura K, Yamane H, Sezaki K. Enhancing wireless location privacy using silent period. in: Pauly L, ed. Proc. of the
IEEE Wireless Communications and Networking Conf. New Orleans: IEEE, 2005. 1187−1192. [doi: 10.1109/WCNC.2005.
1424677]

[8] Huang LP, Yamane H, Matsuura K, Sezaki K. Towards modeling wireless location privacy. in: Danezis G, ed. Proc. of the 5th Int'l
Workshop on Privacy Enhancing Technology (PET). lNCS 3856, Heidelberg: Springer-Verlag, 2005. 59−77. [doi: 10.1007/
11767831_5].

[9]KIDO H, YANAGISAWA Y, SATOH T. An anonymous communication technique using dummies for location-based services[C]//International Conference on Pervasive Services. IEEE, 2005: 88-97.

[10] KIDO H, YANAGISAWA Y, SATOH T. Protection of location privacy using dummies for location-based services [C]//21st International Conference on Data Engineering Workshops. IEEE, 2005: 1248-1248.

[11] You T H, Peng W C, Lee W C. Protecting moving trajectories with dummies [C]/ / Proc of International Conference on Mobile Data Management. Piscataway, NJ:IEEEPress, 2007:278-282.

[12] 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.

[13] 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.

[14] Li, F. H.,Zhang, C.,Niu, B.,Li, F.,Hua, J. F.,Shi, G. Z. An efficient trajectory privacy protection scheme[J]. journal of communication,2015,36(12):114-123.

[15] Lei Kaiyue. Research on the privacy protection scheme of fake trajectories based on spatio-temporal association [D]. xi'an university of electronic science and technology, 2017.

[16] AgrawlR,SrikantR.Miningsequentialpatterns[C] //Procee-dings of the EleventhInternational Conferenece on Data Engi-neering(ICDE'95).Washington,DC;IEEE
ComputerSociety,1995:3-14;

[17] Srikant R, Agrawal R. Mining Sequential Patterns: Generalizations and Performance Improvements[C]// International Conference on Extending Database Technology: Advances in Data-base Technology. Springer-Verlag, 1996:3-17

[18] Pei J,Han J, Mortazaviasl B, et al. PrefixSpan: Mining Sequential Patterns Efficiently by Prefix-Projected Pattern Growth[J].2001:215--224.

[19] Han J, Pei J, Hsu M C, et al. FreeSpan: frequent pattern-projected sequential pattern mining[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.ACM,2000:355-359.

[20] Retrieved from https://www.yongliu.org/datasets/